Vol. 9 No. 3
June 2025

Opportunities Identification & Problem Solutions

INTERNATIONAL JOURNAL OF SYSTEMATIC INNOVATION

ACCSCIENCE PUBLISHING

ISSN(Print): 2077-7973 ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202506_9(3)

https://ijosi.org/index.php/IJOSI

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202506_9(3)

The International Journal of Systematic Innovation

Editorial Office:

- The International Journal of Systematic Innovation
- 6F, # 352, Sec. 2, Guanfu Rd, Hsinchu, Taiwan, R.O.C., 30071
- e-mail:editor@systematic-innovation.org
  ijosi@systematic-innovation.org
- web site: http://www.IJoSI.org
- Tel: +886-3572-3200

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202506_9(3)

INTERNATIONAL JOURNAL OF SYSTEMATIC INNOVATION

CONTENTS

JUNE 2025 VOLUME 9 ISSUE 3

FULL PAPERS

- Innovative solutions for convolutional neural network performance: A TRIZ-based reverse engineering approach

.... M. Cosgun, K. Altun. 1–7

- A Mayan calendar-inspired cyclical TRIZ approach: Enhancing systematic innovation and long-term problem-solving

.... K. Altun. 8–19

- Hybrid intelligence model for traffic management in intelligent transportation systems

.... I. Appaji, P. Raviraj. 20–30

- New product development process: A conceptual framework for automobile industries

.... B. G. Shinde, S. B. Sanap, S. S. Pawar, V. D. Wakchaure. 31–74

- Handwriting match and AI content detection

.... H. Panigrahi, S. Naidu, A. Pandey, P. Shaikh, A.K. Tripathy. 75–95

- Enhancing healthcare efficiency with artificial intelligence: Benefits, challenges, and the future of clinical practice

.... R.A. Awashreh. 96–108

- Ensemble learning for enhanced brain tumor diagnosis: A new approach for early detection

.... N.F. Othman, S.W. Kareem. 109–123

- Sentiment of the largest state, first mover, and largest private banks digital performance in Indonesia: Strategic perspective

.... T.R.I. Putra, M.I. Fajri. 124–141

Innovative solutions for convolutional neural network performance: A TRIZ-based reverse engineering approach

Merve Cosgun, Koray Altun*

Department of Industrial Engineering, Bursa Technical University, Bursa, Turkey

*Corresponding author E-mail: koray.altun@btu.edu.tr

(Received 23 September 2024; Final version received 31 December 2024; Accepted 14 January 2025)

Abstract

Convolutional neural networks (CNNs) are widely used in computer vision for tasks like image classification and detection. These models work well when the number of image classes is small, but as the number of classes increases, accuracy tends to drop due to overfitting. There are several methods to address this issue, such as data augmentation, preprocessing, class weighting, transfer learning, and adjusting technical parameters. This study introduces a novel approach utilizing the theory of inventive problem-solving (TRIZ) methodology to systematically analyze and enhance these existing methods. Using reverse engineering, we deconstructed current solutions and aligned them with TRIZ principles to propose more innovative and effective approaches for improving CNN performance. The results show that TRIZ provides a structured and creative framework for solving accuracy decline issues in CNN models, offering the potential for broader applications in other machine learning architectures.

Keywords: Convolutional Neural Network, Image Classification, Reverse Engineering, Theory of Inventive Problem Solving

1. Introduction

Convolutional neural networks (CNNs) are widely used in computer vision tasks like image classification due to their high performance in learning patterns from data (LeCun et al., 2015). However, as the number of classes in a dataset increases, model accuracy tends to decrease. This drop in accuracy is often caused by overfitting, where the model becomes too specialized in the training data and struggles to generalize to new, unseen data (Krizhevsky et al., 2017). Solving this problem is critical for applications that require accurate classification across many classes, such as medical diagnosis, autonomous driving, and facial recognition systems (Guo et al., 2019; Litjens et al., 2017).

Several approaches have been proposed to address this issue, including data augmentation, preprocessing techniques, and transfer learning (Shorten & Khoshgoftaar, 2019).

"Data augmentation," for example, increases the diversity of the training data by applying transformations such as rotation, scaling, and flipping to existing images. This approach has proven effective in applications such as medical imaging, where obtaining large datasets is difficult (Perez & Wang, 2017).

Traditional augmentation techniques, such as rotation, flipping, and cropping, have been widely used in image classification and segmentation tasks. However, more innovative strategies are continuously being developed to enhance augmentation effectiveness. For instance, Alomar et al. (2023) introduced a new "random local rotation" technique, which improves data diversity while minimizing the boundary artifacts commonly caused by traditional rotation methods. These advancements in augmentation help CNNs generalize better, especially in tasks with limited data availability.

Several approaches have been proposed to improve data augmentation, particularly with automated techniques. Automated data augmentation (AutoDA) methods have been increasingly studied, as they can automatically discover optimal augmentation strategies tailored to specific datasets. For example, a recent comprehensive survey by Yang et al. (2023)

categorizes existing AutoDA methods, highlighting their efficiency in improving image classification tasks by reducing manual intervention and increasing model performance through learned augmentation policies. These AutoDA methods present a promising direction for enhancing data diversity and model generalization while reducing human error in the augmentation process.

"Preprocessing" techniques, such as normalization and cropping, are used to refine input data before training, ensuring that the model receives consistent and high-quality information (Kamnitsas et al., 2017).

Image preprocessing is essential for improving CNN performance by reducing noise and enhancing data quality. Techniques such as noise reduction, histogram equalization, and image hashing have shown notable accuracy improvements in facial recognition tasks, with gains of over 4% in some cases (Tribuana et al., 2024).

"Transfer learning," on the other hand, leverages knowledge from pre-trained models to enhance performance in specific tasks, particularly when labeled data is limited, allowing models pre-trained on large datasets to be fine-tuned for specific tasks (Atasever et al., 2023; Pan & Yang, 2010; Tan et al., 2018).

More innovative approaches are needed to tackle the root of the problem in a systematic way. This is where the theory of inventive problem-solving (TRIZ) methodology comes into play. Developed in engineering, TRIZ offers a structured approach to identifying contradictions and proposing creative solutions based on inventive principles. The method has proven effective in solving technical problems across various fields, yet its application in machine learning, particularly for improving CNN performance, remains underexplored.

The aim of this study is to address the issue of decreasing accuracy in CNN models as the number of classes increases. By applying the TRIZ methodology, we aim to develop innovative and systematic solutions to improve CNN performance. Through reverse engineering, we analyzed current techniques, such as data augmentation, preprocessing, class weighting, and transfer learning, and aligned these with TRIZ principles to propose more effective solutions.

Recent advancements in artificial intelligence (AI) have increasingly focused on overcoming the challenges posed by small datasets, which are prevalent in fields where data collection is restricted. The study by Brad and Brad (2023) explores the use of TRIZ methodologies to address this issue, offering inventive strategies for enhancing AI model performance under these constraints. While their research underscores the adaptability of TRIZ in optimizing AI with limited data, our study extends the application of TRIZ to tackle accuracy degradation in CNN models as class counts increase. This divergence highlights the broad applicability of TRIZ principles across different AI challenges.

The contribution of this study lies in the novel application of TRIZ methodology to the field of machine learning, specifically in solving the overfitting problem in CNNs with increasing class counts. TRIZ, traditionally applied in engineering, provides a structured approach to identifying contradictions and generating innovative solutions, which has not been widely explored in CNN performance issues.

The remainder of this paper is organized as follows: Section 2 details the methodology; Section 3 presents the implementation; Section 4 provides the discussion; and Section 5 concludes the paper, summarizing the findings and offering directions for future research.

2. Methodology

The TRIZ methodology provides a systematic approach to innovation that involves analyzing and categorizing thousands of patents to uncover universal principles of invention. Central to TRIZ is its distinctive method for tackling technical problems by converting specific situations into broader, conceptual challenges. This process requires breaking down the problem into its core elements and then applying TRIZ's set of inventive principles and proven solutions to devise a conceptual solution (Gadd, 2011). The TRIZ methodology is traditionally outlined as:

(i) Specific problem: The initial stage, is where the specific technical problem is identified.
(ii) Conceptual problem (39 parameters): The problem is generalized to a conceptual level by identifying relevant engineering parameters.
(iii) Conceptual solution (40 principles): Solutions are developed based on the 40 inventive principles of TRIZ.
(iv) Specific solution: The conceptual solution is then translated back into a specific practical solution for the initial problem.

The process starts by breaking down a real-world problem into a conceptual format. This simplification allows for aligning the problem with TRIZ's effective solutions, which rely on structured principles instead of random brainstorming ideas. After finding a conceptual solution, it is then polished and converted into a practical solution that specifically addresses the initial problem.

The TRIZ methodology focuses on transforming specific, real-world problems into conceptual challenges, which can then be matched with systematic solutions. Sheu & Lee (2011) proposed a structured process for innovation that incorporates TRIZ principles to help break down complex problems and develop creative solutions. Their work highlights the importance of using TRIZ tools, such as the contradiction matrix and inventive principles, to ensure that the problem-solving process is both organized and comprehensive. By following their systematic process, innovators can consistently arrive at effective solutions for technical challenges.

In this study, we approach TRIZ from a "reverse engineering" perspective, applying its principles not just to generate new solutions but also to reinterpret and reanalyze existing solutions found in the literature. By doing so, we aim to provide a more comprehensive framework for problem-solving that bridges past solutions with inventive methodologies. The methodology proposed in this study is as follows;

(i) Specific solution: Start from an existing solution or product.
(ii) Conceptual solution analysis (40 principles): Deconstruct the solution to understand how TRIZ principles are or can be applied.
(iii) Identification of contradictions (39 parameters): Identify any existing or potential contradictions that the current solution might be causing or not addressing.
(iv) Revised problem statement: Define or redefine problems based on insights gained from the analysis and contradiction identification.

Reverse engineering within the TRIZ framework involves deconstructing existing technical solutions to understand their core principles and then matching these with TRIZ's 40 inventive principles and the contradiction matrix. This approach allows us to assess how well these existing solutions align with TRIZ's systematic process and identify opportunities for improvement or further development. For instance, a solution that addresses one specific technical contradiction may have untapped potential for solving additional contradictions when viewed through the lens of TRIZ.

Fig. 1 demonstrates how the traditional TRIZ process is a forward-thinking approach, starting from problem identification and moving toward a solution. In contrast, the reverse engineering process begins with an existing solution, analyzing it through the TRIZ lens to uncover deeper insights and potentially redefine the problem or improve the solution.

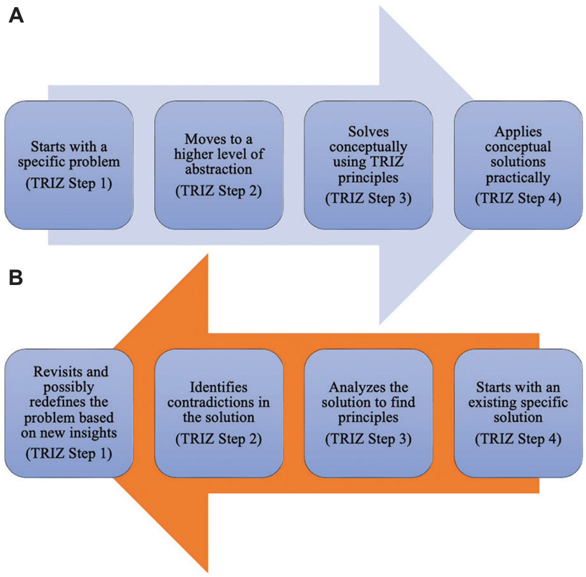


Fig. 1. A comparison of the flows of a (A) TRIZ process and a (B) TRIZ-based reverse engineering process
Abbreviation: TRIZ: Theory of inventive problem solving.

3. Implementation

In this section, we take a different path from the usual forward-thinking problem-solving approaches. This is where reverse engineering comes into play. Reverse engineering, in essence, involves working backward to deconstruct an existing solution to understand its foundational principles. In doing so, we can uncover hidden opportunities for improvement or discover alternative solutions that may not have been obvious initially. Instead of building a solution from scratch, we analyze what already exists, break it down, and explore how it aligns with TRIZ's 40 inventive principles and contradiction matrix.

By reversing the usual flow of thought, we can gain deeper insights into how existing solutions operate and how they can be enhanced or adapted for broader applications. This method is particularly useful in complex problems, such as improving the accuracy of CNN models, where conventional methods may overlook indirect contradictions or potential improvements that TRIZ can highlight.

3.1. TRIZ Step 4: Identifying Specific Solutions

In TRIZ step 4, we focus on identifying concrete, practical solutions to address the issue of decreasing accuracy in CNN models as the number of image classes increases. From a reverse engineering perspective, we analyze how existing methods have been applied to similar problems and how TRIZ principles can guide the improvement of these methods.

One effective solution is "data augmentation," which involves expanding the dataset by generating new training examples through techniques such as rotating, flipping, scaling, cropping, and adding noise. We can see that "Principle 20: Continuity of Useful Action" fits well here, as the method continuously provides useful variations of the data that help the model learn better. “Principle 15: Dynamics” also applies because the transformations increase the adaptability of the model to diverse input scenarios.

Similarly, “dimensional adjustment” is another common preprocessing technique that improves consistency across classes by resizing or normalizing images. This approach aligns with “Principle 17: Another Dimension,” which suggests modifying or using different dimensions to solve a problem. By ensuring that images are of uniform size, we reduce the variability in the input data, which enhances the model’s ability to make accurate predictions.

In cases where there is a class imbalance, “class weighting” can improve accuracy by giving more importance to underrepresented classes. This method reflects “Principle 35: Parameter Changes,” which involves adjusting key parameters to achieve the desired result. Assigning weights to classes based on their representation helps the model treat all classes fairly, reducing bias and improving performance.

Finally, “transfer learning” offers a powerful way to reuse pre-trained models on new tasks, particularly when data is scarce. This approach aligns with “Principle 24: Intermediary,” which suggests using an intermediary to assist with problem-solving. The intermediary model speeds up learning and improves performance, especially when the new task shares similarities with the original one.

3.2. TRIZ Step 3: Identifying Conceptual Solutions

In TRIZ step 3, the goal is to select a conceptual solution based on TRIZ principles that fits the problem. Using a reverse engineering approach, we look back at how similar problems have been solved in the past and apply those insights to find conceptual solutions for the current issue of CNN accuracy decline. Table 1 lists all the candidate principles. Table 2 shows the candidate principles that match the specific solutions presented in step 4.

The candidate principles listed in Table 1 were selected through a structured TRIZ-based process. Initially, principles annotated with the letter “a” (i.e., 15, 17, 20, 24, and 35) were identified based on their relevance to the solution alternatives in Table 2. These principles were mapped to engineering characteristics using a contradiction matrix, highlighting key pairs (Table 3).

For “data augmentation,” we selected Principles 15 and 20 because these principles describe how system flexibility and continuous beneficial actions can enhance performance.

When considering “dimensional adjustment,” Principle 17 stood out as the most appropriate. Adjusting the dimensions of input data can help reduce the computational complexity and make the data more uniform, thereby improving the model’s ability to classify images.

In the case of “class weighting,” Principle 35 was chosen because it directly addresses the problem of imbalanced datasets. Modifying the class weights allows the model to handle rare classes more effectively, which is essential for improving overall accuracy.

Lastly, for “transfer learning,” Principle 24 is highly relevant. Transfer learning enables models to leverage previously learned knowledge, significantly reducing training time and improving accuracy.

Table 1. Candidate principles

No.	Principle	Definition
1	2	Taking out
2	3	Local quality
3	13	The other way round
4	15[a]	Dynamics
5	17[a]	Another dimension
6	18	Mechanical vibration
7	20[a]	Continuity of useful action
8	23	Feedback
9	24[a]	Intermediary
10	27	Cheap short-living objects
11	28	Mechanic substitution
12	29	Pneumatics and hydraulics
13	30	Flexible shells and thin films
14	33	Homogeneity
15	35[a]	Parameter changes
16	36	Phase transitions
17	37	Thermal expansion

Note: [a]Candidate principles.

Table 2. Candidate principles and solution matching

No.	Solution alternatives	TRIZ candidate principle
1	Data augmentation	15, 20
2	Preprocessing	17
3	Assigning class weights	35
4	Transfer learning	24

Abbreviation: TRIZ: Theory of inventive problem-solving.

3.3. TRIZ Step 2: Defining the Conceptual Problem

In TRIZ, step 2 involves identifying the engineering characteristics that are in conflict, resulting in a technical contradiction. Engineering

Table 3. Candidate pairs

Candidates	28	29	35	39
10	10, 23, 24[a], 35[a]	28, 29, 36, 37	15a, 17[a], 18, 20[a]	3, 28, 35[a], 37
26	2, 13, 28	30, 33	3, 15a, 29	3, 13, 27, 29

Note: [a]Candidate principles.

characteristics refer to specific technical parameters or features of a system. A contradiction occurs when improving one characteristic negatively impacts another. In our case, increasing the number of image classes in a CNN model may lead to a decrease in accuracy—these are two conflicting engineering characteristics.

Based on the guidance from the candidate principles identified in TRIZ step 3, we selected a pair of engineering characteristics. Table 4 outlines the engineering characteristics obtained from the normal analysis of the problem. Two major contradictions were identified: one representing the increase in the number of classes (related to engineering characteristic 26, "quantity of substance") and another representing the decrease in model accuracy (related to engineering characteristic 35, "adaptability or versatility"). These contradictions must be addressed to improve the model's performance.

Table 3 shows the conceptual solution combinations derived from these contradictions. The contradiction pair 26 and 35, guided by Principles 15, 17, and 20, was determined to be the most relevant for this study. Engineering characteristic 26 corresponds to the challenge of managing an increasing number of classes while engineering characteristic 35 relates to the issue of decreasing accuracy. These contradiction reflects the balance we seek between increasing model capacity and maintaining high accuracy.

It should be highlighted here that, unlike the traditional TRIZ methodology that starts with engineering contradictions, our approach began with identifying the inventive principles, as shown in Table 3. For instance, the contradiction pair (10, 35) was highlighted due to its strong alignment with Principles 15, 17, and 20.

The inclusion of "Force" was guided by its connection to critical principles such as 15 (Dynamics). While characteristics like "Strength" were considered, no directly related principles were identified, which justified its exclusion.

By identifying the engineering characteristics and their associated contradictions, we can apply TRIZ principles to systematically resolve these conflicts. For example, "Principle 15: Dynamics" helps address flexibility in handling different classes, while "Principle 17: Another Dimension" suggests altering how data are processed to maintain accuracy despite increased complexity.

Table 4. Candidate contradictions

No.	Eng. char.	Definition
1	10	Force (intensity)
2	26	Quantity of substance
3	28	Measurement accuracy
4	29	Manufacturing precision
5	35	Adaptability or versatility
6	39	Productivity

Abbreviation: Eng. Char.: Engineering characteristic.

In comparison, Brad and Brad (2023) explored contradictions between data quantity and system performance (contradiction pairs 26 & 28) in their study. Their analysis of contradictions between system complexity and performance (contradiction pair 36 & 28) further illustrates how TRIZ can be used to systematically resolve technical challenges by focusing on the core engineering characteristics at play.

4. Discussion

In this study, we applied a TRIZ-based reverse engineering approach to address the problem of decreasing accuracy in CNN models as the number of image classes increases. This methodology enabled a systematic examination of existing solutions, revealing opportunities for innovation by aligning these solutions with TRIZ principles.

We further discuss the insights gained through this approach, with a particular focus on the application of transfer learning and its reinterpretation within the TRIZ framework.

4.1. General Insights

The TRIZ-based reverse engineering shifts the focus from conventional problem-solving to a structured analysis of existing solutions. This approach provides a systematic way to identify and resolve contradictions inherent in CNN models, such as the trade-off between model complexity and accuracy. By deconstructing existing methods, such as data augmentation, preprocessing, and class weighting, we identified their alignment with specific TRIZ principles and proposed refinements that address overlooked challenges.

For instance, data augmentation aligns with "Principle 20: Continuity of Useful Action," as it generates continuous variations in the training data, enabling better model generalization. Similarly, preprocessing techniques, such as dimensional adjustments, align with "Principle 17: Another Dimension," addressing variability in input data to improve consistency and prediction accuracy. These connections demonstrate how TRIZ principles provide a creative and structured framework to optimize existing solutions.

4.2. Transfer Learning and Intermediary Principle

We acknowledge that transfer learning is a widely recognized method for improving CNN performance, particularly in cases with limited labeled data. However, its reinterpretation through TRIZ's "Principle 24: Intermediary" offers new perspectives and applications.

Transfer learning is typically seen as a way to reuse pre-trained models for specific tasks. Within the TRIZ framework, we redefine it as an intermediary that bridges two conflicting needs: (i) the scarcity of labeled data in new tasks, and (ii) the requirement for high accuracy in performance. This reinterpretation positions transfer learning not just as a static tool but also as a dynamic mediator that facilitates the resolution of these contradictions. By emphasizing its role as an intermediary, TRIZ provides a structured perspective for enhancing the applicability of transfer learning.

Viewing transfer learning through the lens of TRIZ enables a broader application of this technique. For example, TRIZ principles encourage creative extensions of pre-trained models to address additional challenges, such as:

(i) Reducing bias in class imbalance: By systematically reweighting features learned by the intermediary model, we can mitigate biases present in underrepresented classes.
(ii) Overcoming noise in preprocessing: Transfer learning can serve as a filter to preprocess noisy data more effectively, guided by TRIZ principles, such as "Principle 35: Parameter Changes."

These reinterpretations highlight how TRIZ inspires creative problem-solving by encouraging researchers to think beyond the conventional uses of established techniques.

5. Concluding Remarks

The TRIZ-based reverse engineering approach, proposed in this study, offers a structured framework for analyzing existing solutions, uncovering contradictions, and proposing innovative refinements. Unlike traditional forward-thinking methods, reverse engineering starts with what already exists, systematically deconstructs these solutions, and applies TRIZ principles to identify untapped opportunities. This methodology provides two major advantages:

(i) Systematic problem analysis: Traditional problem-solving approaches often focus on developing new solutions from scratch. In contrast, reverse engineering allows us to examine existing solutions critically, revealing their underlying contradictions or limitations. By aligning these with TRIZ's 40 inventive principles and contradiction matrix, we can systematically identify areas for improvement or innovation.
(ii) Maximizing existing knowledge: This approach avoids reinventing the wheel by leveraging what is already available. It enables researchers to reinterpret existing solutions through the lens of TRIZ principles, uncovering new opportunities for optimization or broader application.

This study implemented the TRIZ-based reverse engineering approach to address the problem of accuracy decline in CNN models as the number of image classes increases. By systematically deconstructing existing solutions and analyzing contradictions between key engineering characteristics, we were able to define specific TRIZ principles to propose practical solutions.

The use of data augmentation, dimensional adjustment, class weighting, and transfer learning proved to be effective strategies for improving model accuracy. These methods, when viewed through the TRIZ framework, provided a structured approach to solving the technical contradictions between class quantity and accuracy. The reverse engineering perspective further enhanced this process by allowing us to identify hidden opportunities for improvement within the existing solutions.

While this study focused on CNN models, the principles of TRIZ could be applied to other machine learning architectures, such as recurrent neural networks, transformers, or generative adversarial networks. Future research could investigate how TRIZ can be adapted to solve contradictions in these more advanced architectures.

References

Alomar, K., Aysel, H.I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. *Journal Imaging*, 9, 46.

Atasever, S., Azginoğlu, N., Terzi, D.S., & Terzi. (2023). A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning. *Clinical Imaging*, 94, 18–41.

Brad, S., & Brad, E. (2023). Using TRIZ to handle small datasets in artificial intelligence. In:

Acta Technica Napocensis Series Applied Mathematics, Mechanics, and Engineering, Technical University of Cluj-Napoca, Romania, p66.

Gadd, K. (2011). *TRIZ for Engineers: Enabling Inventive Problem Solving*. John Wiley and Sons, United States.

Guo, Y., Zhang, L., Hu, Y., He, X., & Gao, J. (2019). A survey on deep learning based face recognition. *Computer Vision and Image Understanding*, 189, 102805.

Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., *et al*. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36, 61–78.

Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., *et al*. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.

Pan, S.J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.

Perez, L., & Wang, J. (2017). *The Effectiveness of Data Augmentation in Image Classification Using Deep Learning*; [arXiv preprint].

Sheu, D.D., & Lee, H.K. (2011). A proposed process for systematic innovation. *International Journal of Production Research*, 49(3), 847–868.

Shorten, C., & Khoshgoftaar, T.M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. In: *International Conference on Artificial Neural Networks*. Springer, Cham, p270–279.

Tribuana, D., & Arda, A.L. (2024). Image preprocessing approaches toward better learning performance with CNN. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 8 (1), 1–9.

Yang, Z., Sinnott, R.O., Bailey, J., & Ke, Q. (2023). A survey of automated data augmentation algorithms for deep learning-based image classification tasks. *Knowledge and Information Systems*, 65, 2805–2861.

AUTHOR BIOGRAPHY

Merve Cosgun completed her undergraduate degree in Industrial Engineering at Uludağ University. She is currently pursuing her Master's degree in Industrial Engineering at Bursa Technical University. Merve works as an industrial engineer at a private company and has a keen interest in the theory and applications of TRIZ.

Koray Altun is an assistant professor of Industrial Engineering at TU Bursa and a research collaborator at GLORAD. He is also the founder and manager of a startup focused on digital innovation and software consultancy. He holds a Ph.D and a BSc. Degree in Industrial Engineering from Gaziantep University and Erciyes University, respectively. He has published papers in well-respected journals and has led consulting engagements in reputable companies with a focus on R&D, technology, and innovation management. His recent research interests include the Innovation Excellence Model, Systematic Innovation, Technology and Innovation Management, and R&D Management.

DOI: 10.6977/IJoSI.202506_9(3).0002
K. Altun/Int. J. Systematic Innovation, 9(3), 8-19 (2025)

A Mayan calendar-inspired cyclical TRIZ approach: Enhancing systematic innovation and long-term problem-solving

Koray Altun[1,2]*

[1]Department of Industrial Engineering, Bursa Technical University, Bursa, Bursa, Turkey

[2]ARDiMER Digital Innovation Limited, Bursa, Bursa, Turkey

*Corresponding author E-mail: koray.altun@btu.edu.tr

(Received 30 October 2024; Final version received 01 February 2025; Accepted 25 March 2025)

Abstract

This paper introduces the Mayan calendar-inspired cyclical theory of inventive problem solving (TRIZ) model, an innovative approach to systematic innovation (SI) that integrates the seven TRIZ pillars into a structured model consisting of "Tzolk'in" (short-term, adaptation), "Haab" (mid-term, harmonization), and "Long Count" (long-term, transformation) cycles. Unlike traditional linear innovation models, this cyclical model enables continuous adaptation, iterative refinement, and sustainable evolution. Each cycle addresses a different level of complexity: The adaptation cycle focuses on rapid, low-cost improvements using available resources. The harmonization cycle resolves deep-rooted contradictions to enhance system functionality. The transformation cycle drives strategic evolution by integrating intelligence and automation. This approach is validated through its alignment with trends of engineering system evolution, demonstrating that innovation naturally progresses through these phases. The model's practical applicability is illustrated through case studies on coffee machine design and automotive seat design, showing how short-term enhancements, mid-term optimization, and long-term transformation collectively contribute to sustainable evolution. By bridging systematic problem-solving with iterative adaptation, the cyclical TRIZ model provides a versatile and scalable SI model for industries seeking to achieve both immediate efficiency gains and long-term innovation resilience.

Keywords: Cyclical TRIZ Model, Mayan Calendar, Systematic Innovation, TRIZ

1. Introduction

Innovation is not just a straightforward journey from point A to point B. It often involves cycles of learning, feedback, and adaptation. This cyclical nature of progress can be seen in ancient philosophies, like the Mayan calendar (Fig. 1), which views time as a repeating cycle rather than a straight line.

The Mayan calendar consists of three main cycles: "Tzolk'in" (short-term), "Haab" (mid-term), and "Long Count" (long-term). Each cycle represents a layer of time that builds upon the previous one, emphasizing the importance of iteration, balance, and continuous growth. This concept of cyclicality not only promotes holistic understanding but also encourages adaptation over time, allowing for flexibility in the face of changing circumstances.

This idea of cycles aligns well with a systematic innovation process (SIP). Similarly, systematic innovation (SI) is not a purely linear process. While it provides a structured approach to problem identification, solution generation, and solution implementation, its effectiveness can be further enhanced by integrating cyclical elements.

Various SI models provide structured pathways to guide the process, typically moving from problem identification to solution implementation. These models often incorporate the theory of inventive problem-solving (TRIZ) tools to enhance creativity, improve idea generation, or manage knowledge transformation. While some focus on incremental problem-solving, others aim for strategic planning or interdisciplinary collaboration. However, a common limitation among these models is their linear

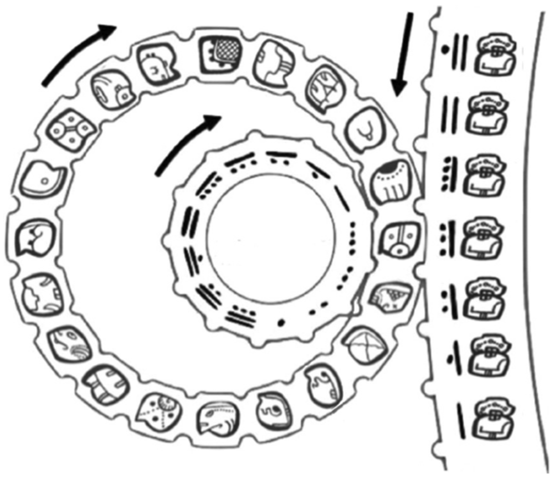


Fig. 1. A representation of the Mayan calendar cycle (adapted from Chanier, 2018)

progression, which lacks a mechanism for continuous feedback and adaptation.

A linear model may not fully support dynamic adaptation or long-term problem-solving. This study seeks to create a more adaptive SIP that can sustain both short-term problem-solving and long-term strategic growth. By integrating a cyclical approach inspired by the Mayan calendar, the TRIZ approach can become more iterative, allowing for continuous feedback and adaptation, which is essential for both short-term and long-term solutions.

The remainder of this paper is organized as follows: Section 2 reviews existing TRIZ-based SIPs. Section 3 presents the new cyclical TRIZ model, detailing its iterative and adaptive structure. Section 4 illustrates the model's application through a case study, showcasing its effectiveness in evolving scenarios. Sections 5-8 discuss the results, comparing the new model to traditional approaches and offering conclusions.

2. Literature Review

SI distinguishes itself from empirical, trial-and-error methods by providing a clear, step-by-step process that guides problem identification, solution generation, and implementation (Sheu & Lee, 2011). TRIZ is built on systematic principles that help identify contradictions in a system and generate inventive solutions. SI reduces the uncertainty and trial-and-error nature of traditional innovation by incorporating well-defined tools and strategies, often integrating principles from TRIZ.

The theory of inventive problem-solving is a structured approach that aims to identify and solve contradictions in design and development. TRIZ traditionally adopts a linear problem-solving approach, focusing on one contradiction at a time. Although TRIZ tools are highly effective, applying the right tool at the right stage of the innovation process can be challenging (Ilevbare et al., 2013). This difficulty often arises due to the complexity and diversity of available TRIZ tools, as well as the need to adapt them to specific problem contexts.

Several SI models have been developed to provide structured pathways for innovation. These models aim to bridge the gap between problem identification and solution implementation, incorporating various tools, including TRIZ elements.

Mann's (2007) model, a systematic creativity process, emphasizes a clear sequence of steps: "Define," "Select Tool," "Generate Solutions," and "Evaluate." The approach integrates TRIZ tools to enhance creativity in problem-solving but primarily focuses on addressing specific problems rather than broader opportunities or strategic adaptation.

The W-model (Brandenburg, 2002) covers a continuous cycle of innovation from goal setting to implementation planning. It follows the stages of future analysis, idea generation, idea evaluation, concept detailing, and implementation planning. While effective in strategic planning, it often lacks practical tools for execution and does not emphasize feedback loops, making it less adaptable to evolving challenges.

The innovation value chain (Hansen & Birkinshaw, 2007) divides the innovation process into three distinct stages: idea generation, idea conversion, and idea diffusion. It focuses on moving ideas from concept to market but lacks specific tools for systematic problem-solving, making it more conceptual than operational.

Roper et al.'s (2008) model outlines innovation as a transformation of knowledge into business value, with stages of knowledge sourcing, transformation, and exploitation. While it emphasizes knowledge management as a driver of innovation, it lacks the structured tools necessary for resolving contradictions or implementing solutions systematically.

Sheu & Lee (2011) proposed a SIP that integrates both TRIZ and non-TRIZ tools to facilitate innovation across different phases. The SIP consists of structured stages that guide innovation from opportunity identification to solution generation, allowing for effective exploitation of developed technologies.

The ACE framework (Zhan et al., 2017) facilitates product innovation by shortening time-to-market, accelerating the understanding of customer needs, and reducing costs. The model uses big data analytics to accelerate innovation processes, enhance customer connection, and create an innovation ecosystem. It emphasizes the importance of faster innovation cycles and stronger customer feedback loops, making the product development process more dynamic and flexible. While this approach effectively uses big data to speed up innovation, it follows a linear process without built-in mechanisms for continuous adaptation or feedback.

Kruger et al. (2019) proposed a SI model that integrates TRIZ and creative problem-solving techniques to enhance innovation management. This model considers psychological factors, such as overcoming psychological inertia, to support creative thinking.

Sun et al. (2020) proposed a SIP specifically oriented toward interdisciplinary research. Their approach aims to solve complex problems by integrating insights from multiple disciplines, using TRIZ tools, the general theory of powerful thinking (OTSM), and patent databases. This process emphasizes both incremental and disruptive innovation, leveraging interdisciplinary collaboration to enhance creativity and generate high-quality solutions.

The radical problem-solving framework (Wang et al., 2024) aims to go beyond the limits of existing design methods to achieve more radical innovation. This model leverages TRIZ tools to drive more radical innovation, addressing some of the challenges associated with traditional TRIZ applications. While it provides a clearer guide for using TRIZ tools effectively, it follows a linear process that lacks built-in adaptation and continuous feedback.

Mann (2023) also proposed a SIP that applies TRIZ principles to chaotic situations, particularly in emergency response scenarios. The approach is based on TRIZ's concept, the "someone, somewhere, has already solved your problem," aiming to identify the most suitable solutions under unpredictable conditions. This process integrates TRIZ with the OODA Loop (a four-step process: observe, orient, decide, and act), suggesting that this combination allows for faster and more effective decision-making in complex environments. This model demonstrates the potential of TRIZ in managing complexity and asymmetric threats.

Existing SI models in the literature typically emphasize linear or mid-term problem-solving approaches, focusing on specific phases such as idea generation, evaluation, and implementation. While some models incorporate TRIZ tools, they often lack mechanisms for long-term adaptation and sustainability. In this study, the proposed Mayan calendar-inspired cyclical TRIZ approach addresses this gap by integrating short-term, mid-term, and long-term cycles, using the seven pillars of TRIZ (Sheu et al., 2020) to continuously navigate and refine solutions over time. This approach not only emphasizes long-term problem-solving but also aligns innovation strategies with enduring outcomes, ensuring more resilience in dynamic environments. It offers a more holistic approach by embedding iterative feedback and adaptability within each phase, bridging the gap between structured inventive problem-solving and sustainable innovation.

3. Proposed Model

According to Sheu et al. (2020), TRIZ is built on seven foundational pillars that guide SI (Table 1): (i) Ideality, striving for the most beneficial outcome with minimal costs; (ii) resources, maximizing the use of available and hidden resources; (iii) functionality-value (FV), enhancing system value by optimizing positive and negative effects; (iv) contradiction, resolving conflicts between system parameters; (v) space, time, domain, and interface (STDI), analyzing problems from spatial, temporal, and interface perspectives; (vi) system transfer, borrowing solutions from other fields; and (vii) system transition, evolving the system's structure or principles.

While these pillars offer a comprehensive foundation, traditional TRIZ is often applied in a linear manner, solving issues step-by-step without continuous feedback or adaptation.

The Mayan calendar-inspired cyclical TRIZ model integrates the seven foundational TRIZ pillars into three overlapping cycles: (i) Short-term (Tzolk'in), (ii) mid-term (Haab), and (iii) long-term (Long Count). Each cycle aims to refine solutions continuously, ensuring ongoing adaptation and sustainability. This model (Fig. 2) draws an analogy to the Mayan calendar, where each cycle represents a different time frame but works together for sustained progress.

Table 1. The theory of inventive problem solving's philosophies and tools (adapted from Sheu et al. 2020)

Philosophy	**Practical tools**
Ideality	Ideal final result
Resources	Harmful resource usage, DS-TPQ
Functionality-value	Patent regeneration
Contradiction	Separation principles, contradiction matrix, parameter deployment
STDI	Effect/resource database, STIC, smart little people, multi-screen viewpoints
System transfer	Feature transfer
System transition	Trends of engineering system evolution

Abbreviations: DS-TPQ: Demand-supply thought-provoking question; STDI: Space, time, domain, and interface; STIC: Space, time, interface, and cost.

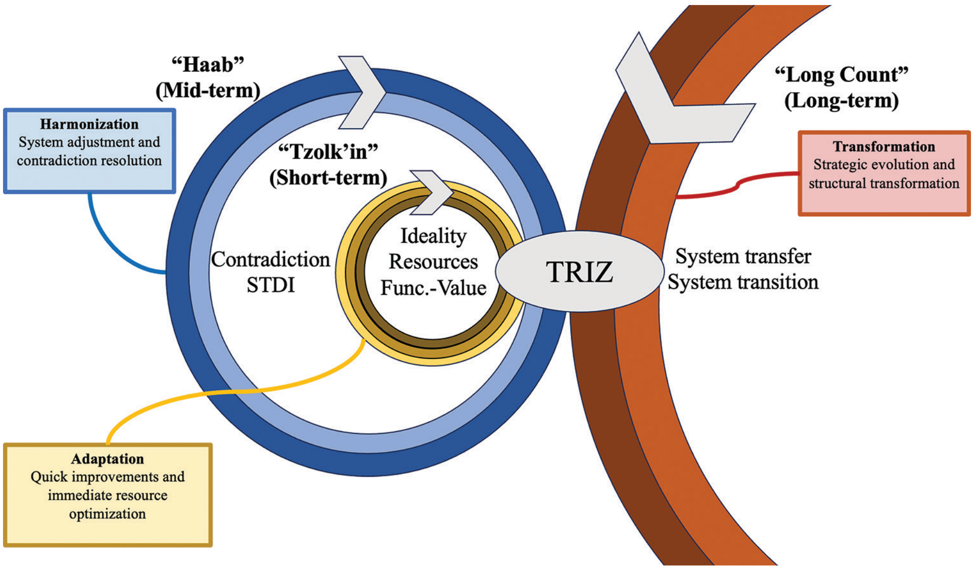


Fig. 2. An overview of the Mayan calendar-inspired cyclical TRIZ model
Abbreviations: Func.: Functionality; STDI: Space, time, domain, and interface; TRIZ: Theory of inventive problem solving.

3.1. Cycle Overview of the Model

i. Short-term cycle (Tzolk'in)
Focus: Quick improvements and immediate resource optimization.
Theory of inventive problem solving's pillars used:
- Ideality: Implement the ideal final result to achieve immediate benefits with minimal costs.
- Resources: Apply tools, such as harmful resource usage analysis and demand-supply thought-provoking questions (DS-TPQ) to maximize existing resources.
- FV: Use patent regeneration to optimize functionality quickly.

Objective: To generate rapid solutions that enhance present performance and reduce inefficiencies.

ii. Mid-term cycle (Haab)
Focus: System adjustment and contradiction resolution.
Theory of inventive problem solving's pillars used:
- Contradiction: Use separation principles and the contradiction matrix (CM) to resolve conflicts between system parameters.
- STDI: Tools such as space, time, interface, and cost (STIC), smart little people (SLP), and the effect/resource database (ERD) help analyze problems from different perspectives.

Objective: To address contradictions and adjust system components for better integration.

iii. Long-term cycle (Long Count)
Focus: Strategic evolution and structural transformation.
Theory of inventive problem solving's pillars used:
- System transfer: Apply tools such as feature transfer to adapt successful solutions from other fields.
- System transition: Use trends of engineering system evolution (TESE) and multi-screen viewpoints to guide long-term transformation.

Objective: To achieve sustainable innovation by evolving the system's core principles.

3.2. Algorithm with Parametric Expressions

Phase 1: Initialize short-term cycle (Tzolk'in):
Objective: Achieve rapid improvements (R).
Inputs: Present system state (S), available resources (Res), immediate contradictions ($C1$)
Process:

Step-1: Apply Ideality (I):
- Target: Maximize beneficial outcomes and minimize costs.
- Formula: $R = max(I) - min(Cost)$

Step-2: Use resource optimization ($ResOpt$):
- Analyze and utilize available resources.
- Formula: $ResOpt = f(DS\text{-}TPQ, HarmfulUsage)$

Step-3: Implement FV improvements:
- Focus on rapid generation of functionality.
- Formula: $FV = f(PatentRegeneration)$

Outputs: Rapid improvements (*R*), refined resources (*Res*), improved value (*V*).

Feedback: Measure short-term impact and feed results (*F1*) to the next cycle.

Phase 2: Transition to mid-term cycle (Haab):

Objective: Address systemic contradictions and refine solutions (*Ref*).

Inputs: Refined system state (*S*), unresolved contradictions (*C2*), feedback (*F1*).

Process:

Step-1: Apply contradiction analysis (CA):

- Use separation principles (*Sep*), parameter deployment and manipulation (*PDM*), and *CM* to resolve conflicts between system parameters.
- Formula: *CA = f(Sep*, *PDM*, *CM)*

Step-2: Implement STDI analysis:

- Use STIC, *SLP*, *ERD*
- Formula: *STDI = f(STIC, SLP, ERD)*

Outputs: Refined solutions (*Ref*), reduced contradictions (*C*↓)

Feedback: Gather mid-term evaluation (*F2*) and feed results to the next cycle.

Phase 3: Advance to long-term cycle (Long Count):

Objective: Achieve sustainable solutions (*Sust*).

Inputs: Refined solutions (*Ref*), strategic goals (*G*), feedback (*F2*).

Process:

Step-1: Use system transfer (*SysTransfer*):

- Integrate features from other fields.
- Formula: *SysTransfer = f(FeatureTransfer)*

Step-2: Implement system transition (*SysTransition*):

- Adapt and evolve core principles.
- Formula: *SysTransition= f(TESE, MultiScreen)*

Outputs: Sustained solutions (*Sust*), evolved system structure (*E*).

Feedback: Establish long-term evaluation (*F3*) and inform the next short-term cycle to use feedback from the long-term cycle to refine initial solutions.

- Formula: *R = f(Sust, F3)*

4. An Illustrative Example

The design and improvement of a coffee machine provide an illustrative example of how the Mayan calendar-inspired cyclical model can be applied to an actual product development process. The coffee machine represents a typical engineering system that can benefit from ongoing refinement across short-term, mid-term, and long-term cycles.

4.1. Cycle Overview of the Example

i. Short-term cycle (Tzolk'in)

Objective: Achieve quick improvements that enhance the coffee machine's immediate performance and reduce costs.

Theory of inventive problem solving's pillars used:

- Ideality: The goal is to create a coffee machine that provides the optimal taste with minimal energy and water usage. Here, we can apply the ideal final result by identifying features that can be improved quickly. For example, optimizing the water heating system to use less energy while maintaining the ideal brewing temperature.
- Resources: To maximize the use of available resources, we can use tools like DS-TPQ for a deep search for technical and physical quantities. For example, utilizing residual heat from the brewing process to pre-heat the water for the next cycle can be a quick improvement.
- FV: By applying patent regeneration, we can analyze existing patents related to coffee machines and adapt features that enhance value, such as an adjustable pressure control for better extraction.

Output: Improved energy efficiency, better taste consistency, and faster brewing times. These short-term improvements are immediately tested and refined based on user feedback.

ii. Mid-term cycle (Haab)

Objective: Resolve contradictions and optimize system components to enhance the coffee machine's functionality over a medium time frame.

Theory of inventive problem solving's pillars used:

- Contradiction: One common contradiction in coffee machines is the balance between high pressure for espresso and lower pressure for drip coffee. Using separation principles and the CM, the system can be designed to separate the two pressure levels within the same unit. For example, a dual-pressure mechanism that adjusts based on the brewing mode can solve this contradiction.
- STDI: By analyzing the machine's spatial design and interface, we can identify areas where space can be better utilized. For example, using a more compact design for water reservoirs without compromising capacity. Tools such as STIC and SLP can help simulate user interactions and improve the design for a better user experience.

Output: More versatile brewing options, improved user interface, and reduced space requirements. Mid-term solutions focus on resolving deeper issues and adapting the design based on accumulated user feedback.

iii. Long-term cycle (Long Count)
Objective: Evolve the coffee machine's core principles and introduce long-term sustainable innovations.
Theory of inventive problem solving's pillars used:

- System transfer: Borrow solutions from other fields to introduce advanced features. For example, adopting feature transfer from smart appliances to include Internet of Things (IoT) capabilities, allowing the coffee machine to connect to a smartphone app for remote control and personalized brewing preferences. For another example, some functions of the coffee machine can be integrated with other kitchen appliances to create a new product category (e.g., combining a coffee machine with a water filter).
- System transition: By applying TESE, the machine can be designed to evolve from a manual operation to a fully automated, artificial intelligence (AI)-driven coffee maker that learns user preferences over time. This can include features such as self-cleaning, auto-refill, and predictive maintenance based on sensor data.

Output: A smart, adaptable coffee machine that is capable of evolving with user needs and technological advancements. Long-term solutions ensure the machine remains relevant and competitive in the market.

The short-term cycle focuses on quick wins and immediate efficiency improvements. The mid-term cycle targets more complex contradictions and system integration issues, making the machine more versatile and user-friendly. The long-term cycle drives strategic evolution and sustainable innovation, making the product future-proof. This cyclical approach ensures that short-term improvements, mid-term adjustments, and long-term evolution are interconnected, creating a continuous loop of SI that adapts to changing user needs and technological trends.

4.2. Algorithm with Parametric Expressions

Phase 1: Initialize short-term cycle (Tzolk'in):
Objective: Achieve rapid improvements (*R*) in the coffee machine's performance and efficiency.

Inputs: Present system state (*S*), available resources (*Res*), immediate contradictions (*C1*)
Process:
Step-1: Apply Ideality (*I*):

- Formula: *R = max(I) – min(Cost)*
- Application: Optimize the water heating system for energy efficiency.
- Example: Implement a feature to reuse residual heat for pre-heating, reducing energy consumption.
- *R = f(IdealHeatUsage, min(EnergyCost))*

Step-2: Use resource optimization (*ResOpt*):

- Formula: *ResOpt= f(DS-TPQ, HarmfulUsage)*
- Application: Maximize available resources, such as the water used in brewing cycles.
- Example: Use DS-TPQ to identify underutilized heat or pressure that can be optimized.
- *ResOpt = f(WaterPressure, HeatResidual)*

Step-3: Implement FV improvements:

- Formula: *FV = f(PatentRegeneration)*
- Application: Adapt existing patents to enhance functionality.
- Example: Adjust pressure control to optimize extraction for better taste.
- *FV = f(OptimalPressure, ExtractionTime)*

Outputs: Rapid improvements (*R*), optimized resources (*Res*), and enhanced value (*V*).
Feedback (F1): Measure energy savings, brewing speed, and taste improvement.

Phase 2: Transition to mid-term cycle (Haab):
Objective: Address systemic contradictions (*C*) and refine components for better system integration (*Ref*).
Inputs: Refined system state (*S*), unresolved contradictions (*C2*), feedback (*F1*).
Process:
Step-1: Apply CA:

- Formula: *CA = f(Sep, PDM, CM)*
- Application: Resolve the contradiction between high pressure for espresso versus low pressure for drip coffee.
- Example: Implement a dual-pressure system that adjusts automatically.
- *C = f(high, low) → Resolved C*

Step-2: Implement STDI analysis:

- Formula: *STDI = f(STIC, SLP, ERD)*
- Application: Analyze space and interface for better user experience.
- Example: Redesign the water reservoir to save space while maintaining capacity.
- *Space = f(ReservoirDesign, Compactness)*

Outputs: Refined solutions (*Ref*), reduced contradictions ($C\downarrow$).

Feedback (*F2*): Evaluate usability, versatility, and system adjustments.

Phase 3: Advance to long-term cycle (Long Count):

Objective: Achieve sustainable solutions (*Sust*) by evolving core features and principles.

Inputs: Refined solutions (*Ref*), strategic goals (*G*), feedback (*F2*).

Process:

Step-1: Use system transfer (*SysTransfer*):

- Formula: *SysTransfer = f(FeatureTransfer)*
- Application: Integrate features from other smart appliances.
- Example: Add IoT capabilities for remote operation and customization.
- *Sust = f(AppIntegration, UserPreferences)*

Step-2: Implement system transition (SysTransition):

- Formula: *SysTransition=f(TESE, MultiScreen)*
- Application: Evolve from a manual to an AI-driven machine.
- Example: Use predictive maintenance to notify users about potential issues.
- *AI-Transition = f(MaintenancePred., Adapt.)*

Outputs: Sustained solutions (*Sust*), evolved system structure (*E*).

Feedback (*F3*): Assess long-term sustainability, user adaptation, and technological integration.

5. Application in Various Industries

The cyclical TRIZ model introduced in this study extends beyond consumer products, such as coffee machines, and can be effectively applied in automotive engineering. One of the key challenges in automotive seat design is achieving an optimal balance between comfort, safety, durability, and energy efficiency while meeting consumer expectations and regulatory standards.

Through the application of cyclical TRIZ principles, automotive seats can be systematically improved by addressing contradictions and developing adaptive, smart seat functions for both conventional and autonomous vehicles.

Phase 1: Adaptation/short-term optimization (Tzolk'in)

Objective: To improve comfort and energy efficiency using existing materials and minor design changes.

Examples:

- Breathable seat materials that dynamically regulate airflow based on external temperature and body sweat levels.
- Passive seat ventilation that does not require additional power but improves air circulation to prevent discomfort.

Innovation impact:

- These low-cost improvements offer immediate functional benefits without requiring complex redesigns.
- They align with TRIZ principles of FV optimization, ensuring better comfort without energy waste.

Phase 2: Harmonization/mid-term contradiction resolution (Haab)

Objective: To resolve contradictions between comfort, durability, and ergonomics while enhancing seat functionality.

Examples:

- Active lumbar support systems that automatically adjust based on driving conditions and driver fatigue.
- Pressure-sensitive seat cushions that dynamically redistribute weight to prevent discomfort during long-distance driving.

Innovation impact:

- These solutions apply TRIZ contradiction resolution principles, improving both comfort and durability.
- The implementation of adaptive support mechanisms enhances seat ergonomics while maintaining long-term structural stability.

Phase 3: Transformation/long-term system innovation (Long Count)

Objective: To introduce intelligent and autonomous seat functions for future mobility solutions.

Examples:

- A smart posture detection system that adjusts the seat position automatically based on the driver's or passenger's body alignment.
- Artificial intelligence-driven personalized comfort settings, using biometric sensors to analyze fatigue levels and stress, adapting temperature, lumbar support, and massage functions accordingly.

Innovation impact:

- This cycle applies TRIZ system transition and transfer principles, ensuring that seat design evolves with future autonomous vehicle needs.
- Artificial intelligence-based comfort management integrates real-time user feedback to create a fully personalized seating experience.

The cyclical TRIZ approach effectively structures the automotive seat innovation process, providing a clear roadmap for short-term optimizations, mid-term contradiction resolution, and long-term transformations.

6. Validation of the Proposed Approach

Validation is essential to establish the scientific credibility and real-world applicability of the cyclical TRIZ model. While a retrospective validation using patent analysis is still in progress, which will be published in a separate study, this section demonstrates that the model is already verifiable within TRIZ itself.

By integrating the TESE, it is shown that the cyclical TRIZ phases naturally align with how engineering systems evolve over time. This ensures that the model is not just a conceptual framework but a structured reflection of real innovation processes.

6.1. TESE: Its Relevance to Validation

The TESE is a systematic methodology that explains how technological systems develop over time through specific evolutionary stages. TESE is an extension of TRIZ and provides a scientific basis for predicting innovation pathways by identifying patterns in system development.

The TESE defines multiple trends that describe how engineering systems evolve (Ghane et al., 2022; Mann, 2003; Sheu and Chiu, 2017), including:

- Increasing ideality: Systems improve functionality while reducing cost and complexity
- Resolving contradictions: Successful systems find ways to eliminate trade-offs between opposing requirements
- Dynamization and adjustability: Systems evolve to become more flexible and adaptable
- System transition: Systems undergo structural and functional transformations into new generations over time
- Integration of intelligence and automation: The highest stage of evolution, where AI and self-learning systems replace manual processes.

By aligning our cyclical TRIZ phases with TESE trends, we provide a strong theoretical foundation for the validity of the model.

6.2. Mapping Cyclical TRIZ Phases to TESE

Each phase of the cyclical TRIZ model aligns with specific TESE evolutionary trends, confirming its consistency with real-world engineering advancements.

Table 2 summarizes that each cyclical phase corresponds to a specific TESE trend, provided that the model follows established evolutionary laws in engineering.

In the adaptation cycle (Tzolk'in), systems evolve by maximizing functionality while minimizing complexity and costs, which aligns with the TESE principle of "increasing ideality." This phase focuses on quick and low-cost improvements that enhance efficiency without requiring major structural changes. In addition, the "resource utilization" principle is evident as early innovations rely on better use of existing materials and functions before undergoing significant modifications. In coffee machines, this is seen in improved water heating efficiency and thermal insulation, while in automotive seats, breathable seat materials and passive ventilation enhance comfort without additional energy consumption.

The harmonization (Haab) cycle corresponds to TESE's "resolving contradictions" principle and "dynamization and adjustability" principles. Engineering systems do not merely optimize existing solutions but instead eliminate inherent trade-offs to improve overall performance. A key example is adaptive lumbar support in automotive seats, which resolves the contradiction between comfort and stability by dynamically adjusting to driver posture and fatigue levels. Similarly, coffee machines evolve to brew different coffee types within the same system, eliminating the trade-off between versatility and efficiency. Over time, systems become more flexible and adaptable, as demonstrated by pressure-sensitive automotive seats that redistribute weight in real time to enhance user comfort.

The transformation (Long Count) cycle aligns with "system transition" and "integration of intelligence and automation" in TESE. At this stage, systems undergo fundamental structural changes, marking a transition from traditional mechanical designs to intelligent and autonomous systems. Coffee machines have evolved from manual brewing methods to AI-powered devices that self-adjust brewing parameters, while automotive seats incorporate smart posture detection systems that autonomously configure seating positions based on biometric data. This trend culminates in AI-driven, self-learning solutions, where both coffee machines and automotive seats dynamically adapt to user preferences and external conditions, ensuring maximum efficiency and personalization.

By mapping the cyclical TRIZ model to TESE evolutionary trends, we demonstrate that the proposed approach is not only conceptually valid but also systematically structured according to engineering system evolution principles. This alignment reinforces

Table 2. Mapping three phases of the proposed cyclical theory of inventive problem-solving model to trends of engineering system evolution (TESE)

Phases	Objective	Relevant TESE trend	Example: Coffee machine design	Example: Automotive seat design
Adaptation (Tzolk'in)	Quick, low-cost improvements using available resources.	"Increasing ideality" "Resource utilization"	Optimizing water heating efficiency. Reducing energy waste using thermal insulation.	Breathable seat materials that adapt to humidity. Passive ventilation improves comfort.
Harmonization (Haab)	Resolving functional contradictions to balance efficiency and performance.	"Resolving contradictions" "Dynamization and adjustability"	Dual brewing system for multiple coffee types. Adjustable pressure control for different flavors.	Active lumbar support adapts to driver fatigue. Pressure-sensitive cushions optimize comfort.
Transformation (Long Count)	Fundamental innovation and future-ready system evolution.	"System transition" "Integration of intelligence and automation"	AI-powered coffee makers that adjust brewing settings automatically. IoT-enabled self-learning coffee machines	AI-based seat adjustment using biometric sensors. Fully autonomous seat modes for different driving conditions.

Abbreviations: AI: Artificial intelligence; IoT: Internet of things.

the scientific legitimacy of the model and its applicability across multiple industries.

7. Comparative Analysis against Existing Models

To critically evaluate the cyclical TRIZ model, a comparison with relevant TRIZ-based methodologies is necessary. However, not all SI models are directly comparable. Instead, we selected three models that best align with the core characteristics of the cyclical TRIZ model in terms of problem-solving approach, adaptability, and sustainability. The chosen models, the classical TRIZ, OTSM, and Sheu and Lee's (2011) SIP, were included for the following reasons.

The classical TRIZ, developed by Genrich Altshuller, is the origin of SI methodologies. It is the baseline for all SI models. It provides a reference point to demonstrate the improvements made by the cyclical TRIZ in terms of iterative adaptability and sustainability.

The OTSM is an extension of the classical TRIZ to handle complex, multi-variable problems (Cavallucci et al., 2015; Khomenko & Ashtiani, 2007; Khomenko & Kucharavy, 2002). It introduces problem networks and system-level problem-solving, making it a useful benchmark for evaluating whether the cyclical TRIZ offers a more structured iterative process. Unlike the OTSM, the cyclical TRIZ emphasizes evolutionary, time-dependent problem-solving rather than static network modeling.

Sheu and Lee's (2011) SIP is a structured, phase-wise approach to innovation that incorporates both TRIZ-based tools and non-TRIZ methodologies. This model was selected because it represents an industry-oriented innovation framework with business and technology considerations. Comparing it to the cyclical TRIZ demonstrates how our model enhances adaptability and long-term sustainability beyond structured stage-wise innovation.

This section compares these models against the cyclical TRIZ, focusing on:

- Efficiency: How well the model optimizes resources and eliminates contradictions
- Adaptability: The model's ability to handle evolving innovation challenges dynamically
- Sustainability: How the framework supports long-term innovation processes.

7.1. Efficiency

Unlike the classical TRIZ, which operates in a single problem-solving cycle, the cyclical TRIZ model (i) enables continuous optimization at multiple levels, (ii) feeds short-term improvements into mid-term contradiction resolution to ensure deeper refinements, and (iii) ensures that long-term transformation integrates past refinements into future design improvements.

For example, in automotive seat design, the classical TRIZ can optimize seat adjustability using inventive principles. Meanwhile, the OTSM can model complex interactions between comfort, safety, and cost. On the other hand, Sheu and Lee's (2011) SIP provides a structured process to integrate both TRIZ and non-TRIZ methodologies. In contrast, the cyclical TRIZ continuously refines seat design across multiple innovation cycles, ensuring progressive and sustainable improvements.

Example in automotive seats:

- Short-term adaptation cycle: Optimized ventilation efficiency in seat cushions

Table 3. Comparison table of the proposed model and selected models

Model	Problem-solving approach	Efficiency	Adaptability	Sustainability
Classical TRIZ	Linear contradiction resolution using 40 IPs	High (Effective contradiction elimination)	Low (Does not inherently support iterative adaptation)	Low (One-time problem resolution without built-in iteration)
OTSM	Network-based problem-solving for complex, multi-layered systems	High (Handles interconnected problems well)	Medium (Uses problem flow modeling but lacks cyclical iteration)	Medium (Provides structured knowledge mapping but no built-in renewal mechanisms)
Sheu and Lee's (2011) systematic innovation process	Stage-wise systematic innovation process, integrating TRIZ and non-TRIZ tools	High (Structured phases improve efficiency)	Medium (Stepwise adaptability but lacks cyclical iteration)	High (Integrates business opportunity exploration and cross-industry application)
Cyclical TRIZ model	Iterative cycles for short-term optimization, mid-term contradiction resolution, and long-term transformation	High (Leverages existing TRIZ tools for continuous improvement)	High (Ensures adaptive problem-solving through iterative cycles)	High (Facilitates continuous innovation/evolution rather than static problem-solving)

Abbreviations: IP: Inventive principle; OTSM: General theory of powerful thinking; TRIZ: Theory of inventive problem-solving.

- Mid-term harmonization cycle: Resolving contradictions between comfort and durability
- Long-term transformation cycle: AI-driven smart seating systems that dynamically adjust based on real-time biometric data.

7.2. Adaptability

Most existing TRIZ methodologies, including Sheu and Lee's (2011) SIP, follow a linear, stage-wise structure. While these approaches provide clear problem-solving pathways, they lack built-in iterative mechanisms.

The cyclical TRIZ model ensures (i) short-term improvements feed into mid-term refinements, making it highly adaptive, and (ii) innovation is not a one-time process but a continuous, self-improving loop.

For example, in coffee machine design, Sheu and Lee's (2011) SIP identifies innovative business opportunities but treats problem-solving as a stepwise approach. In contrast, the cyclical TRIZ continuously iterates through adaptation, harmonization, and transformation cycles, ensuring sustained technological evolution (e.g., from manual espresso machines to AI-driven smart coffee systems).

7.3. Sustainability

Unlike the classical TRIZ and OTSM, which focus on single-instance problem resolution, the cyclical TRIZ model provides a mechanism for continuously refining past solutions and structures innovation into a long-term iterative model, ensuring continuous renewal.

For example, in automotive seat design, Sheu and Lee's (2011) SIP provides a structured method to evaluate contradictions but does not inherently recycle improvements into future iterations. The cyclical TRIZ incorporates a continuous refinement mechanism, ensuring that past advancements dynamically influence future designs.

Table 3 presents a comparison table. The comparative analysis confirms that the cyclical TRIZ model provides superior adaptability and sustainability compared to traditional TRIZ methods while maintaining high efficiency. The key advantages of the proposed approach include:

- A structured, iterative approach to innovation that continuously refines solutions
- Higher adaptability than the classical TRIZ, OTSM, and Sheu and Lee's (2011) SIP
- A long-term sustainable innovation model that ensures continuous learning and adaptation.

8. Conclusion

The Mayan calendar-inspired cyclical TRIZ model represents a significant advancement in the field of SI, offering a structured and flexible approach that ensures both short-term efficiency and long-term adaptability. Unlike traditional TRIZ applications, which often focus on solving individual contradictions in a linear manner, this model emphasizes a continuous feedback loop, ensuring that each phase builds upon and refines previous innovations: (i) Short-term cycle (adaptation) allows for quick wins by maximizing available resources and making low-cost optimizations; (ii) mid-term cycle (harmonization) targets more complex contradictions, enabling system-wide improvements and integration; and (iii) long-term cycle (transformation) ensures sustainable innovation by aligning product evolution with future technological advancements.

The practical effectiveness of this model is demonstrated through its application in coffee machine design and automotive seat design, where each cycle contributes to a progressive and self-sustaining improvement process. Furthermore, the alignment of these cycles with TESE principles confirms that the proposed model is not only theoretically sound but also practically validated through real-world technological evolution.

By providing a dynamic and adaptive problem-solving approach, the cyclical TRIZ model extends beyond conventional TRIZ applications, making it an effective tool for industries requiring SI. This model bridges the gap between structured problem-solving and sustainable innovation, offering a scalable solution for industries seeking both rapid improvements and long-term strategic evolution.

Future research should focus on retrospective validation through patent analysis to quantify the model's effectiveness. Previous studies have successfully employed such analysis to validate technological evolution patterns (Rahim & Iqbal, 2023). Future research could also further explore its application across various industries, examining its effectiveness in more complex systems and diverse contexts.

References

Brandenburg, F. (2002). *Methodology for Planning to Technological Product Innovation*. Shaker Verlag, Aachen.

Cavallucci, D., Fuhlhaber, S., & Riwan, A. (2015). Assisting decisions in inventive design of complex engineering systems. *Procedia Engineering,* 131, 935–983.

Chanier, T. (2018). A possible solution to the mayan calendar enigma. *The Mathematical Intelligencer,* 40, 18–25.

Ghane, M., Ang, M.C., Cavallucci, D., Kadir, R.A., Ng, K.W., & Sorooshian, S. (2022). TRIZ trend of engineering system evolution: A review on applications, benefits, challenges and enhancement with computer-aided aspects. *Computers and Industrial Engineering*, 174, 108833.

Hansen, M.T., & Birkinshaw, J. (2007). The innovation value chain. *Harvard Business Review*, 85(6), 121–130.

Ilevbare, I.M., Probert, D., & Phaal, R. (2013). A review of TRIZ, and its benefits and challenges in practice. *Technovation*, 33, 30–37.

Khomenko, N., & Ashtiani, M. (2007). *Classical TRIZ and OTSM as a Scientific Theoretical Background for Non-Typical Problem-Solving Instruments*. ETRIA Future, Frankfurt.

Khomenko, N., & Kucharavy, D. (2002). *OTSM-TRIZ Problem Solving Process: Solutions and their Classification*. *Proceedings of TRIZ Future Conference*. Strasbourg, France.

Kruger, L.L.S.J., Pretorius, J.H.C., & Erasmus, L.D. (2019). Towards a comprehensive systematic innovation model: A literature review. *SAIEE Africa Research Journal*, 110, 39–46.

Mann, D. (2003). Better technology forecasting using systematic innovation methods. *Technological Forecasting and Social Change,* 70, 779–795.

Mann, D. (2007). *Hands-on Systematic Innovation*. IFR Press, Devon.

Mann, D. (2023). TRIZ and chaos: First principles for emergency first responders. In: Cavallucci, D., Livotov, P., & Brad, S., editors. *Towards AI-Aided Invention and Innovation, IFIP Advances in Information and Communication Technology*. Springer Nature Switzerland, Cham, p453–464.

Rahim, Z.A., & Iqbal, M.S. (2023). Hydrogen fuel cell technology revolution and intervention using TRIZ S-curve analysis for automotive system innovation. *Chemical Engineering and Technology,* 46(12), 2590–2599.

Roper, S., Du, J., & Love, J. (2008). Modeling the innovation value chain. *Research Policy*, 37(6–7), 962–977.

Sheu, D.D., & Chiu, S.C. (2017). Prioritized relevant trend identification for problem solving based on quantitative measures. *Computers and Industrial Engineering,* 107, 327–344.

Sheu, D.D., & Lee, H.K. (2011). A proposed process for systematic innovation. *International Journal of Production Research,* 49, 847–868.

Sheu, D.D., Chiu, M.C., & Cayard, D. (2020). The 7 pillars of TRIZ philosophies. *Computers and Industrial Engineering*, 146, 106572.

Sun, J., Li, H.Y., Du, Y.J., Song, Z., & Tan, R. (2020).

A systematic innovation process oriented to inter-discipline. In: Cavallucci, D., Brad, S., & Livotov, P., editors. *Systematic Complex Problem Solving in the Age of Digitalization and Open Innovation, IFIP Advances in Information and Communication Technology*. Springer International Publishing, Cham, p257–267.

Wang, F., Tan, R., Wang, K., Cen, S., & Peng, Q. (2024). Innovative product design based on radical problem solving. *Computers and Industrial Engineering*, 189, 109941.

Zhan, Y., Tan, K.H., Ji, G., Chung, L., & Tseng, M. (2017). A big data framework for facilitating product innovation processes. *Business Process Management Journal,* 23, 518–536.

AUTHOR BIOGRAPHY

Koray Altun is an assistant professor of Industrial Engineering at Bursa Technical University and a Research Collaborator at GLORAD. He is also the founder and manager of a startup focused on digital innovation and software consultancy. He holds a Ph.D. and a B.Sc. degree in Industrial Engineering from Gaziantep University and Erciyes University, respectively. He has published papers in well-respected journals and has led consulting engagements in reputable companies with a focus on R&D, technology, and innovation management. His recent research interests include the innovation excellence model, systematic innovation, technology and innovation management, and R&D management.

DOI: 10.6977/IJoSI.202506_9(3).0003
I. Appaji, P. Raviraj/Int. J. Systematic Innovation, 9(3), 20-30 (2025)

Hybrid intelligence model for traffic management in intelligent transportation systems

Impana Appaji[1,2], Pandian Raviraj[3]*

[1]GSSS Institute of Engineering and Technology for Women, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

[2]Infosys Limited, Mysuru, Karnataka, India

[3]Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

*Corresponding author E-mail: impana.appaji@gmail.com

(Received 13 January 2024; Final version received 16 March 2024; Accepted 16 July 2024)

Abstract

A typical traffic environment in an intelligent transportation system (ITS) involves various infrastructural units that generate a vast amount of sophisticated traffic data. Such a form of complex data is challenging to analyze and hence poses a potential issue in designing an effective and responsive traffic management system. Therefore, this paper develops an analytical modeling approach to harness the potential of artificial intelligence and computational intelligence. The scheme presents a simplified predictive approach that is meant to mitigate current issues and promote intelligent traffic management. The simulated outcome of the study showcases that the proposed scheme offers a significant advantage in its predictive performance in ITS.

Keywords: Artificial Intelligence, Computational Intelligence Technologies, Intelligent Transportation System, Traffic Management

1. Introduction

With the rising demands of incorporating modern technologies, the transportation system has been witnessing a new paradigm shift toward revolutionizing itself from a conventional to an intelligent transportation system (ITS) (Shaaban et al., 2021). In contrast to the conventional transportation system, the significance of ITS in the future is considerably high owing to its beneficial features of enhanced safety, smart traffic management, positive impact on the environment, higher efficiency and time-saving, effective infrastructure management, emergency response, and management (Barodi et al., 2023; Chowdhury et al., 2023; Karthikeyan & Usha, 2023; Kumari et al., 2020). However, it is not a simplified process to deploy ITS as there is a heavy infrastructural cost underlying its implementation process (e.g. modern technologies, communication networks, sensors, etc.) (Njoku et al., 2023). Another bigger challenge is to ensure complete traffic management operations, knowing that devices and services from different manufacturers and vendors are included. At present, there are no standard and generalized technologies and protocols for the effective implementation of ITS, which poses a primary concern for smart traffic management. There is an involvement of various intricate systems in conducting the maintenance and updating operation, which is challenging. Automating the entire process that runs heterogeneous forms of services is a heavily challenging task (Sharma et al., 2021). Another important aspect of traffic management in ITS is the need for reliable and effective decision-making (Rajkumar & Deborah, 2021). To do so, there is a need for real-time traffic data with higher accuracy from various sources. Even if all these issues are solved someday, the bigger challenge will still reside in data privacy and data security in ITS (Kołodziej et al., 2022). It is a computationally challenging task to offer security to the massive amount of data collected from infrastructural devices of different designs and operations. Hence, traffic management will always be a potential research challenge when it is related to a highly distributed and large-scale deployment environment of ITS. However, there are some upcoming research studies where traffic management problems have been attempted to be addressed (Zulkarnain & Putri, 2021).

It has been seen that artificial intelligence (AI) and computational intelligence (CI) have been acting as an evolving solution toward addressing various issues related to dynamic problems, which are considered computationally complex. The prime role played by AI and CI-based schemes is in traffic prediction and optimization, adaptive traffic control systems, intelligent routing and navigation, detection and management of an automated event, smart parking solutions, and analysis and management of specific vehicular or driver behavior (Damaj et al., 2022). It is also noted that such schemes also assist data analysis for effective decision-making, optimizing public transportation networks, and integrating with autonomous vehicles (Bıyık & Yigitcanlar, 2020). Hence, the proposed scheme contributes toward developing a novel framework of traffic management by harnessing the problem-solving capabilities of AI and CI-based schemes. The core notion of this model is to develop a generalized model that offers higher coverage to host smart and intelligent predictive operations on large traffic data in the ITS environment. The contribution of the proposed study is as follows:

1. The proposed model introduces a novel traffic management model that can not only monitor the dynamic condition of lanes but also assist in smarter decision-making in complex ITS scenarios
2. The development of the proposed model is conducted considering the practical constraints as well as parameters, such as location, direction, speed, and lane density, which assist in a better form of predictive decision-making
3. The model implements an AI-based scheme where a long short-term memory (LSTM)-based attention network is used along with a reinforcement learning strategy to make intelligent decision-making for vehicles
4. The presented scheme uses CI-based methodology with a definitive objective function applied on an integrated ITS network for optimized service relaying.

The organization of the manuscript is as follows: Section 2 discusses the frequently used methodologies in traffic management in ITS, followed by a briefing of prominent research issues in Section 3. The discussion of the adopted research methodology is carried out in Section 4, while accomplished simulation outcomes are briefed in Section 5. Finally, Section 6 gives a conclusive summary of the complete work and its contribution.

2. Related Work

The discussion of the related work is restricted only to the state-of-the-art methods witnessed in the current era using CI-based, analytical approaches, and AI-based approaches in traffic management.

- CI-based schemes: This scheme is mainly adopted to enhance the overall transportation system in terms of congestion reduction and improving the flow of traffic. The work carried out by Zhou et al. (2021), Ojala et al. (2020), and Haghighat et al. (2020) has used neural networks as a CI scheme. The majority of the CI schemes have reportedly used fuzzy logic systems to formulate a unique traffic management scheme. The adoption of fuzzy logic has also been witnessed in the work of Tang et al. (2021), where a neural network is used alongside a fuzzy inference system for traffic management. Standalone adoption of fuzzy logic as a CI scheme is reported in the work of Simić et al. (2021), Shi et al. (2020), Lu et al. (2020), Liu et al. (2022), Servizi et al. (2021), and Li et al. (2021), considering various aspects of the urban transportation system. Li et al. (2021) have also used a fuzzy scheme to design a controller system in ITS. Optimization using bio-inspired approaches in clustering the ITS is reported in the work of Jain et al. (2022), Husnain et al. (2023), and Mohammadi & Farahai (2020). The work carried out by Chen et al. (2022) used an evolutionary computing scheme to manage the collaborative traffic framework over the cloud and edge.
- Analytical-based schemes: There are various analytical-based schemes in traffic management in ITS that emphasizes modeling capable of handling various issues. The work of Lee et al. (2021) used a clustering approach where multi-tasking is performed during sequential learning operations to mitigate positioning problems of vehicles. Qadri et al. (2020) emphasized the importance of traffic signals.

Eom & Kim (2020) have a solution for traffic congestion associated with the intersection point. Iliopoulou & Kepaptsoglou (2019) have discussed the likelihood of joint operation of various approaches in mitigating coordination issues in the public transport system. An emergent intelligence scheme is implemented by Chavhan et al. (2021), where the solution is focused on mitigating allocation issues for public transport vehicles in the ITS.

- AI-based schemes: Different variants of AI-based schemes have proven beneficial in making predictive decisions for traffic management in ITS. Adoption of the deep learning approach was witnessed in the work of Malek et al.

(2021), where the prediction of vehicular velocity was carried out using LSTM. The work of Tak et al. (2021) used deep learning for the detection and tracking of vehicular nodes in a traffic scene. Severino et al. (2021) and Sayed et al. (2023) have analyzed the traffic flow and studied its impact on autonomous vehicles using predictive methods. Lytras et al. (2020) used AI to investigate the indicators for catering to dynamic demands in ITS. The model presented by Benterki et al. (2020) uses neural networks and LSTM to perform trajectory prediction by considering the driving behavior of a vehicle for facilitating autonomous maneuvering. Discussion on the importance of the adoption of AI is also seen in the work carried out by Akhtar & Moridpour (2021).

3. Research Problem

After reviewing the existing approaches in traffic management, it is noted that different variants of approaches have been developed. Although such approaches are witnessed with some significant contribution, they are also witnessed with certain impending issues as follows:

- It has been noted that the majority of existing approaches toward ITS have been carried out considering a specific form of research environment and application that restricts the generalized application in practical utilization
- The frequent adoption of approaches in mitigating traffic issues in ITS is carried out using AI-based approaches, which are found to be adopted slightly more often than CI-based approaches. Unfortunately, all these approaches are highly focused on accomplishing their objective function rather than proving their computational efficiency in the presence of a dynamic traffic scenario of ITS
- The adoption of intelligent schemes in ITS is mainly focused on vehicle-to-vehicle communication and not much on the embedded wireless units. This is a prime obstacle to constructing a novel controller and testing it on a large-scale ITS infrastructure
- The majority of existing studies using AI use deep learning approaches where predictive decision-making demands a higher training operation with the inclusion of extensive computational resources. Such schemes, although proven with higher predictive accuracies, could not be used for emergency applications on vehicles in ITS
- Not many existing learning-based approaches are found to emphasize the data quality perspective. It is quite evident that data collected from traffic scenes is highly complex and continuous in form, which imposes another bigger problem in understanding its pattern. Hence, mining proper knowledge from such complex forms of traffic data is a computationally complex process with a lack of studies on optimization approaches.

Therefore, all the above-mentioned statements are prime research problems in carrying out a smart and intelligent traffic management system in ITS. The next section presents a discussion about the problem, solution, and methodologies in the proposed model.

4. Research Methodology

The prime aim of the proposed study is to design a simplified, intelligent, and robust computational framework for traffic management associated with the ITS environment. To facilitate effective decision-making toward modeling highly dynamic traffic management, the proposed scheme places more importance on realizing the importance of traffic data analysis, as well as evolving with a scheme for mitigating complex traffic scenarios in ITS. Hence, the implementation of the proposed scheme is performed using analytical research methods that will offer more flexibility to implement and extensively test the proposed model of traffic management. Fig. 1 highlights the architecture of the proposed scheme.

A closer look at Fig. 1 shows that it is classified into two core operational blocks. The first operational block is meant to harness the potential of the machine learning approach to facilitate effective traffic management in ITS, while the second operational block is meant to harness the potential of CI for further optimizing the performance. According to this architecture, the first step of implementation is constructing a traffic model that considers essential elements of the road network and the characteristics of vehicles. This is followed by constructing a novel conditional logic toward traffic environment formulation. The further reinforcement learning model is used in framing up triple attributes (reward, action, and state) that are further subjected to an LST network

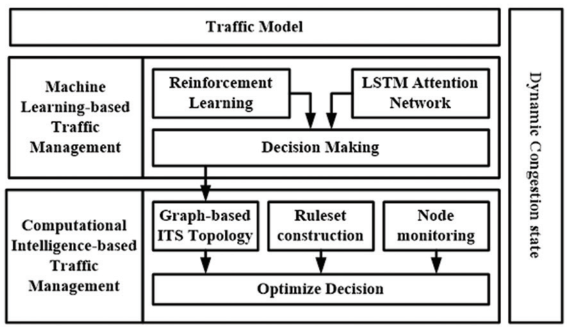


Fig. 1. Proposed architecture of traffic management

with an attention graph to facilitate effective decision-making. On top of the first operation block, the proposed scheme constructs a novel topology of ITS considering triple parameters (location data, vehicle direction, and vehicle speed) that are finally deployed to construct a graph.

This part of the implementation further constructs a novel objective function in incorporating CI with a sole agenda to frame an integrated ITS network system. The scheme also introduces a requestor and non-requestor node in the integrated network that is further subjected to the module of optimizing the service relaying in an ITS environment. Hence, the second operational block acts as a complementary model to the first operational block, where the overall network structure assists in traffic management. A unique use case is adopted for the practical scenario of vehicular movement affected by various road conditions on varying lane densities. Hence, the adopted research methodology targets generating an effective decision-making process that positively and intelligently structures traffic management in the ITS environment.

From the prior discussion, it is noted that the overall implementation of the proposed scheme was performed using two discrete operational blocks in the presented architecture. This section elaborates on the core system implementation associated with the operational blocks of the architecture.

4.1. Machine Learning-Based Traffic Management

This is the primary implementation module, which is responsible for undertaking an intelligent decision-making system in traffic management using a machine learning approach (Fig. 2). The first step in implementation is defining three variants of congestion in traffic, such as high, low, and medium, while specific dimensions, speed, gap between the vehicles, and time of journey characterize each vehicle. These parameters are practically chosen to suit the model deployment, considering actual vehicular characteristics in ITS. The scheme makes use of the CityFlow simulator [40] in a traffic environment where a different number of vehicles are deployed on multiple junctions of the road with a definitive speed and computed congestion. Furthermore, the scheme applies a reinforcement learning approach where the formulation of state, action, and reward is carried out. The primary logic behind applying the reinforcement algorithm is to optimize the signal quality and control the power consumption in the presence of various degrees of congestion. The scheme designs its objective function to ensure lower congestion and higher signal quality to choose the path for the movement of vehicles. Different from any existing scheme, the scheme utilizes average queue length to represent the congestion attribute. The proposed scheme uses a Memory Graph Attention Network, where LSTM is used in its three hidden layers to predict the upcoming possibilities of traffic scenarios for appropriate decision-making. The prime novelty of this approach is that the majority of existing studies have used only LSTM, in which case the system can only undertake decisions on a local level with respect to its agent. However, by applying attention-based LSTM, the system is now capable of realizing the global perspective of the ITS environment to make more accurate decisions. Another significant contribution is that all the agents of the reinforcement learning model can now share their biases and weights using graph attention. This method eventually offers a better understanding of global traffic parameters, considering interdependencies among congestion levels on each lane. Therefore, the system contributes toward yielding a discrete mode of decision with respect to the current congestion level on each route, which is not seen in any existing research work on ITS-based traffic management. The outcome of the study is in the form of optimal paths that offer a better movement option with the least distance and the least fuel consumption, along with reduced power utilization among the infrastructural units.

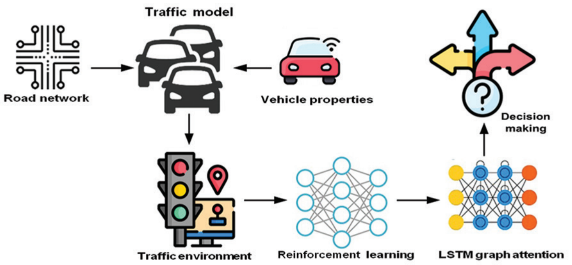


Fig. 2. Machine learning-based traffic management

4.2. CI-Based Traffic Management

This is an extension of the prior implementation module, which is mainly meant to optimize decision-making in ITS. Different from existing methodologies, the system constructs a graph-based ITS topology considering vehicular speed, vehicular direction, and location data (Fig. 3). The system assesses the effective

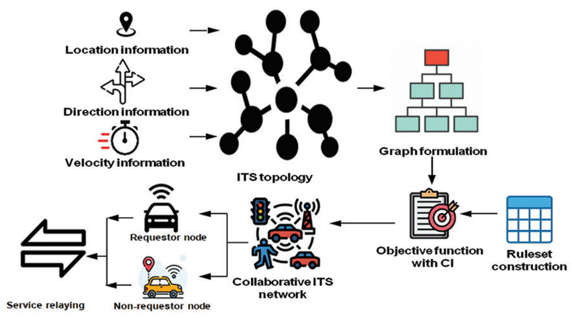


Fig. 3. Computational intelligence-based traffic management

path and position considering spatial distances among the vehicles, while the optimal probability of a better road is determined considering fuzzy logic. The core idea of this approach is to reduce the waiting time for vehicles on specific routes. The study model considers the requestor node as a vehicle seeking assistance with navigation by forwarding a request message to neighboring vehicles. Such a neighboring vehicle can be both a requestor node by itself and a non-requestor node (such a vehicle node does not forward any request for navigational assistance).

The proposed scheme uses CI to construct its objective function. It is to be noted that fuzzy-based rulesets are constructed as a part of CI to guide the type of traffic state. This is further followed by assigning a specific neighboring channel to formulate the outcome toward optimal routes. The next part of the implementation is associated with dispersing the ultimate traffic command over the intersection. The novelty of this study's contribution is that the proposed system generates a unique route for each vehicle on the basis of their best route to a destination, and thereby, it generates distinct and unique traffic commands for each vehicle. The majority of existing schemes generate the same traffic commands for all the waiting vehicles in a lane, while the proposed scheme allows the propagation of unique traffic commands for each vehicle waiting at an intersection point. This clearance signal and its distinctiveness in relaying are significant novelties of the proposed system that is not witnessed in existing methodologies for traffic management in ITS. Apart from this, the scheme also constructs a temporary memory system where each decision to be relayed is stored and subjected to its match with the existing traffic condition. This process also offers a benefit in lowering the computational processing effort by optimizing the analytical operation undertaken for each traffic unit ready to forward optimized navigational information. Hence, a better form of optimized traffic management is presented in the proposed scheme.

5. Results

The complete implementation mentioned in the prior section has been scripted in a Python environment, while the analysis of the optimized outcome is performed in a MATLAB environment. The study considers 800 vehicular nodes bearing the IEEE 802.11 wireless standard that is spread over a monitoring area of 1,000 × 1,000 m^2. Considering synthetically generated data, the analysis was carried out on the mean travel time for both emergency and normal vehicles, as well as power consumption. The study outcome is compared with the relevant work of Wei et al. (2019), where the CoLight model has been presented using a learning-based approach to mitigating similar traffic-related problems in ITS.

Fig. 4 showcases the simulation outcome with a four-channel system for one junction point in the ITS environment. The sample instantaneous visuals of the simulations show various vehicle deployments in a two-lane system with vehicles traversing in two opposite directions on the lanes. The numbers within the simulation show the total number of vehicles that received the traffic command from the controller, either to clear the junction, to wait, or to stop. The numerical outcome of the study is shown in Tables 1-5.

Fig. 5 showcases the clearance controller value for the distinct number of vehicles at each junction point with respect to all four test channels. The computation of this parameter is carried out as

$$C_{cv} = avg\sum_{i=1}^{m} t_c^{\,m} \tag{1}$$

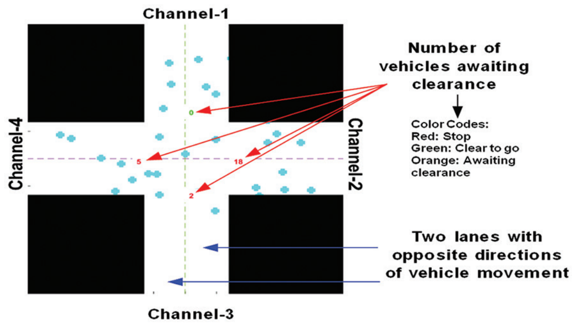


Fig. 4. Simulation outcome of the proposed model

Table 1. Mean duration of journey

Model	6×6 uniflow MD-1	6×6 biflow MD-2	New York MD-3	Hangzhou MD-4
Normal	198	197	1210	210
Dynamic	150	149	600	185

Table 2. Cumulative power utilization

Model	6×6 uniflow MD-1	6×6 biflow MD-2	New York MD-3	Hangzhou MD-4
CoLight	220	225	1210	165
Proposed	209	210	1,009	151

Table 3. Algorithm processing time

Model	6×6 uniflow MD-1	6×6 biflow MD-2	New York MD-3	Hangzhou MD-4
CoLight	0.59	0.52	0.83	0.62
Proposed	0.34	0.37	0.54	0.39

In the above Eq. (1), the computation of the clearance controller value is carried out considering *m* as the maximum number of junction points where distinct traffic commands, t_c, are relayed by the controller. It is to be noted that the value of traffic command t_c is unique for each requestor vehicle. The accomplished outcome in Fig. 5 showcases that there is no significant difference in clearing the vehicle from overall junction points. It implies that the outcome of relaying the clearance signal offers higher consistency for the overall traffic system. This trend of consistency can also be seen in the outcome shown in Fig. 6.

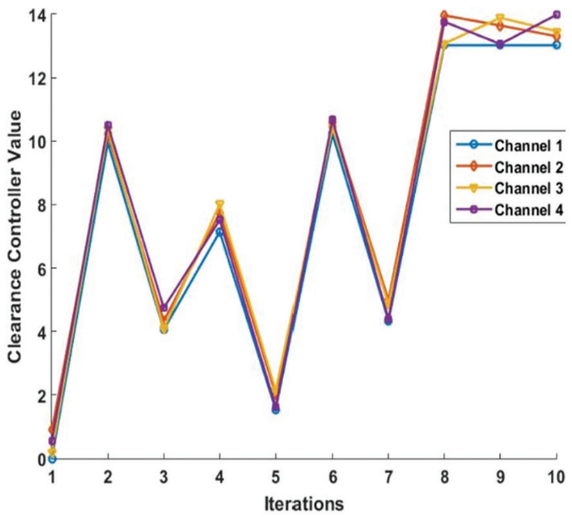


Fig. 5. Simulation outcome of clearance controller value

Table 4. Clearance counter and number of vehicles cleared

Channels	Clearance counter	No. of vehicles cleared
Ch_1	34	451
Ch_2	44	700
Ch_3	34	500
Ch_4	39	750

The outcome exhibited in Fig. 6 is an extensive version of the prior result of controller values. It is mathematically exhibited as,

$$cons\ (C_{cv}) = [array\ C_{cv}{}^{m}] \quad (2)$$

The above Eq. (2) showcases the computation of the consistency of generating a clearance signal. Cons (C_{cv}) is fundamentally a secondary assessment parameter considering the array of obtained traffic commands. However, this computation is carried out with respect to the time of generating the traffic command, and hence, its unit is seconds. The prime notion is to exhibit consistency of duration spent in junction points for relaying the clearance signal.

The graphical results shown in Figs. 7 and 8 show individual outcomes for four different forms of map datasets considered for analysis with respect to the mean duration of journey and cumulative power utilization, respectively. The computation is carried out as follows:

$$M_{dj} = mean\ (d_{ch})^{m} \quad (3)$$

$$P_u = \sum_{i=1}^{m} \Delta p_{v_m} \quad (4)$$

Eqs. (3) and (4) are meant to exhibit the computation of the mean duration of journey, M_{dj}, and power utilization, P_u. The system computes the duration for each channel *dch* with respect to all the junction points *m* and finds its average. It is to be noted that this duration computation starts from the instance the traffic command is relayed. From the perspective of power computation, the system allocates an initialized power (*pinit*) and considers that amount of power (*pd*) is required to forward *d* size of data in one second. Hence, the individual power utilized by a vehicle is computed as Eq. (5)

$$\Delta p_v = pinit\text{-}pd \quad (5)$$

This computation is carried out for all *m* junctions to finally obtain the cumulative utilized power P_u.

The outcome in Figs. 7 and 8 shows that the mean travel time for the proposed scheme is approximately 25% reduced compared to the CoLight model. An extensive test environment has

Table 5. Comparison with modern methods

Approaches	Computational effort
Proposed	Low–Medium
Neural network (Bıyık & Yigitcanlar, 2020; Damaj et al., 2022; Zhou et al., 2021)	High
Fuzzy (Haghighat et al., 2020; Ojala et al., 2020; Servizi et al., 2021; Shi et al., 2022; Simić et al., 2021; Tang et al., 2021)	High
Computational intelligence evolutionary (Chen et al., 2022; Husnain et al., 2023; Jain et al., 2022)	High
Analytical learning method (Chavhan et al., 2021; Eom & Kim, 2020; Iliopoulou & Kepaptsoglou, 2019; Lee et al., 2021; Qadri et al., 2020)	Medium–High
Artificial intelligence (Akhtar & Moridpour, 2021; Benterki et al., 2020; Lytras et al., 2020; Malek et al., 2021; Tak et al., 2021)	High

been considered, adopting 6×6 uniflow, 6×6 biflow, New York, and Hangzhou as the frequently adopted urban transportation maps, where both the proposed and existing CoLight models were assessed. The prime reason for the accomplished outcome is that CoLight uses reinforcement learning in combination with a graph attentional network (which is also used in the proposed scheme), and it offers a benefit to perform prediction for hundreds of traffic signals without much dependency on indexing neighboring intersections.

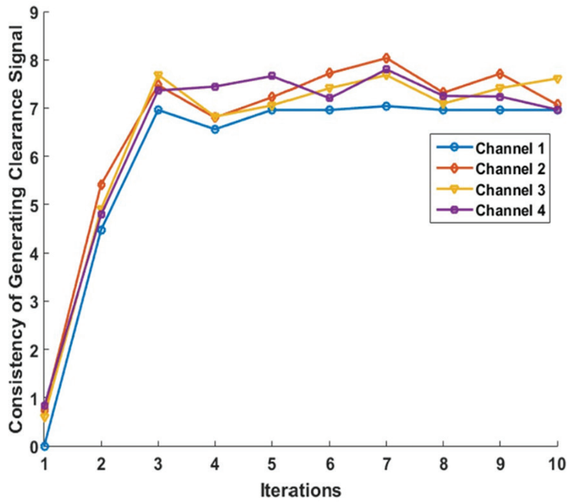


Fig. 6. Simulation outcome of consistency

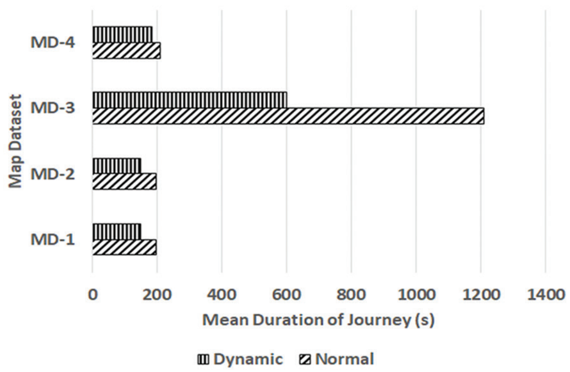


Fig. 7. Analysis of the mean duration of journey

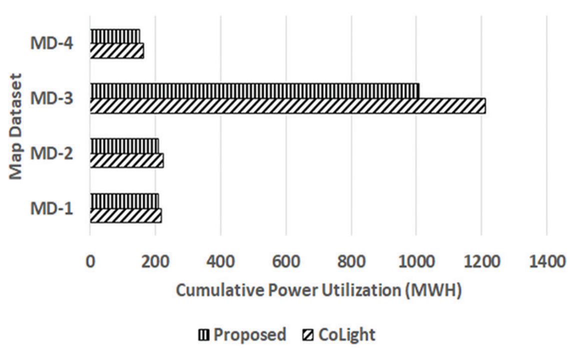


Fig. 8. Analysis of cumulative power utilization

However, the proposed scheme adds upon LSTM that can further handle complex forms of variable-length sequences by allocating different weights to distinct parts of sequences. This significantly improves the predictive performance of the proposed system in shorter time ranges. Apart from this, the proposed model is also proven highly adaptable to different traffic scenarios (e.g., 6 × 6 uniflow, 6 × 6 biflow, New York, and Hangzhou), which causes less power consumption even in the presence of fluctuating traffic conditions. From a computational burden viewpoint, the proposed scheme can be used on different locations and is highly adaptive to different traffic, with only one-time training.

The next part of the implementation performs an analysis for a counter for clearance and quantity of vehicles being cleared in a defined traffic location, as exhibited in Fig. 9. The outcome in Fig. 9 showcases four test channels, which represent adjoining lanes at an intersection point of traffic. Fig. 9a shows that the proposed scheme is capable of maintaining nearly similar consistency of vehicle occupation and clearance on all four tests connected lane systems in an intersection, while Fig. 9b exhibits a maximum number of vehicles being observed to be cleared from each channel. The outcome evidently showcases that the proposed scheme can maintain better consistency on multiple lanes irrespective of any dynamic congestion occurrences in the ITS environment.

Furthermore, Fig. 10 showcases the analysis of time complexity by evaluating the algorithm processing time consumed by the existing CoLight and

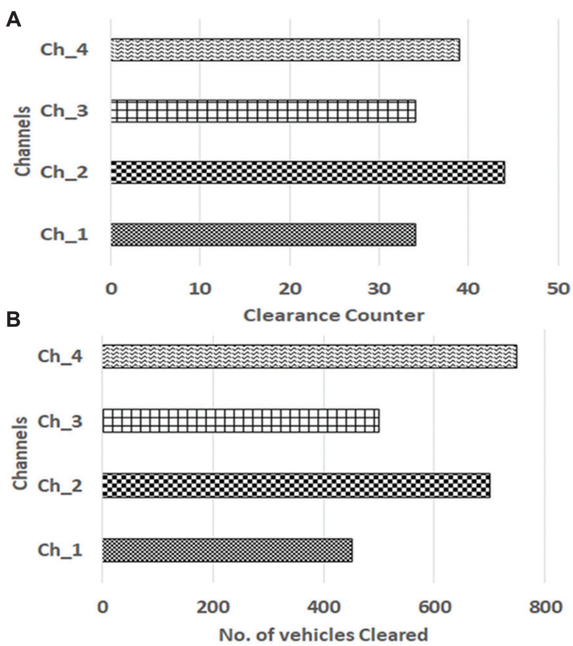


Fig. 9. Analysis of vehicle clearance system. (A) Counter for clearance. (B) Quantity of cleared vehicles

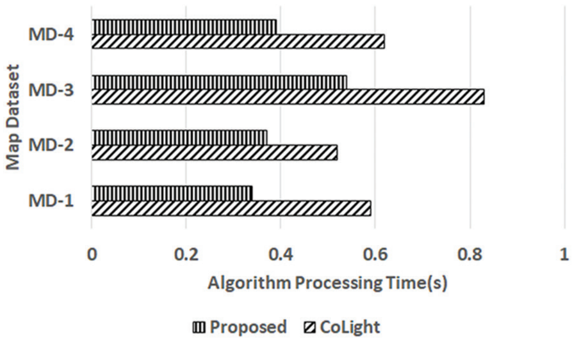


Fig. 10. Analysis of time complexity

the proposed system. It shows that the proposed system consumes approximately 23% of reduced processing time as compared to the existing CoLight. The prime reason behind this is that the existing CoLight does not offer a comprehensive attribute of traffic to undertake global decisions, whereas the proposed system uses both local and global attributes using CI to make the decision easier and faster. Further, the computational effort of the proposed scheme has been compared with the relevant approaches to find that the proposed scheme excels in reduced computational effort in Table 5. The prime reason behind this is that the proposed scheme extracts various intrinsic features, e.g., direction, dimension, and speed of vehicles, along with consideration of lane capacity, to carry out modeling. Acquisition of such information takes less time as it can be derived straight from the ITS infrastructure, while less training is required owing to highly adaptive operations carried out by the LSTM-based attention graph model. This phenomenon turns the proposed scheme into less iterative and more progressive, causing less involvement of computational resources while performing dynamic traffic management in the ITS environment.

6. Conclusion

This paper presented a novel computational model for performing autonomous traffic management with an explicit deployment of a vehicular network in the ITS environment. The proposed study model was designed using analytical research methodology by harnessing the potential of AI- and CI-based approaches. The essential and novel contribution of the study model implementation is as follows:

- Deployment of LSTM with attention network and reinforcement learning assists in realizing an appropriate state of dynamicity in the traffic environment that can offer precise decision-making in the ITS environment.
- The proposed model has been tested over multiple traffic environment data (Hangzhou, New York, 6 × 6 biflow, 6 × 6 uniflow). The outcome exhibited approximately 96% minimized power consumption and 85% minimal travel time.
- The scheme presents a generalized evaluation platform where traffic management problems can be assessed and solved effectively.
- The proposed system model also presents a novel ITS distributed graph-based topology constructed using the velocity, direction, and location information of vehicles.
- A novel CI-based scheme has been implemented using a fuzzy ruleset toward facilitating effective decision-making for clearing the congestion and thereby contributing toward optimized service relaying in ITS.

References

Akhtar, M., & Moridpour, S. (2021). A review of traffic congestion prediction using artificial intelligence. *Journal of Advanced Transportation*, 2021, 1-18. https://doi.org/10.1155/2021/8878011

Barodi, A., Zemmouri, A., Bajit, A., Benbrahim, M., & Tamtaoui, A. (2023). Intelligent transportation system based on smart soft-sensors to analyze road traffic and assist driver behavior applicable to smart cities. *Microprocessors and Microsystems*, 100, 104830. https://doi.org/10.1016/j.micpro.2023.104830

Benterki, A., Boukhnifer, M., Judalet, V., & Maaoui, C. (2020). Artificial intelligence for vehicle behavior anticipation: Hybrid approach based on maneuver classification and trajectory prediction. *IEEE Access*, 8, 56992–57002. https://doi.org/10.1109/access.2020.2982170

Bıyık, C., & Yigitcanlar, T. (2020). Intelligent transport systems in Turkish urban environments: a comprehensive review. *International Journal of Knowledge-Based Development*, 11(4), 382. https://doi.org/10.1504/ijkbd.2020.115037

Chavhan, S., Gupta, D., Chidambaram, R.K., Khanna, A., & Rodrigues, J.J.P.C. (2021). A novel emergent intelligence technique for public transport vehicle allocation problem in a dynamic transportation system. *IEEE IEEE Transactions on Intelligent Transportation Systems*, 22(8), 5389–5402. https://doi.org/10.1109/tits.2020.3011198

Chen, Z.G., Zhan, Z.H., Kwong, S., & Zhang, J. (2022). Evolutionary computation for intelligent transportation in smart cities: *A survey [review article]. IEEE Computational Intelligence Magazine*, 17(2), 83–102. https://doi.org/10.1109/mci.2022.3155330

Chowdhury, A., Kaisar, S., Khoda, M.E., Naha, R., Khoshkholghi, M.A., & Aiash, M. (2023). IoT-based emergency vehicle services in

Intelligent transportation system. *Sensors (Basel, Switzerland)*, 23(11), 5324. https://doi.org/10.3390/s23115324

Damaj, I., Al Khatib, S.K., Naous, T., Lawand, W., Abdelrazzak, Z.Z., & Mouftah, H.T. (2022). Intelligent transportation systems: A survey on modern hardware devices for the era of machine learning. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 5921–5942. https://doi.org/10.1016/j.jksuci.2021.07.020

Eom, M., & Kim, B.I. (2020). The traffic signal control problem for intersections: A review. *European Transport Research Review*, 12(1), 50. https://doi.org/10.1186/s12544-020-00440-8

Haghighat, A.K., Ravichandra-Mouli, V., Chakraborty, P., Esfandiari, Y., Arabi, S., & Sharma, A. (2020). Applications of deep learning in intelligent trans- portation systems. *Journal of Big Data Analytics in Transportation*, 2(2), 115–145. https://doi.org/10.1007/s42421-020-00020-1

Husnain, G., Anwar, S., Sikander, G., Ali, A., & Lim, S. (2023). A bio-inspired cluster optimization schema for efficient routing in vehicular ad hoc networks (VANETs). *Energies*, 16(3), 1456. https://doi.org/10.3390/en16031456

Iliopoulou, C., & Kepaptsoglou, K. (2019). Combining ITS and optimization in public transportation planning: state of the art and future research paths. *European Transport Research Review*, 11(1), 27. https://doi.org/10.1186/s12544-019-0365-5

Jain, R., Chakravarthi, M.K., Kumar, P.K., Hemake-Savulu, O., Ramirez-Asis, E., Pelaez-Diaz, G., *et al*. (2022). Internet of Things-based smart vehicles design of bio-inspired algorithms using artificial intelligence charging system. *Nonlinear Engineering*, 11(1), 582–589. https://doi.org/10.1515/nleng-2022-0242

Karthikeyan, H., & Usha, G. (2023). A secured IoT-based intelligent transport system (IoT-ITS) framework based on cognitive science. *Soft Computing*, 28, 13929–13939. https://doi.org/10.1007/s00500-023-08410-7

Kołodziej, J., Hopmann, C., Coppa, G., Grzonka, D., & Widłak, A. (2022). Intelligent transportation systems-models, challenges, security aspects. In: *Cybersecurity of Digital Service Chains*. Berlin: Springer International Publishing. pp56–82.

Kumari, S., Kumari, S., Vikram, V., Kumari, S., & Gouda, S.K. (2020). Smart traffic management system using IoT and machine learning approach. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3647656

Lee, S., Lim, D.E., Kang, Y., & Kim, H.J. (2021). Clustered multi-task sequence-to-sequence learning for autonomous vehicle repositioning. *IEEE Access*, 9, 14504–14515. https://doi.org/10.1109/access.2021.3051763

Li, C., Zhao, H., Huang, K., & Chen, Y.H. (2021). Optimal design for anti-skid control of electric vehicles by fuzzy approach. *Chinese Journal of Me-Chanical Engineering*, 34(1), 125. https://doi.org/10.1186/s10033-021-00642-8

Li, S., Shu, K., Chen, C., & Cao, D. (2021). Planning and decision-making for connected autonomous vehicles at road intersections: A review. *Chinese Journal of Mechanical Engineering*, 34(1), 133. https://doi.org/10.1186/s10033-021-00639-3

Liu, Y., Xu, Q., Guo, H., & Zhang, H. (2022). A type-2 fuzzy approach to driver-automation shared driving lane keeping control of semi-autonomous vehicles under imprecise premise variable. *Chinese Journal of Mechanical Engineering*, 35(1), 46. https://doi.org/10.1186/s10033-022-00706-3

Lu, Y., Lang, M., Sun, Y., & Li, S. (2020). A fuzzy intercontinental road-rail multimodal routing model with time and train capacity uncertainty and fuzzy programming approaches. *IEEE Access*, 8, 27532–27548. https://doi.org/10.1109/access.2020.2971027

Lytras, M.D., Chui, K.T., & Liu, R.W. (2020). Moving towards intelligent transportation via artificial intelligence and internet-of-things. *Sensors (Basel, Switzerland)*, 20(23), 6945. https://doi.org/10.3390/s20236945

Malek, Y. N., Najib, M., Bakhouya, M., & Essaaidi, M. (2021). Multivariate deep learning approach for electric vehicle speed forecasting. *Big Data Min. Anal*., 4(1), 56–64. https://doi.org/10.26599/bdma.2020.9020027

Mohammadi, S., & Farahani, G. (2020). Computational intelligence-based connectivity restoration in wireless sensor and actor networks. *EURASIP Journal on Wireless Communications and Networking*, 2020(1), 198. https://doi.org/10.1186/s13638-020-01831-0

Njoku, J.N., Nwakanma, C.I., Amaizu, G.C., & Kim, D.S. (2023). Prospects and challenges of Metaverse application in data-driven intelligent transportation systems. *IET Intelligent Transport Systems,* 17(1), 1–21. https://doi.org/10.1049/itr2.12252

Ojala, R., Vepsalainen, J., Hanhirova, J., Hirvisalo, V., & Tammi, K. (2020). Novel convolutional neural network-based roadside unit for accurate pedestrian localisation. *IEEE Transactions on Intelligent Transportation Systems,* 21(9), 3756–3765. https://doi.org/10.1109/tits.2019.2932802

Qadri, S.S.S.M., Gökçe, M.A., & Öner, E. (2020). State-of-art review of traffic signal control methods: Challenges and opportunities. *European Transport Research Review*, 12(1), 1-23. https://doi.org/10.1186/s12544-020-00439-1

Rajkumar, S.C., & Deborah, L.J. (2021). An improved public transportation system for effective usage of vehicles in intelligent transportation system. *International Journal of Communication Systems*, 34(13), e4910. https://doi.org/10.1002/dac.4910

Sayed, S. A., Abdel-Hamid, Y., & Hefny, H.A. (2023). Artificial intelligence-based traffic flow prediction: A comprehensive review. *Journal of Electrical Systems and Information Technology*, 10(1), 13. https://doi.org/10.1186/s43067-023-00081-6

Servizi, V., Pereira, F.C., Anderson, M.K., & Nielsen, O.A. (2021). Transport behavior-mining from smartphones: A review. *European Transport Research Review*, 13(1), 57. https://doi.org/10.1186/s12544-021-00516-z

Severino, A., Curto, S., Barberi, S., Arena, F., & Pau, G. (2021). Autonomous Vehicles: An analysis both on their distinctiveness and the potential impact on urban transport systems. *Applied Sciences (Basel, Switzerland)*, 11(8), 3604. https://doi.org/10.3390/app11083604

Shaaban, K., Elamin, M., & Alsoub, M. (2021). Intelligent transportation systems in a developing country: Benefits and challenges of implementation. *Transportation Research Procedia*, 55, 1373–1380. https://doi.org/10.1016/j.trpro.2021.07.122

Sharma, V., Kumar, L., & Sergeyev, S. (2021). Recent developments and challenges in intelligent transportation systems (ITS)-A survey. In: *Algorithms for Intelligent Systems*. Berlin: Springer Singapore. pp37–44.

Shi, Q., Wang, M., He, Z., Yao, C., Wei, Y., & He, L. (2022). A fuzzy-based sliding mode control approach for acceleration slip regulation of battery electric vehicle. *Chinese Journal of Mechanical Engineering*, 35(1), 72. https://doi.org/10.1186/s10033-022-00729-w

Simić, V., Lazarević, D., & Dobrodolac, M. (2021). Picture fuzzy WASPAS method for selecting last-mile delivery mode: A case study of Belgrade. *European Transport Research Review*, 13(1), 43. https://doi.org/10.1186/s12544-021-00501-6

Tak, S., Lee, J.D., Song, J., & Kim, S. (2021). Development of AI-based vehicle detection and tracking system for C-ITS application. *Journal of Advanced Transportation*, 2021, 1–15. https://doi.org/10.1155/2021/4438861

Tang, J., Zhang, X., Yin, W., Zou, Y., & Wang, Y. (2021). Missing data imputation for traffic flow based on combination of fuzzy neural network and rough set theory. *Journal of Intelligent Transportation Systems*, 25(5), 439–454. https://doi.org/10.1080/15472450.2020.1713772

Wei, H., Xu, N., Zhang, H., Zheng, G., Zang, X., Chen, C., *et al*. (2019). CoLight: Learning network-level cooperation for traffic signal control. In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management.*

Zhou, F., Yang, Q., Zhong, T., Chen, D., & Zhang, N. (2021). Variational graph neural networks for road traffic prediction in intelligent transportation systems. *IEEE Transactions on Industrial Informatics*, 17(4), 2802–2812. https://doi.org/10.1109/tii.2020.3009280

Zulkarnain, & Putri, T.D. (2021). Intelligent transportation systems (ITS): A systematic review using a Natural Language Processing (NLP) approach. *Heliyo*n, 7(12), e08615. https://doi.org/10.1016/j.heliyon.2021.e08615

DOI: 10.6977/IJoSI.202506_9(3).0003
I. Appaji, P. Raviraj/Int. J. Systematic Innovation, 9(3), 20-30 (2025)

AUTHORS BIOGRAPHIES

Impana Appaji holds an M.Tech in Software Engineering, Post-Graduation from SJCE Mysuru, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, INDIA in 2014. She also received a B.E. degree in Computer Science from GSSSIETW, Mysuru, and was affiliated with Visvesvaraya Technological University, Belagavi, Karnataka, India, in 2009. She has 13 years of Teaching Experience. She is currently working as a Senior Analyst at Infosys. She has published 12 research papers in international journals and conferences from 2009 to 2023. She can be contacted at email: impana.appaji@gmail.com.

Dr. Raviraj Pandian completed his doctorate degree in Computer Science and Engineering in the area of Image Processing. He holds the position of Professor and Head in the Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysore. In addition, he also serves as an IQAC Director at GSSS Institute of Engineering and Technology for Women, Mysore. He has 16 years of teaching and research experience and has published over 90 papers in International journals and conferences and 8 book chapters. He serves as an Editorial Board Member and Reviewer for over 15 International Journals. He is also a life member of professional bodies such as ISTE and CSI. He has been guiding Ph.D. research scholars in the areas of Image Processing, Pervasive and Cloud Computing, Bio-Inspired Algorithms, and Robotics. He received a project grant of Rs 5 Lakhs from the VGST, Government of Karnataka for the "Underwater Robotic Fish for Surveillance and Pollution monitoring." He served as the chairperson and keynote speaker for many National and International Conferences and as a Board of Studies member, Subject expert, and Ph.D thesis evaluator for various universities and autonomous institutions. He has reviewed many articles in International conference proceedings and journals. He has received various awards such as "Rhastriya Gaurav Award-2015," "Shri P.K.Das Memorial Best Faculty Award-2012," "Young Achiever Award-2016," and "Best Circuit Faculty Finalist Award-2017" for his outstanding academic achievements and contributions.

DOI: 10.6977/IJoSI.202506_9(3).0004
B.G. Shinde, S.B. Sanap, etc./Int. J. Systematic Innovation, 9(3), 31-74 (2025)

New product development process: A conceptual framework for automobile industries

Balasaheb G. Shinde[1]*, Sudarshan B. Sanap[1], Sachin S. Pawar[1], Vishnu D. Wakchaure[2]

[1]Department of Mechanical Engineering, MIT School of Engineering and Sciences, MIT Art, Design and Tecnology University, Pune, Maharashtra, India

[2]Department of Mechanical Engineering, Amrutvahini College of Engineering, Sangamner, Maharashtra, India

*Corresponding author E-mail: balasaheb.shinde@avcoe.org

(Received 15 August 2024; Final version received 04 April 2025; Accepted 23 April 2025)

Abstract

India is emerging as a key destination for global automobile makers, prompting businesses to improve their abilities in product design and development to grow within the technology-focused automobile sector. Managing new product development (NPD) poses significant challenges within the dynamics to remain competitive. A well-defined and proven NPD process in the automobile industry results in high-quality, cost-effective, and timely product delivery to the market. Various frameworks have been proposed in the literature, and limitations highlight the need for a more flexible, integrated, and adaptive NPD model. Utilizing Cooper's highly efficient Stage-Gate framework, this research proposes a new NPD process framework to enhance the performance of the automobile industry. Based on the limitations of existing stages and gates used and a survey among the NPD professionals, detailed activities of the stages and associated gates have been presented.

Keywords: New Product Development, New Product Development Process Framework, Stage-Gate System

1. Introduction

New product development (NPD) is the process of transforming identified market opportunities into a profitable product ready for sale, typically involving a series of steps that companies can utilize to achieve commercialization objectives (Khannan et al., 2021). The literature offers various definitions and explanations regarding the boundaries of NPD. To ensure the effectiveness of NPD, it is crucial to establish seamless coordination among various departments such as manufacturing, engineering, research and development (R&D), marketing, finance, and purchasing. The marketing department takes the lead by conducting an assessment of the new product, followed by the formation of a cross-functional team dedicated to the development of the said product (Gurbuz, 2018). The development of new products is the basis of manufacturing companies as it is the key to leading the market (Rahim & Baksh, 2003). The survival and growth of a company in today's rapidly evolving market heavily rely on the creation of novel and enhanced products (Zhu et al., 2019). Year by year, technological advancements alter the market landscape, causing fluctuating customer demands and increasing market flexibility, which in turn makes NPD more complex to manage. Functional teams are often deployed to streamline the complex tasks involved in NPD, from design to launch (Cano et al., 2021). The true facts are that, out of every seven new product ideas, approximately four undergo development, one and a half are introduced to the market, and only one achieves success (Agrawal & Bhuiyan, 2014). No company likes to develop defective products or cancel the launch because of them. The cancellation of projects in the product development (PD) phase has a negative impact on the industry; such cancellations result in the loss of valuable resources, create a competitive disadvantage by not introducing new or improved products to the market, and lead to overall financial losses (Almeida et al., 2020). Hence, having a systematic NPD process framework is essential for reducing risks, optimizing resources, improving collaboration, and increasing the likelihood of developing successful products.

Intense international competition, driven by rapid technological progress and ever-changing consumer demands, highlights the critical need for companies to develop innovative and competitive new products to succeed. NPD gives organizations an excellent chance to maximize profits and enhance efficiency. To meet the increasing demand for high-quality products that address evolving customer needs, organizations must deliver superior products quickly, leaving no room for errors in the NPD process (Lapunka et al., 2023). Studies consistently show that companies that align their new products with shifting consumer demands are more likely to succeed than those that neglect NPD investments (Brown & Eisenhardt, 1995; Poolton & Barclay, 1998; Yadav et al., 2007). In today's fast-paced business environment, where customer preferences change rapidly, companies must be agile in adapting to these shifts (Singh & Garg, 2015). To succeed, organizations must strengthen their PD capabilities and create products that reflect evolving consumer demands. The NPD process is crucial, especially for companies in markets where product changes happen rapidly (Yadav et al., 2007). Traditional PD methods are slow and prone to significant inefficiencies, with many companies taking three to 4 years to bring a product to market. A study by Anand & Kodali (2008) found that nearly 50% of PD costs are wasted in the NPD process. Furthermore, Rajeshwari (2017) revealed that fewer than 15% of generated ideas succeed in the market, although leading companies achieve an 82.2% success rate, while others only reach 52.9%. These inefficiencies lead to substantial costs, wasted effort, and energy due to high failure rates. While formal NPD processes are standard in most companies and no longer provide a competitive edge (only 6% report lacking such processes), organizations are increasingly adopting flexible, customized approaches tailored to the complexity and size of each project (Brem & Kurzdorfer, 2016). In modern organizations, NPD presents significant risks, but successful companies attribute their success to strong, effective processes that emphasize quality and efficiency.

A well-structured NPD process framework is crucial for guiding functional teams through the product design and development phases with clear, logical, and sequential activities, often illustrated using flowcharts (Yin & Zhang, 2021). When activities are unclear, illogical, or disorganized, it leads to wasted effort, misdirected work, frequent clarification meetings, inaccurate resource and schedule estimates, excessive task dependencies, and ongoing problem-solving, all of which hinder the NPD process. Therefore, it is vital to focus on systematic screening, monitoring, and progression frameworks to address these challenges effectively (Owens & Davies, 2000). Companies continuously refine their frameworks, with a focus on time, cost, and quality, to meet market demands efficiently and reduce the risk of failure. Leading companies consistently leverage robust NPD frameworks to maintain their competitive advantage. Adopting the right framework enhances planning and decision-making, optimizes technology use, allows for evaluation at key milestones, improves cost-efficiency, encourages creativity, strengthens market penetration, boosts revenue, and drives quality improvements. NPD frameworks come in various forms, tailored to the specific needs of different organizations, industries, and products. The choice of framework depends on factors such as product characteristics, market volatility, organizational culture, and the level of innovation desired.

NPD in the automobile industry is a complex and highly competitive process that must consider several key factors due to the industry's rapid technological advancements, evolving consumer preferences, and intense global competition. The nature of competition in the automobile industry influences how companies approach NPD and refine their strategies. According to the 2017 Project Management National Conference in India, the automobile PD process is complicated due to factors such as simultaneous engineering, concurrent activities in PD, the need to introduce the right product to the right market, ensuring each product contributes to overall organizational profitability, and managing the risks associated with running multiple projects or programs concurrently. All these factors make it essential to have a robust process with a continuous plan for improvement. The NPD process must be flexible and adaptable to changing market and customer demands. An efficient, simplified, and flexible NPD process is crucial not only for the survival of a company but also for distinguishing successful firms from less successful ones. As a result, top-performing companies are evolving their NPD processes by incorporating elements of adaptability, agility, and speed into the traditional framework, which is considered rigid (Smolnik & Bergmann, 2020).

The expanding range of NPD frameworks reflects a continuous drive to improve and streamline the process of bringing innovative products to market. The automobile industry has continually worked to improve its PD processes, aiming to reduce failure risks while enhancing efficiency and product quality. The NPD process is essential to this effort, offering a systematic approach to conceptualizing, designing, and launching products. Despite the structured approach, failure cases are not uncommon in the industry, often arising from gaps in the utilization of the NPD framework. The literature review explores the industry scenarios, identifying challenges and analyzing the role of the NPD process in addressing these challenges.

Scholars such as Cooper (2001) have highlighted the Stage-Gate model as a common approach within the industry, emphasizing iterative development and decision-making checkpoints. Effective NPD utilization has been linked to reduced time-to-market, improved product quality, and enhanced customer satisfaction (Clark & Fujimoto, 1991). Failure cases in the automobile industry, such as technical failures, market failures, regulatory failures, and supply chain disruptions, reveal recurring themes of insufficient risk management, inadequate stakeholder engagement, and lack of iterative feedback loops within the NPD process. Despite the popularity of the Stage-Gate framework and efforts to improve NPD performance, innovation failure rates remain high, ranging from 70% to 98%. For an NPD team, Stage-Gate is not merely a process but a series of decision points aimed at aligning resource allocation with business potential. As a result, NPD teams require a more dynamic environment that emphasizes guidance, problem-solving, and knowledge-based risk management, extending beyond the constraints of the traditional Stage-Gate system (Rigen & Welo, 2013). Implementing the Stage-Gate NPD framework can be challenging. It is important not only to modify the Stage-Gate process but also to examine the company's existing business model and capabilities. Companies must take a dynamic approach, continuously reassessing their business models and capabilities, to effectively address challenges and leverage external collaborations (Jaksic et al., 2014).

Stage-Gate is often misunderstood, with one common misconception being that it is a linear and inflexible process (Cooper, 2008). It focuses primarily on internal organizational factors and follows a linear process from exploration to commercialization (Masyhuri, 2022). Common issues with Stage-Gate include:

(1) Too many projects are driven by customer or sales force demands, often leading to quick, uncritical project initiation.
(2) A lack of mechanisms to terminate projects once they have started, resulting in them continuing without clear Go/Kill decision points.
(3) Insufficient criteria for Go/Kill and prioritization decisions, with nearly 50% of firms admitting weaknesses in this area.
(4) Senior management not being sufficiently engaged in the decision-making process, often due to time constraints, lack of understanding, or unpreparedness to make crucial decisions.
(5) The difficulty in terminating projects that seem promising is due to pressure to bring projects to market (Cooper, 2002).

The traditional Stage-Gate process has proven to be cumbersome and less effective in today's complex, unpredictable, non-linear, and interactive market environment (Wind & Mahajan, 1997). Traditional gates are often overly rigid or focused on financial metrics, making the system excessively controlling, bureaucratic, and burdened with unnecessary paperwork, checklists, and tasks that add little value. Consequently, many leading companies are re-evaluating and redesigning their gating systems (Cooper, 2014). Since its introduction 40 years ago, leading firms have refined the model, incorporating techniques like value stream mapping to eliminate bureaucracy and adopting concurrent and parallel processes (Cooper, 2022). Automobile companies that have modified their Stage-Gate processes tend to report various enhancements, including:

(1) The utilization of virtual teams.
(2) The integration of collaborative and virtual tools for NPD.
(3) The establishment of formal strategies dedicated to NPD.
(4) The adoption of structured procedures to guide the NPD process (Ettlie & Elsenbach, 2007).

By addressing the gaps in NPD utilization and drawing lessons from past failure cases, the automobile industry can enhance its resilience and drive sustainable growth. Both management and engineers view the current process as lacking the flexibility and scalability needed to handle the diverse range of projects undertaken by an automobile company. Consequently, there is a need to study the stages and activities of the NPD process and their alignment with various project types. This analysis aims to pinpoint critical challenges and gather the necessary insights to develop a framework for process improvement. Therefore, the objective of this study is to optimize the existing Stage-Gate NPD process within an automobile company while incorporating cutting-edge practices in automobile PD. To achieve this, the study will focus on the following goals:

(1) Identify best practices in automobile NPD through an in-depth literature review and analysis of industry applications.
(2) Conduct a performance measurement survey to uncover areas for improvement in the current NPD process framework.
(3) Propose an improved NPD process framework that addresses the key areas of improvement identified.

1.1. Challenges of NPD

With the evolution of technology and the growing need for flexibility in response to market demands, products and processes are becoming increasingly intricate. This increased complexity introduces additional risks to the NPD process. The ever-increasing customer demands can be viewed as a challenge. Another hurdle is related to gaining knowledge and efficiently handling uncertainty to reduce the chances of failure in the PD process (Cooper & Scott, 2003). The focus of the organization has now turned to short-term objectives, leveraging technological advancements such as rapid prototypes, virtual-reality prototypes, digital twins, artificial intelligence, computer animations and simulations, product lifecycle management, product data management, and other tools. This shift has led to a notable rise in the quantity of "small projects" and improvements in NPD methodologies within the development process (Cooper, 2022). Ensuring the transfer of knowledge between different functional areas and during smaller exchanges between individual team members is a fundamental and recurring challenge in NPD (Ringen & Welo, 2013). The current NPD process includes many time-wasting activities with bureaucratic procedures, a lack of focus, and limited learning opportunities (Gronlund et al., 2010). The PD process, when executed in sequence, frequently leads to extended lead times and increased product costs (Kazimierska & Grębosz-Krawczyk, 2017). The current NPD process is typically outlined in phases and documented in brochures or procedure manuals. However, these documents are often created by central staff not involved in actual development, leading to their perception as mere management procedures. Consequently, project managers and engineers may ignore them, and they quickly become outdated as new practices evolve (Tennant & Roberts, 2003). The existing NPD process also suffers from a lack of various management procedures, well-defined feedback loops for transferring information, and data metrics for assessing and evaluating performance (Ranjan, 2014). Chirumalla (2017) has identified nine significant hurdles in the management of the NPD process. These challenges pertain to resources, time-readiness, and schedule, gated administration, ways of working, communication and time-sharing, learning, business case, coordination and alignment, and competencies. Engineer-to-order companies encounter difficulties in introducing new products due to their customized manufacturing process. These difficulties encompass shorter design cycles, quicker market entry, enhanced product quality, and ongoing cost reductions to maintain competitiveness (Kumar & Wellbrock, 2009). As a result, managing the process of NPD has become increasingly challenging for businesses due to the significant investment of time, finances, and human resources it demands. As per Yang (2016), process complexity is one of the issues in NPD. The NPD process presents a unique set of challenges in comparison to other processes. It requires careful navigation and effective management (Wynn & Clarkson, 2018).

1.2. NPD Process

The NPD process is essential for the success of businesses, especially in today's competitive global market driven by rapid technological progress and evolving consumer preferences. According to Phillips et al. (1999), the product's quality (product performance) is greatly influenced by the quality of the NPD process. To meet the increasing need for high-quality products that meet the changing requirements of customers, organizations must ensure the delivery of improved products within tighter deadlines, allowing no margin for error in the process of NPD. The NPD processes require the participation of essential functional departments within the company, which encompass strategic planning, marketing, product design and development, manufacturing, maintenance, quality, sales, and financial planning. It is organized in a manner that involves engagement with both internal and external stakeholders, including customers and suppliers (Ulrich et al., 2009). The NPD process encompasses all activities involved in bringing a new product to the marketplace, including idea generation, screening, testing, and obtaining customer approval (Wijewardhana et al., 2021). In general, NPD requires eight stages. At the conclusion of every phase, a company must decide whether to proceed to the subsequent phase, abandon PD, or seek additional information. Fig. 1 illustrates the eight stages of the NPD process and is explained below.

(1) Generation of new product ideas: This is the first stage of the NPD process. Various ideas are created using idea-generation techniques, which will help to satisfy needs and examine the evolving technologies. Internal sources, that is, internal departmental members; external sources, that is, customers and competitors; and other sources such as seminars, universities, and investors, are the main sources for idea generation. The survey, which involved 750 interviews with Chief Executive Officers (CEOs) of global businesses, revealed that employees

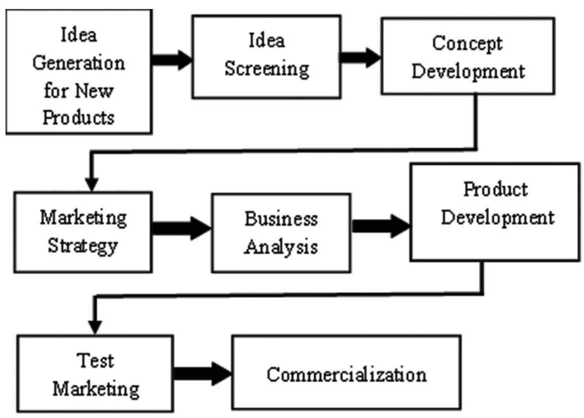


Fig. 1. New product development process

were responsible for generating 41% of new product ideas, customers contributed 36% of ideas, and the R&D department generated only 14% of ideas (Gurbuz, 2018).

(2) Screening and evaluation of ideas: It is a systematic criterion to evaluate the potential of the new product idea. Differentiation between useful and less useful ideas based on various types of feasibility criteria decides the selection of the best ideas.

(3) Concept development and testing: Once the best ideas are selected, product concepts are developed using specific concept selection techniques.

(4) Marketing strategy: At this stage, marketing strategies are determined, including market identification, pricing for the new product, and long-term strategic planning.

(5) Business analysis: It involves studying the estimated economic feasibility of a new product idea, including finalizing the capital and revenue budget based on a make-or-buy decision.

(6) PD: This involves upgrading the product's characteristics to align with customer preferences, along with further development in manufacturing, cost estimation, packaging, and distribution. Prototype development and testing activities are completed, and facilities, tooling, and gauges are designed and developed.

(7) Test marketing: It evaluates major market acceptance through market research, assessing all marketing elements such as the new product concept's target market, market position, advertising, distribution, packaging, and costs.

(8) Commercialization: It involves the actual introduction of the product into the market, including all related resources and decisions. The product launch and associated plans are finalized and executed.

1.3. Review of NPD in The Automobile Industry

The automobile industry's PD process is inherently complex, with varying levels of difficulty depending on the project. Even in the simplest cases, teams must design and develop hundreds, if not thousands, of components. This process must also incorporate customer requirements, design feasibility, performance and safety standards, product quality and reliability, and real-world usage conditions, resulting in an extensive list of factors to consider. Several key factors make this process even more critical: the integration of simultaneous engineering and concurrent activities in vehicle development, the necessity of launching a flawless product in the right market on the first attempt, ensuring that every product contributes to the organization's overall profitability, and effectively managing the risks associated with running multiple projects or programs simultaneously (Koranne & Shende, 2017). PD in automobile NPD differs from other industries due to its complexity, long development cycles, and high regulatory requirements (Ulrich & Eppinger, 2020). Unlike consumer electronics or software, where products can be developed and launched within months, automobile NPD often takes several years due to extensive research, engineering, prototyping, and rigorous safety testing (Thomke & Feinberg, 2006). The process involves collaboration across multiple disciplines, including mechanical, electrical, and software engineering, as modern vehicles integrate advanced technologies such as autonomous driving, electric powertrains, and connectivity features (Nieuwenhuis & Wells, 2015). In addition, automotive NPD must comply with strict government regulations and safety standards across different markets, making testing and validation more intensive than in many other industries. The high costs associated with tooling, manufacturing, and supply chain coordination further differentiate it from industries with lower capital investment requirements (Clark & Fujimoto, 1991). Furthermore, customer expectations for durability, performance, and reliability in automobiles necessitate extensive real-world testing, unlike industries where updates and patches can be released post-launch. These factors make automobile NPD more resource-intensive, risk-prone, and time-consuming compared to many other industries.

An effective NPD process framework, coupled with robust communication, data management, and knowledge handling, stands as a basis for success in NPD projects. Within the dynamic landscape of India's manufacturing industry, particularly in the automobile sector, managing NPD poses a significant challenge. India's automobile market was valued at Rs. 10,000 crore in 2021 and is expected to reach Rs. 16,000 crore in 2027, registering a compound annual growth rate of 8.1% over the forecast period 2022 – 2027. As per the Ministry of External Affairs declaration in August 2023, India's auto industry is expected to rank 3rd in the world by 2030. Due to the rise in middle-class income and the rising young population, the Indian automobile market is expected to witness strong growth. In addition, due to the rising demand for automobiles, exports from the country have also seen a significant increase. The automobile sector accounts for 7.1% of India's gross domestic product and 49% of the manufacturing gross domestic product. This industry provides direct and indirect employment to 1.9 crore people in India. As India emerges as a hub for international car manufacturers, businesses are looking to enhance their capabilities in product design and development within the country to expand their presence in the knowledge-driven automobile industry.

Developing a new product in the automobile industry is a highly complex process. It involves designing and manufacturing hundreds of components, coordinating communication across multiple functions, and ensuring the seamless synchronization of various functional deliverables, each triggering a series of dependent tasks. With this complexity comes inherent risks. These challenges, along with many others, make it crucial to have a well-structured process in place, supported by a continuous improvement plan (Koranne & Shende, 2017). Numerous researchers have previously focused on developing NPD processes tailored to the automobile industry. For instance, the self-assessment NPD process was modified and implemented by Tennant & Roberts (2003) at Land Rover. This led to a notable enhancement in the performance measures of the Freelander compact sport utility vehicle program. Sumantran (2004) conducted a study on Tata Motors' new product introduction process, focusing on the development of a new Sedan Indigo within a span of approximately 20 months. The study highlighted the effectiveness of Tata Motors' approach, which involved a formal vehicle development process, concurrent engineering, math-based simulations, and disciplined manufacturing planning. The implementation of this approach has resulted in significant benefits for the company. A review was conducted on the application of process models, such as the Stage-Gate NPD process framework, to effectively structure and guide the NPD process in the automobile industry. The review concluded with lessons from a benchmarking study on implementing these process models and offered suggestions for future research to enhance their implementation and effectiveness (Chao & Ishii, 2005). An investigative study aimed at enhancing the NPD process in an automobile company, supported by a comprehensive literature review and both qualitative and quantitative research, identified areas for improvement and facilitated the successful implementation, adoption, and adaptation of the process by key drivers (Williams, 2008). The findings of a qualitative meta-analysis involving 16 empirical studies on the success of NPD in the automobile industry reveal a gap between knowledge about the practical relevance of NPD dimensions such as the development process, resources, and strategy and their systematic assessment in practice (Gerhard et al., 2008). A comprehensive PD process (PDP) for the automobile industry was presented, encompassing all stages from market research to sales. This includes customer input, conceptual design, detailed design and engineering, manufacturing process and production, and sales and distribution. The essential features of these five stages were outlined, and specific techniques for their implementation were detailed from a macro perspective with real-life examples (Liang, 2010). A thorough analysis of the six-stage Stage-Gate approach to PD across six different companies found that organizations structured around cross-functional teams tend to favor a low-phased approach, whereas those with a strong functional structure tend to adopt a higher number of phases and gates, indicating a high-phase approach (Phillips et al., 1999). A case study of the Indian automobile industry investigated the impact of original equipment manufacturer collaborations on the convergence of PD processes. It mapped the relationships between automobile companies and other auto original equipment manufacturers in India, concluding that a standardized PD process helps in reducing development cycle time, minimizing platforms, involving suppliers throughout the development stage, promoting cross-functional collaboration, implementing concurrent engineering, and standardizing work practices (Loganathan & Jayakrishnan, 2014). Lean PD was adopted to transform the process of developing Indian automobile products, leveraging foundational lean principles, waste reduction strategies, and a fundamental framework for applying lean practices within the PD process in the Indian automotive industry (Anand et al., 2009). A framework known as Automotive-PDP, developed to oversee the PD process in the automotive industry, was validated through a literature review and a study involving three global automakers from Asia, Europe, and America (Silva & Kaminski, 2017). A study of a small automaker in Malaysia highlighted that its approach to NPD, although aligned with generic NPD processes, integrates concurrent engineering practices. This emphasizes the importance of a structured NPD framework for frequent new product introductions and effective management of risks and uncertainties (Boejang et al., 2017). The following section explores the literature, highlighting the need to redesign NPD process frameworks in the automobile industry.

1.4. The Need for Redesigning The NPD Process Frameworks

Research consistently shows that companies that adapt their new product offerings to meet evolving consumer demands are more likely to succeed than those who fail to invest in NPD initiatives (Brown & Eisenhardt, 1995; Fernandes et al., 2017; Poolton & Barclay, 1998). In every automobile industry, the NPD process has significant value because it greatly influences the whole value chain and decisions on fundamental aspects such as quality, cost, and time. Thus, it is essential to customize the NPD process according to the specific industrial sector. Manufacturing firms within the automobile industry face the challenges of improving their NPD procedure, particularly when operating in markets that demand innovation, reduced time-to-market, an extensive product portfolio, and superior quality (Fernandes et al., 2017). Today, many automobile companies are redesigning their NPD process to be more agile through better governance and portfolio management practices. Some organizations have also created open innovation modifications to their stage procedures or have shifted toward fully automated systems for NPD. Creating new products presents a more significant obstacle than simply extending product lines. Therefore, it requires a more efficient risk management strategy in the NPD process. By improving their grasp of risks and critical factors that may hinder the success of the NPD process, companies can boost their operational effectiveness and improve their ability to predict potential challenges that may affect NPD process results (Salavati et al., 2016). Hence, it is essential to consider redesigning your NPD process if it has been in place for over 5 years or if your idea-to-launch system does not incorporate current best practices (Cooper, 2008). Many reasons discussed in the literature define the need for redesigning NPD process frameworks. The arrival of multiple foreign companies in the Indian automobile sector has brought about a diverse mix of domestic and international firms through foreign direct investment through greenfield investments, mergers and acquisitions, joint ventures, or contract manufacturing. This has made it necessary to incorporate innovative process improvement techniques in the PD frameworks of these companies (Anand et al., 2009). The expansion of regional and global customer and resource markets due to globalization has increased international competition, leading to significant changes in the competitive landscape and practices across various industries. The collaborations between Indian home-grown automobile companies and major American, European, and Japanese players have brought about significant transformations in the PD processes within the Indian automobile industry (Loganathan & Jayakrishnan, 2014). Manufacturers are in competition to enhance performance by reengineering their processes for NPD to accelerate speed, lower costs, and better meet customer demands (Arnold & Floyd, 1997). The implementation of business strategies such as downsizing, outsourcing, and reengineering has resulted in the creation of more streamlined, cross-functional organizations and has fundamentally changed the relationship between companies, employees, customers, and other stakeholders. This has had a significant impact on all aspects of life and business operations, requiring a reevaluation of NPD process frameworks to adapt to these evolving changes (Wind & Mahajan, 1997). In the current fast-moving environment, upper management demands a shortened NPD cycle, necessitating the immediate delivery of new product projects to the market. However, as highlighted in a Forbes article, speeding up innovation carries risks, it may hinder creativity. Hence, the process must be redesigned instead of just condensing the conventional process (Mayer, 2020). In the automobile industry, NPD is a highly complex process involving numerous part developments, managing communication across functions, synchronizing deliverables, and addressing associated risks, necessitating a robust process with continuous improvement plans (Cooper et al., 2003). Hence, top-performing companies are transforming their NPD processes by incorporating adaptability, flexibility, and speed into the traditional Stage-Gate model, known for its inflexible structure (Smolnik & Bergmann, 2020). The next section focuses on a discussion of the five most used NPD process frameworks, out of which the Stage-Gate process offers several advantages.

2. NPD Process Frameworks

Yin & Zhang (2021) defined the NPD process framework, which represents unambiguous, logical, and clearly defined stages in the form of flow charts that describe the design and development activities of the product to be performed by functional teams. Framework translates theory into practice through systematic means and clearly depicts the leadership goal for the organization (Wang & Kourouklis, 2012). Numerous studies have been undertaken and published about the NPD process frameworks. The crucial task for any organization is to have a proper NPD process framework to guide the NPD team so that quality products are successfully introduced into the market with a reduction in cost and development time. According to Shepherd & Ahmed (2000), a robust NPD framework not only sustains product advantage but also enhances new product success, improves company health, and serves as a significant source of competitive advantage, offering benefits such as reduced PD costs, accelerated time to market for first-mover advantages and new product benefits. The five most used NPD process frameworks are selected for discussion in this section, highlighting the importance of the Stage-Gate NPD process framework as compared to Booz, Allen, and Hamilton (BAH), Lean startup, IDEO, and exploratory PD model (Ex-PD) frameworks.

Table 1 compares these five NPD process frameworks across various aspects, including applicability, implementation process, decision-making, objectives, management, and costs, all of which are critical for the successful execution of NPD. These factors vary depending on the industry, company, product type, and market conditions. The information discussed about each of the five frameworks serves

Table 1. Comparison of five macro-categories of new product development frameworks

Aspect	**Stage-Gate System** **Cooper (1990)**	**Lean Start-Up** **Ries (2011)**	**IDEO** **David Kelley (1991)**	**BAH** **Booz, Allen and Hamilton (1982)**	**Exploratory product development model (Ex-PD)** **Mary Drotar and Kathy Morrissey (2015)**
Applicability	Well-funded, established companies with diverse product lines thrive in stable market conditions.	Entrepreneurs, intrapreneurs, governmental bodies, and emerging businesses	IDEO offers design services to clients across diverse sectors, spanning medical, computer, retail, food, and automotive industries in both public and private domains.	The luxury industry, in particular, merits distinctive attention and focus.	Ex-PD is better suited for product development in volatile and unpredictable markets.
Process	Divided into typically five stages- initial inquiry, thorough examination, development, testing and validation, production and launch. The product undergoes progressive development and refinement at each step.	It begins with the entrepreneur's vision, taking a leap of faith. The Build-Measure-Learn approach is embraced to test fundamental assumptions or hypotheses. The product undergoes continual modifications based on frequent and early customer feedback.	IDEO structured the product development process into five key steps: understanding/ observing, visualizing/ realizing, evaluating/ refining, implementing/ detailed engineering, and implementing/ manufacturing liaison.	The BAH model unfolds in seven steps: new product strategy, idea generation, screening, business analysis, development, testing, and commercialization.	The Ex-PD approach comprises three interconnected segments within its process: strategy, idea generation and selection, and exploration and creation.
Decisions	Following each stage, senior managers make a pivotal decision to either proceed or halt the project (Go/Kill).	Choose to persist or pivot. Should the hypotheses prove incorrect, iterate on the project and test refined hypotheses. If rejected, alter the strategy – this shift is referred to as pivoting.	Prototyping and brainstorming stand as crucial approaches in the product development process. Brainstorming sessions assist the project team in generating and finalizing solutions.	The feedback obtained from testing offers nonprofit executives an additional chance to prepare their products for market entry.	Ex-PD operates under the assumption that the product team lacks sufficient knowledge or awareness of the factors contributing to uncertainty and risk.
Goal	Efficiently and effectively launch a polished final product.	Launch a product to validate your assumptions, refine it using feedback, and minimize the wastage of both time and capital resources.	IDEO encourages designers and engineers to rapidly produce prototypes, focusing on various small sections of the project.	Reducing risk can lead to assured long-term growth and eventual profitability through new product introductions.	Ex-PD's main objective is to diminish uncertainty and risk by minimizing the unknown factors.
Management	A gatekeeper, someone with a vested interest in	Owners usually double as project managers, with no	The IDEO project team regularly convenes meetings	Implementing a consistent framework that centers	Ex-PD is characterized as a dual-focused, integrated systems approach that requires

(Cont'd...)

Table 1. *(Continued)*

Aspect	Stage-Gate System Cooper (1990)	Lean Start-Up Ries (2011)	IDEO David Kelley (1991)	BAH Booz, Allen and Hamilton (1982)	Exploratory product development model (Ex-PD) Mary Drotar and Kathy Morrissey (2015)
	the product but not directly managing it, determines each gate.	specifically assigned gatekeepers. Decisions hinge on customer responses to the Minimum Viable Product (MVP).	with clients to gather and incorporate their feedback.	on a comprehensive 360° vision, particularly led by an initial stage entirely dedicated to strategy, allows the process to engage with external networks and the environment.	holistic management, integrating vital elements: strategy, portfolio management, organization/teams/ culture, metrics, market/customer understanding, and process.
Expense	Costs rise with each subsequent gate, and generally, the product does not generate revenue during its development phase.	As the product scales up, expenses grow, but there is a simultaneous generation of revenue during the process.	Costs escalate with rapid prototyping. Generally, the product does not yield revenue during its development phase.	Costs rise at each stage. Generally, the product does not produce revenue during its development.	Costs escalate with each stage. Generally, revenue is not generated during the product's development phase.

as the basis for comparing them against each of these aspects. The comparison reveals that these five NPD process frameworks are well-suited for NPD in Indian industries, and the Stage-Gate NPD process framework proves to be more potent as compared to others. Although other frameworks are effective and useful under certain conditions, the Stage-Gate process offers several advantages.

(1) Performance and operations: The Stage-Gate process provides a highly structured, risk-averse, and systematic approach to NPD, offering clear decision points, quality control, and predictability, making it well-suited for complex projects and large organizations. While BAH, Ex-PD, Lean Startup, and IDEO offer flexibility and rapid iteration, they may lack the formal structure and control that Stage-Gate provides, making them less suitable for projects requiring rigorous oversight and detailed planning.

(2) Documentation and traceability of decisions and progress: The Stage-Gate process stands out in NPD for its rigorous documentation practices that ensure thorough traceability of decisions, progress, and project history. This structured approach supports accountability, auditing, learning from past experiences, and effective communication across teams and stakeholders. In contrast, BAH, Ex-PD, Lean Startup, and IDEO, while offering flexibility and innovation, may not prioritize or achieve the same level of comprehensive documentation and traceability in NPD processes.

(3) Cross-functional collaboration: Stage-Gate excels in fostering structured and formalized cross-functional collaboration through defined stages and gates, clear roles and responsibilities, and formal communication channels. This structured approach helps integrate diverse perspectives and expertise across functions, ensuring alignment and informed decision-making throughout the NPD process. In contrast, while BAH, Ex-PD, Lean Startup, and IDEO also promote collaboration, they may prioritize other aspects such as financial analysis, technical development, customer feedback, or creative design, potentially at the expense of comprehensive cross-functional integration in NPD.

(4) Documentation and regulatory compliance: The Stage-Gate process stands out for its comprehensive documentation and structured approach to regulatory compliance. Its formalized stages and gates ensure that all necessary documentation is in place, making it easier to meet regulatory requirements and maintain a clear audit trail. In contrast, while BAH, Ex-PD, Lean Startup, and IDEO offer valuable approaches to

Table 2. New product development process framework specifications and requirements

Framework specifications	Framework requirements
New product development (NPD) framework stages involve lots of technical activities that play an important role in the successful conversion of product ideas into product function	Technical capabilities
Each activity in each stage is completed within the stipulated period	Product production at an appropriate time and cost
Smooth flow of stages, reduced reverse flow, easy to follow	A clear and common understanding of the NPD process
Framework stages emphasize up-front homework	Up-front homework
All new product development framework stages meet consumer needs by the exploitation of new, unique technology	Use of unique technology
Framework stages successfully converted product ideas into product functions quickly and on time	NPD process speed
Deliverables of framework stages include a clear definition of the functions of the product developed	Clear definition of the functions of the product
NPD activities in the framework stages followed all necessary quality standards	Implementation of quality standards
NPD activities in the framework stages followed all necessary regulatory practices	Regulatory practice
Deliverables of framework stages include a clear definition of the functions of the product developed	Clear definition of the functions of the product
NPD activities in the framework stages followed all necessary quality standards	Implementation of quality standards
NPD activities in the framework stages followed all necessary regulatory practices	Regulatory practic
Framework stages emphasize market research and customer involvement	Market research and customer involvement
Deliverables of framework stages include a clear definition of the target market	Clear definition of the target market
Each activity in the framework included a focus on the customer	Focus on the customer
Objectives and deliverables of the NPD framework based on appropriate marketing strategy	Appropriate marketing strategy
Market size defines the objectives of the framework stages	Market size
Senior management provides internal legitimacy and momentum for the new idea and concept	Senior management commitment and involvement
The framework involved cross-functional teams, i.e., marketing, purchasing, sales, after-sales service, design, and manufacturing	Involvement of cross-functional teams
The framework stages used structured new product development activities	Use of structured NPD process
Human factors such as Experience and a dedicated team are used for the successful conversion of product ideas into product function	Dedicated team members with relevant experience in the NPD process and activities
The framework perceives NPD as a strategy for the long term	Long-term vision and strategy
NPD goals that are clearly outlined and visible across the company	Presence of clear goals and milestone measurement
The framework process included effective internal communication with team members and management by properly linking activities	Effective internal communication among team members and management
NPD activities defined in the framework stages support an entrepreneurial culture in the organization	Entrepreneurial culture in the organization
Framework stages aligned with strategy	Alignment of NPD process activities with strategy
Human factors such as teamwork, cooperation, support, and guidance used for the successful conversion of product ideas into product function	Cooperation, support, and guidance within the team

(Cont'd...)

Table 2. *(Continued)*

Framework specifications	**Framework requirements**
Framework stages support teamwork by maintaining the proper organizational environment	Organizational environment to support teamwork
Each stage of the framework is passed with Go/Kill decision points in the process	Go/Kill decision points in the process
Framework stages considered the time of replacement of the product, considering the product life cycle	The time of replacement, considering the product life cycle
Framework stages utilized the research and development budget for NPD activities	Research and development budget for NPD activities
Project scheduling and monitoring activities are considered as NPD activities in the framework stages	Project scheduling and monitoring
Framework stages support teamwork by maintaining an innovative climate and culture	An innovative climate and culture
Framework stages ensured the availability of project/NPD resources and their management	Availability of project/NPD resources and their management
Framework stages involved innovating ideas in their activities	Innovate idea generation by expert groups
Developing and launching a product within the proper time frame	The right time to launch
The framework stage declared product scoring through the competitor benchmarking tool	Product scoring through benchmarking (competitor)
The framework stage declared product scoring through the internal benchmarking tool	Product scoring through benchmarking (internal)
NPD framework stages require financial resources that play an important role in the successful conversion of product ideas into a product function	Availability of financial resources
NPD framework stages involve lessons learned from past projects that play an important role in the successful conversion of product ideas into product function	Applying lessons learned from past projects
The framework filtered the activities after launch, and lessons learned were captured	Refining a product after launch and having a long-term view
Development of a product within the proper time frame	Time to market
NPD framework stage activities are defined considering social responsibility	Social responsibility
NPD framework stage activities involved statutory and environmental compliance	Statutory and environmental compliance
NPD framework stage activities defined considering Cultural competence	Cultural competence

NPD, they may lack the detailed documentation and formal compliance framework that Stage-Gate provides, potentially complicating regulatory compliance and audit processes.

(5) Specific needs and characteristics of the industry: The Stage-Gate process offers a structured and adaptable framework that can be tailored to meet specific industry needs in NPD, particularly in regulated sectors requiring rigorous risk management, quality assurance, documentation, and cross-functional collaboration. While BAH, Ex-PD, Lean Startup, and IDEO offer innovative approaches to NPD, they may not inherently address the complex requirements and challenges unique to specialized industries without additional customization and integration of industry-specific practices and standards.

A total of 60% of all investigated NPD functions implemented some form of the Stage-Gate process to enhance product innovation (Adams-Bigelow, 2005; Griffin, 1997; Kahn et al., 2012). The implementation of Stage-Gate frameworks provides a top-level overview to facilitate decision-making at key review points, dividing the overall process into more manageable stages to direct information-generating tasks (Phillips et al., 1999). The Stage-Gate process is characterized by low risk, immediate rewards, and a focus on incremental projects (McDermott & O'Connor, 2002). According to Harmancioglu et al. (2007), the level of competition in the industry is directly correlated with the implementation of formal stage gate processes, and Hamidizadeh et al. (2018) highlighted that the Stage-Gate model is the most famous model of NPD. The research findings, especially concerning the

Indian automobile sector, support existing literature, confirming the Stage-Gate framework's suitability for industries that prioritize meticulous planning and control. Consequently, the next section investigates identifying the crucial framework requirements and specifications essential for developing the NPD process framework as outlined in this study.

2.1. NPD Process Framework Specifications and Requirements

The NPD process has gradually developed from the implementation of product design and development activities. Hence, the design of the NPD process is greatly impacted by various elements, including technology, maturity, management strategies, business objectives, policies, culture, beliefs, and more. Manufacturing companies typically adjust or utilize current PD procedures based on the operational environment, regulations, and policies in the real world (Yin & Zhang, 2021). Successful companies demonstrate increased awareness of customer needs, prioritize marketing and advertising efforts, excel in PD, actively seek external expertise, and place trust in responsible and experienced employees. Integration of technical, commercial, organizational, marketing, and social factors can be deemed as an important multifunctional management mechanism that boosts the accumulated knowledge exchange, in the success of NPD in all aspects (Kadwe et al., 2017). Effective and repeatable NPD demands a balance between strategic effectiveness, functional excellence, and operational competence (Connell et al., 2001). The framework specifications and requirements, shown in Table 2, are derived from the literature based on seven dimensions delineating NPD success, such as strategy, research, commercialization, NPD process, project climate, company culture, and metrics and performance measurement. These dimensions are shaped by benchmarking studies, aiming to identify the best NPD process framework, expecting companies to adopt and sustain them (Adams-Bigelow, 2005; Barczak et al., 2009; Cooper et al., 2002; Cooper et al., 2004a; 2004b; 2004c).

3. Review on The Importance of The Stage-Gate Approach For NPD in Manufacturing Industries

The Stage-Gate development framework is applied within companies to streamline their PD processes, facilitating the efficient transition of new products from conception to market launch. This integrated approach combines project management principles with necessary processes for product realization, serving as a conceptual model increasingly adopted by organizations to mitigate challenges such as performance issues, rising development costs, and delays, thus minimizing risks associated with NPD. Literature is available on the importance of the Stage-Gate approach for NPD as a best practice in different manufacturing industries. Top-performing companies excel in NPD with an average success rate of 60.2%, while those in the bottom 20% struggle with over 3.5 times the failure rate, highlighting the importance of effective NPD management, leading many firms to adopt Stage-Gate processes as blueprints (Cooper, 1990; Smolnik & Bergmann, 2020). Several organizations have adopted phased-review workflow procedures that cover multiple functions to improve the development of new products. One widely accepted method in this regard is referred to as the "Stage-Gate" process (O'Connor, 1994). Griffin (1997) discovered that a majority of the NPD functions examined utilized a Stage-Gate process to enhance product innovation in manufacturing industries. Barringer & Gresock (2008) found that the acceptance rate of the stage-gate model in the United States industrial sector, currently at 73 percent, emphasizes the valuable advantages of utilizing a solid conceptual model for industrial NPD processes. According to Pietzsch et al. (2009), the primary development model utilized in the medical device industry is the stage-gate process. When the Stage-Gate process was initially introduced, numerous companies such as Procter and Gamble, Polaroid, The Royal Bank of Canada, Lego, and Shell adopted either certain components or the entire process (Broum et al., 2011). Without a Stage-Gate model, managing cross-functional communication, stability, and the efficient handling of NPD processes for a company as large as Ericsson while maintaining synergy between Stage-Gate, lean, and agile processes would be extremely challenging (Davoodi & Aslanzadeh, 2014). According to the findings by Wuest et al. (2014), implementing the Stage-Gate model for manufacturing and assembly processes in industry suggests that an adapted version of this model can significantly support product and process quality improvement. It has been determined that the Stage-Gate model is utilized by 70 – 85% of the prominent companies in the United States to oversee the complete journey of developing and introducing new products or services to the market (Stosic & Milutinovic, 2014). The Stage-Gate system is widely recognized as an efficient tool utilized often by leading companies and is currently regarded as the norm for a structured NPD process in the present market (Kazimierska & Grębosz-Krawczyk, 2017). To address market fluctuations and uncertainty, some companies have adopted a hybrid approach using Stage-Gate and Agile Scrum to reduce time-to-market and respond more swiftly to changing customer requirements (Eljayar & Busch, 2021).

4. Usage of The Stage-Gate Approach in Manufacturing Industries

According to the literature, the Stage-Gate approach is very important for NPD in various manufacturing industries as it enhances efficiency and minimizes the risk of failure, whether it is for launching new products or technologies that can reshape competitive positioning, introducing new products to boost company revenue, or defending market share with significantly improved products. The Stage-Gate model provides a strategic and practical guide for the process of developing new products, starting from the initial idea to the final launch, acting as a blueprint for managing the innovation process to improve productivity and performance. Cooper's Stage-Gate model, the primary focus of this study, can be considered an essential example of this era. The Stage-Gate method divides the innovation process into stages that involve coordinated, cross-functional, and simultaneous tasks, with each stage commencing at a checkpoint accountable for ensuring quality and making Go/Kill/Hold/Recycle decisions (Stosic & Milutinovic, 2014). A typical Stage-Gate model is illustrated in Fig. 2 (Cooper, 1990).

Each stage of the diagram represents a component of the PD process, encompassing a group of activities, while each gate serves as a review point for the preceding stage, where decisions are made based on the information generated. This framework enables the organization to enhance output quality by focusing on the process itself, eliminating non-value-added activities, and reducing risks associated with PD. Many organizations use a type of Stage-Gate process to guide their PD. Table 3 summarizes the Stage-Gate approach implemented in different industries for developing new products. Although similar, each has its own uniqueness in its implementation.

The NPD models of different organizations, as shown in Table 3, show both similarities and differences. The number of stages ranges from four to 11, and the number of gates ranges from 4 to 10. The quantity and titles of the stages might differ among organizations, but they all share a common underlying goal. Based on the number of stages and gates, the process is classified into high Stage-Gate and low Stage-Gate approaches. In most of the processes, the stage numbers are exactly the same or more or less equal to the number of gates. Most of the stages and gates are common for all the models. In some models, the number of activities from idea generation to launch of the product is divided into sub-activities to make the NPD simpler and more effective, which increases the number of stages and gates in the process. Each of the NPD process models mentioned above is unique and is implemented in its own distinct way by companies. All these various NPD process models provide a structured framework with standardized principles and methods for PD, ensuring timely and cost-effective market entry while meeting customer needs. Implementing a model promotes uniformity across an organization, supported by sound management decisions and effective risk management.

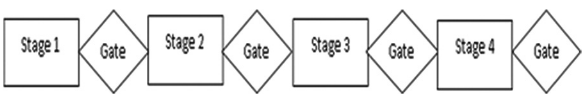


Fig. 2. The Stage-Gate approach

5. The Need to Develop A NPD Process Framework

Observed limitations of this benchmarked NPD process framework are: (i) very rigid and bureaucratic, (ii) stage-wise lead roles are not defined, (iii) all activities from the first stage to the last stage and their one-to-one proper integration is not shown, (iv) lack of functional integration between departments, (v) inadequate flexibility for IT Integration, (vi) lack of feedback process, (vii) lack of involvement of appropriate stakeholders, (viii) lack of generational learning, (ix) lack of organized and structured data, (x) missing management processes, (xi) not designed for rapidly growing and uncertain conditions of the market, (xii) not designed for higher risk initiatives, (xiii) required inputs for each stage are not clearly defined, (xiv) stage-wise deliverables are not clearly defined, (xv) most of the models are based on the traditional Stage-Gate process, and (xvi) coordination complexity.

Launching new and innovative products into the market quickly, cost-effectively, and with minimal risk is essential to meet the targeted demands; companies are consistently upgrading from existing PD models to newer ones that are characterized by increased agility, flexibility, and alignment with their company's structure and operations (Munoli, 2017). Shorter life cycles, changing portfolios of new product opportunities, and associated risks continue to pressure the NPD teams to produce a wider range of products. Shepherd & Ahmed (2000) observed that in many companies, primarily small and medium-sized, the way products are developed is completely unstructured. There is no steady terminology; each company uniquely defines its NPD process framework, even though many are similar. An inconsistent NPD process framework leads to wasted effort, misdirected work, more clarification meetings, failure in estimating resource requirements and schedules, excessive task interdependence, and fire-fighting. To navigate these challenges successfully and efficiently, it is

Table 3. Benchmarking of new product development process frameworks followed by industries (Chao & Ishii, 2005; Phillips et al., 1999; Williams, 2008; Loganathan & Jayakrishnan, 2014)

Company name	Process title	High/low stages	Number of stages	Number of gates
General motor	Product development process	Low	4	-----
Tata motor	New product introduction	High	7	8
Ashok leyland	Genmod	High	6	-----
Daimler's	Commercial vehicle development system	High	8	7
Mitsubishi fuso	Product creation process	High	8	7
Volvo'	Global development process	Low	5	5
Malaysian small-sized automaker	Product development process	Low	4	8
Mahindra and Mahindra	Mahindra product development system	Low	5	13
Renishaw	Renishaw new product development process	High	10	10
ABB	ABB gate model	High	6	7
GE	New product introduction	High	9	10
Lucent	New product introduction	Low	4	8
Rover	Project management guidelines	High	8	-----
BMW	Gateway in new product development	High	7	-----
Chrysler	New product development strategy	Low	4	-----
Ford	Global product development system	High	11	-----
Honda	Programmed milestone philosophy	High	8	-----
IBM	-----	Low	5	-----
Xerox	-----	Low	3	-----
Lucas	Product introduction management	Low	5	-----
Motorola	-----	Low	6	-----
Renault	Project management system	High	6	-----
Toyota	Generic development process	High	9	-----
NASA	Technical design review process	Low	5	-----
Whirlpool Corporation	C2C product creation process	Low	4	4

crucial to focus on systematic screening, monitoring, and progression frameworks. The structured and documented approach provides a clear roadmap for successful NPD (Owens & Cooper, 2001). To remain competitive, best-in-class companies carefully select and use the basic attributes of an effective NPD framework and try for continuous improvements on multiple fronts to retain the leadership position (Griffin, 1997). Implementation of an effective NPD framework improves planning and decision, technology usage, evaluation at key milestones, overhead and labor costs, quality of goods and services, creativity and innovation, need for engineering and design changes, ability to penetrate new markets, revenue and margins, inventory cost, and so on (Shepherd & Ahmed, 2000). Managers managing the introduction of new products within the manufacturing engineering department, as well as other tasks or stages in NPD, face growing pressure to enhance the efficiency of the process. The nature of progress required differs between organizations. The most common and widely cited improvement objectives are as follows: reduced costs and cycle time and increased market share and product quality. In the next section, a structured form is developed to gather the inputs of industrial experts to address the highlighted limitations of existing NPD models and to meet the needs of NPD experts. The initial section gathers fundamental background details regarding the industry and participants (Table 5), while the subsequent section gathers feedback from the participants against the questionnaire (Table 6) based on the scale matrix (Table 4), highlighting the important measures of NPD process frameworks.

5.1. Empirical Investigation to Identify The Need to Address The Modifications of A NPD Framework

Interviews were performed to verify the acceptability and practical use of the NPD process framework specifications and requirements. The interviews were structured as follows: Specific questionnaires for best NPD practice, selection of participants, collection of data, organization of data collected, and presentation and discussion of the results. A set of 37 questions was formed as the foundation for modifications of a PD framework. After answering each of the questions, the average is calculated. High score points are considered as acceptance of framework requirements by industrial experts and are an input in design and development activities for the proposed new theoretical NPD process framework for the automobile Industry.

The selected respondents (Table 5) went through the defined questions (Table 6) and were assigned a scale referring to Table 4.

On reviewing the data presented in Table 6, it is apparent that all survey participants agreed that the NPD process framework for automobile companies should include identified specifications and requirements in design and development activities. In light of this, the proposed NPD process framework effectively addresses the concerns expressed by the participants, as detailed in the following section.

Table 4. Scale matrix

Scale	Definition
5	Strongly agree
4	Agree
3	Neutral
2	Disagree
1	Strongly disagree

6. Proposed NPD Process Framework for The Automobile Industry

Fig. 3 depicts the proposed NPD process framework model, distinguishing between the stages and the gates for the automobile industry. The proposed NPD process framework consists of eleven Stages and Gates across the NPD. The development routines vary from one stage to another. In each stage, there are significant, related main activities that are progressing in parallel with the process as described in Tables 7-17. For each stage, there are gateway(s) indicating activities of monitoring and controlling the ongoing development process. Tables 7-17 show the descriptions of the activities involved in each of the Stages, while Fig. 4 to Fig. 14 shows Stage-Gate NPD activities integration in the form of frameworks.

Each of these eleven stages and gates, as shown in Tables 7-17, is explained as follows;

(1) Stage one, named market research and concept inception, includes the lead role of the marketing department, supported by other departments such as product planning, finance and budgeting, design and development, human resources, launch planning, and the core team. In this stage, an idea is generated after the study and compilation of customer data or customer requirements. The objective of stage one is to complete twenty-one activities. The deliverables of stage one are reviewed in gate one, named idea review, where targets are proposed, and an assessment of affordable alternatives is weighed, including program strategy, product vision, product content, and program goals.

(2) Stage two, named concept verification, includes the lead role of the design department, supported by other departments such as marketing, product planning, development, and the core team. In this stage, the design team develops design features and characteristics after reviewing

Table 5. Profile of respondents

No.	Designation of respondent	Type of company	The sector of the company	Product type
1	Retired manager	Multinational	Automobile/Ancillaries	Highly standardized
2	Vice president	Domestic private	Electrical and Electronics	Standard with custom options
3	Manager	Domestic private	Automobile/Ancillaries	Highly standardized
4	Manager	Domestic private	Automobile/Ancillaries	Standard with custom options
5	CEO	Domestic private	Others	Highly customized
6	GM plant head	Domestic private	Automobile/Ancillaries	Highly standardized
7	Plant head	Domestic private	Automobile/Ancillaries	Highly standardized
8	Director	Multinational	Others	Standard with custom options
9	Manager	Multinational	Automobile/Ancillaries	Highly standardized
10	Manager	Domestic private	Automobile/Ancillaries	Standard with custom options

Table 6. New product development process survey result analysis

Question	**How important is the question? average score**
There should be a process for undertaking portfolio management	4.6
New product development (NPD) goals should be clearly defined and visible within the company	4.9
The company should consider NPD as a long-term strategy	4.6
The mission and strategic plan should help to define strategic arenas for new opportunities	4.8
NPD goals should clearly align with the company's mission and strategic plan	4.8
Projects in a portfolio that should be aligned with the NPD strategy	4.6
NPD projects and programs should be reviewed on a regular basis in the company	4.8
Shall Opportunity identification be an ongoing process? and can redirect the strategic plan in real-time to respond to market forces and new technologies	4.5
There should be a ranking or prioritization of NPD projects	4.6
There should be a consideration for balancing the number of projects and available resources	4.7
Concept, product, and market testing should be consistently undertaken and expected with all NPD projects	4.6
Research should be high if any market research is undertaken	4.6
Customer/user should be an integral part of the NPD process	4.7
Studies of customers and users should be focused on both current and future customer needs and problems	4.7
The company should avoid changing marketing budget decisions dramatically and up to the point of launch	4.0
The launch team should be cross-functional in nature	4.7
Cross-functional teams should make decisions concerning manufacturing, logistics, marketing, and sales	4.6
A project post-mortem meeting should be held after the new product is launched	4.3
Commercialization should be a formal part of the NPD process	4.3
Go/No-Go criteria should be clear and pre-defined for each review gate	4.6
The NPD process should be flexible and adaptable to meet the needs, size, and risk of individual projects	4.6
The NPD process should be visible and well-documented	4.8
Information technology (IT) infrastructure with appropriate hardware, software, and technical support should be available to all NPD personnel	4.7
A clear NPD process should exist	4.7
The company should review projects at the point of completion	4.9
The core project team should work on the NPD project from beginning to end	4.8
Each project should have a clearly identifiable stage-wise project leader	4.7
There should be enough formal communication to properly coordinate NPD activities	4.6
The company should appear to have the right number of projects individually assigned to NPD personnel	4.5
NPD should be a Top priority of management	4.4
The company should actively work with customers to identify new product opportunities	4.8
All NPD ideas should welcome those that come from within and outside of the company	4.6
Management should not primarily be focused on operational efficiency and cost savings	4.0
There should be standard criteria for evaluating individual NPD projects	4.6
There should be standard criteria for evaluating individual NPD efforts	4.6
All NPD project evaluations should be stage-wise and by the CFT team	4.3
NPD projects should be killed before they reach launch if they fail to achieve the stage-wise target	3.3

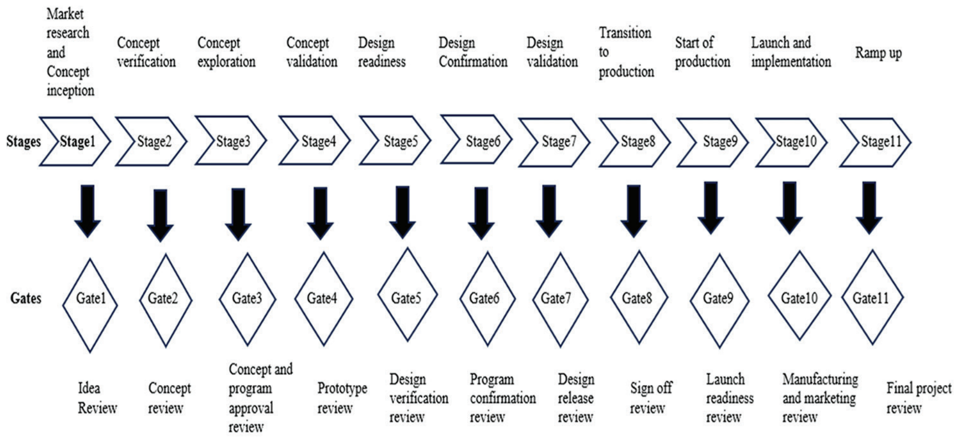


Fig. 3. The proposed new product development process framework model distinguishes between the gates and the stages for the automobile industry

engineering requirements and assessing potential manufacturing problems. The objective of stage two requires completing 20 activities. The deliverables of stage two are reviewed in gate two, named concept review, after gate one (Idea Review) approval, where overall targets are set (Performance, Functional, and Financial), the preliminary project plan is prepared, R&D funds get approved, manufacturing and purchasing opportunities and capacity are reviewed and confirmed, make versus buy decision and manufacturing location confirmed, and project feasibility is confirmed. If found not feasible, the closing report is preserved in the knowledge of the management repository.

(3) Stage three, named concept exploration, includes the lead role of the design department, supported by other departments such as marketing, product planning, development, core team, supplier, launch planning, finance and budget, and manufacturing engineering. In this stage, concept selection and concept analysis activities are completed by the design team. The objective of stage three requires completing 21 activities. The deliverables of stage three are reviewed in gate three, named concept and program approval review, after gate two (concept review) approval, where product specification and styling are frozen, program cost target is frozen, and project budget code is created, and quality and reliability targets are signed off.

(4) Stage four, named concept validation, includes the lead role of the design and development department, supported by other departments such as product planning, HR, supplier, launch planning, finance and budget, and manufacturing engineering. In this stage, design features and characteristics are finalized after reviewing engineering requirements and assessing potential manufacturing problems, and comprehensive and effective manufacturing systems are developed, ensuring that the manufacturing system meets customer requirements. The objective of stage four requires completing 20 activities. The deliverables of stage four are reviewed in gate four, named prototype review after gate three (concept and program approval review) approval, where based on the voice of the customer and customer requirements, technical specifications are developed in the form of PDB, safety data sheet and design input, and customer input requirements and preliminary drawings under PPRF number are released.

(5) Stage five, named design readiness, includes the lead role of the design department, supported by other departments such as product planning, HR, supplier, development, launch planning, and manufacturing engineering. In this stage, a detailed design is ready. The objective of stage five requires completing 15 activities. The deliverables of stage five are reviewed in gate five, named as design verification review, after gate four (Prototype Review) approval, where based on the EP build and test experience, drawings are released for procurement of new parts under PPRF number and final BOM is prepared thus confirming the designs for further activity of facility and process planning.

(6) Stage six, named design confirmation, includes the lead role of the design and development department, supported by other departments such as product planning, HR, supplier, launch planning, and manufacturing engineering. The objective of this stage is design analysis, simulation, and manufacturing planning (FTG planning). The objective of stage six requires completing 14 activities. The deliverables of stage six are reviewed in gate six, named

Table 7. Stage 1 and Gate 1 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 1 – Market research and concept inception	Objectives- (1) To define an idea (2) To determine customer needs (3) To define a quality program		
Gate 1 – Idea review	Objectives – Based on inputs from the voice of the customer (VOC), Market Research, High Priority Customer Wants, and Competitive Benchmark, targets are proposed, and an assessment of affordable alternatives is weighed. The idea review gateway includes (1) Program strategy, (2) Product vision, (3) Product content, and (4) Program goals.		
NPD activities	Lead role	Inputs	Deliverables
(1) Market research (2) Market survey of existing products and similar products (3) Decide Market Segment and Volume (4) Refer J. D. Power (Consumer Survey) and related inputs in the similar segment (5) Identify competitor product (6) Decide benchmark and benchmark testing (7) Receive strategy from management (8) Collate the voice of the customer (VOC) (9) Convert VOC into customer requirements (10) Initiate and obtain QFD (11) Identify competitor vehicle (12) Generation of Product Request Form (13) Collect data on- Market volumes, segment, competition, Design Input, Customer Inputs, Business intent, Business Environment, Financials, Selection of partners, Resources, etc. (14) Decide benchmark and conduct benchmarking studies (15) Creation of ideas (16) Development of the concept based on the requirement request (17) Search for Patents	(1) Product planning/Program management (2) Top management (3) Customer (4) Marketing (5) Finance and Budgeting (6) Design and Development (7) Human resource (8) Core team (9) Launch planning service and training	(1) Voice of the Customer (2) Business Plan/Marketing Strategy (3) Product/Process Benchmark Data (4) Product/Process assumptions (5) Product Reliability studies (6) Customer inputs	(1) Product intent (2) Direction from Top Management (3) Collation of Voice of Customer (Corporate, Regulatory, Final Customer, Manufacturing etc.) (4) Market intelligence inputs considered (5) Limited CFT in place (6) QFD analysis leading to Concepts (7) Forecasting techniques (volume) (8) Product specifications outlined on the Product Request Form (9) Required Job1 date indicated (10) Overall “Boundary levels Targets” for Costs, Weights, and Investments set

(Cont’d...)

Table 7. *(Continued)*

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
(18) Product Portfolio analysis (19) Collect data for the Business Case (20) Gateway review meetings (21) Register as a new development program			

Abbreviations: CFT: Cross-functional team; J1: Job 1; PRF: Product Request Form; QFD: Quality function deployment; VOC: Voice of customer.

Table 8. Stage 2 and Gate 2 new product development activities

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
Stage 2 – Concept verification	Objectives- (1) To develop design features and characteristics. (2) To critically review engineering requirements and assess potential manufacturing problems, if any		
Gate 2 – Concept review	Objectives – Based on Gate 1 (Idea Review) approval, (1) Overall “Targets” are set (Performance, Functional and Financial), (2) Preliminary project plan prepared, (3) R&D funds are approved, (4) Manufacturing and purchasing opportunities and capacity reviewed and confirmed – make v/s buy decision and manufacturing location confirmed,		

(Cont’d...)

DOI: 10.6977/IJoSI.202506_9(3).0004
B.G. Shinde, S.B. Sanap, et al./Int. J. Systematic Innovation, 9(3), 31-74 (2025)

Table 8. *(Continued)*

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
	(5) Product content is sufficiently detailed to drive the project and (6) Project feasibility is confirmed. If found not feasible, the closing report is preserved in the Knowledge Management Repository.		
NPD activities	Lead role	Inputs	Deliverables
(1) Refer to QFD and translate to technical specifications (2) Design Review Concept – DR0 (3) Creation of Unique Selling Points (USPs) (4) Styling options and Concept layout completed (5) Buck components made; Static buck build started – “Mock” ready for approval (6) Buck evaluation (7) Define the scope of the project (Product and Process – Vehicle Level Assumptions), initiate PDB/VDS/SDS development (8) Identify Technology options (Concept Ready and Implementation Ready) (9) Identify powertrain options (10) Identify Regulatory and Product Liability requirements (11) Refer to Lessons learned from previous projects (12) Evaluate the feasibility and revision of the Product Request Form (13) Identify consultants, as applicable (14) Check manufacturing feasibility, competency, and capacity (In-house/vendor) (15) Declare the project team (ensure representation from essential functions) (16) Initiate timing plan (17) Make/buy decision (18) Prepare a business case with a sensitivity analysis (19) Prepare project budget (20) Gateway review meetings	(1) Customer (2) Management (3) Marketing (4) Product design (5) Order management (6) Program Management	(1) Voice of the Customer (2) Business Plan/Marketing strategy (3) Product/Process Benchmark Data (4) Product/Process assumptions (5) Product Reliability studies (6) Customer inputs	(1) Concept (layout) frozen (along with SWOT analysis), Product Request Form Sign-off (2) Manufacturing Musts and Wants Submitted (3) Functional Image targets set based on signed-off Product Request Form (4) Engine, transmission, and Axle options selected (5) Consultants, if any, are identified (6) DR0 (Concept) conducted and outputs recorded (7) Project/Program Team declared; Prod. Launch Manager Identified (8) High Impact-High Risk vendors identified and are on board (9) Mock-up approved (10) Preliminary Project Plan outlined, including firm Job1 date (11) Project key deliverables frozen, including reusability targets (12) Assumptions for PDB (Product Description Book), Concept Styling, Functional Image Targets evolved (13) Product Cost target, Manufacturing Location, make v/s Buy Decision, volume commitment, including cannibalization, decided (14) Business Case details collated, including sensitivity analysis (15) Investments, IRR and Business Case Approved

Abbreviations: DR: Design review; IRR: Internal rate of return; NPD: New product development; PALS: Product attribute leadership strategy; PCF: Process flow chart; PDB: Product data book; SDS: System design specification; SWOT: Strengths, weaknesses, opportunities, and threats; TCE: Teamcenter engineering; VDS: Vehicle design specification.

DOI: 10.6977/IJoSI.202506_9(3).0004
B.G. Shinde, S.B. Sanap, etc./Int. J. Systematic Innovation, 9(3), 31-74 (2025)

Table 9. Stage 3 and Gate 3 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 3 – Concept exploration	Objectives- 1. Concept selection, 2. Concept analysis		
Gate 3 – Concept and Program approval review	Objectives – Based on Gate 2 (Concept review), the program is approved. 1. Product specification and styling are frozen 2. The Program Cost target is frozen, and the project budget code is created 3. Quality and Reliability targets signed off		
NPD activities	Lead role	Inputs	Deliverables
(1) Sign-off targets for functional image, cost, quality, durability, reliability, warranty, time, etc. (2) Decide reusability strategy for facilities, parts, technology, resources, etc. (3) Identify and release hardpoints (4) Identify and complete RWUP (5) Finalize Product Specifications (VDS and SDS, as applicable) (6) Ensure EP readiness of Powertrain (for new development) (7) Internal and external (supplier) manufacturing and assembly objectives consistent with the timing, resource, and investment targets of the vehicle (8) Complete PDB (Product Description Book) and sign-off (9) Make/Buy decision (10) Identify the scope, budget, resource plan, and timing for verification/validation (11) Verify feasibility for CMVR/Homologation and safety (12) Prepare development timing plan and Work Breakdown Structure (WBS) (13) Identify the critical activity path (14) Complete and present Project Budget, Resource Plans, Timing Plans, Deliverables, Benefits, and Business Case for Approval of Corporate Projects (15) Preliminary decision on allocation of manufacturing responsibilities (16) Investigation of necessary production capacity (17) Initiate DFA/MA analysis, as applicable (18) Go ahead/Hold decision for design (19) Creation of budget code (20) Lessons learned recorded (21) Gateway review meetings conducted	(1) Product design (2) Product development (3) Program Management	(1) Voice of the Customer (2) Business Plan/ Marketing Strategy (3) Product/Process Benchmark Data (4) Product/Process Assumptions (5) Product Reliability studies (6) Customer inputs	(1) Define and complete RWUP (2) Quality and Reliability Targets Signed Off (3) Project budget approved by Corporate Projects and budget code created (4) Engine, Transmission and Axle EP ready (if under development) (5) Cost targets for VDS and SDS defined (6) Baseline Plan including Work Breakdown Structure/PDB Signed Off (7) Lessons Learnt Recorded

Abbreviations: APQP: Advanced product quality planning; CMVR: Central motor vehicle regulation; CP: Clay prototype; DFA: Design for assembly; DMA: Design for manufacturing.; EBOM: e-Bill of material; EP: Engineering prototype; FTG: Facility tooling gauge; HIHR: High-risk high impact; IPR: Intellectual property rights; P&H: Packaging and hard point; PALS: Product attribute leadership strategy; PDB: Product data book; PR: Prototype request; RFQ: Request for quotation; RWUP: Real-world usage pattern; SAM: Serviceability, accessibility, and maintainability; SDS: System design specification; SOR: Statement of requirement; TCF: Teamcenter engineering; VDS: Vehicle design specification; WBS: Work breakdown structure.

Table 10. Stage 4 and Gate 4 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 4 – Concept validation	Objectives- 1. To finalize design features and characteristics. 2. To critically review engineering requirements and assess potential manufacturing problems, if any 3. To develop a comprehensive and effective manufacturing System. 4. To ensure that the manufacturing system meets customer requirements		
Gate4 – Prototype review	Objectives – Based on the VOC and Customer Requirement (CR) converted to technical specification in the form of PDB, SDS, Design Input, and Customer Input Requirement, the design teams prepare the preliminary drawings under the PPRF number		
NPD activities	Lead role	Inputs	Deliverables
(1) Preparation of preliminary Bill of Materials and add-delete list (2) Cascade vehicle design specifications to system and component levels. Document SDS and CDS. (3) Initiate and complete DFMEA with recommended actions (4) Vehicle, system, and sub-system level layouts ready (5) Develop and release System and Component level designs (PPRF drawings) for EP (6) MP level build and MP testing (7) Record MP learnings (8) Design Review (Product Design) – DR1 (9) Verify specifications and update PDB based upon SDS	(1) Customer (2) Management (3) Marketing (4) Product design (5) Order management (6) Program Management	(1) Design goals (2) Reliability and quality Goals (3) Preliminary Bill of Material (4) Preliminary process flow chart (5) Preliminary listing of special product and process Characteristics (6) Product Assurance Plan (7) Management support (8) Design failure mode and effects analysis (DFMEA) (9) Design for manufacturability and Assembly (10) Design verification	(1) Availability of preliminary BOM/ “Add-Delete” list (2) Cascading of VDS - SDS – CDS (3) Develop Customer Input Requirements (4) MP preparation (if required) (5) Design Review (Product)-DR1 and recommendations available (6) DFMEAs Initiated and Completed with recommended actions (7) QFD - system level completed for the identified system (8) Vehicle-level design ready

(Cont’d...)

DOI: 10.6977/IJoSI.202506_9(3).0004
B.G. Shinde, S.B. Sanap, etc./Int. J. Systematic Innovation, 9(3), 31-74 (2025)

Table 10. *(Continued)*

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
(10) EP development planned (11) Customer Input Requirement finalization (include performance, functional, appearance, regulatory, reliability, capacity, serviceability, manufacturability, warranty, etc., targets) (12) Prepare concept model for CAE (13) Selection of high-risk-high Impact Suppliers (Sourcing decision) (14) Release specific Customer Input Requirements to the supplier (15) Initiation of APQP at HI-HR suppliers (Prepare Quality Plan) (16) System-level and Component-level DVP completed (17) CAE Analysis complete		(11) Design review (12) Prototype Build – Control Plan (13) Engineering drawings (Including Math Data) (14) Engineering Specifications (15) Material specifications (16) Drawing and specification changes (17) New Equipment, Tooling, and Facilities Requirements (18) Special Product and Process Characteristics (19) Gages/Testing Equipment Requirements (20) Team Feasibility, commitment, and management support	(9) Systems and subsystem-level layouts are ready (10) SDS, CDS level designs ready for release (11) System Level DVP (12) Component Level DVP ready (13) Completion of CAE Analysis (14) MP learnings recorded (15) Long lead tooling and Facilities identified (16) Sourcing Decision [HI-HR] made and APQP initiated for suppliers and in-house. (17) PPRF released for new parts. (18) Preliminary BOM

Abbreviations: APQP: Advance product quality planning; BOM: Bill of material; CAE: Computer-aided engineering; CDS: Component design specification; CIR: Customer input requirement; CP: Clay prototype; DFMEA: Design failure mode effect analysis; DMU: Digital mock-up; DR: Design review; DVP: Design verification plan; EP: Engineering prototype; FTG: Facility, tooling, gauge; HIHR: High-risk high impact; MIMR: Medium risk medium impact; MP: Mechanical prototype; PPRF: Prototype request form; QFD: Quality function deployment; RFQ: Request for quotation; SAM: Serviceability, accessibility, and maintainability; SAP: System applications and products in data processing; SDS: System design specification; VDS: Vehicle design specification; VP: Validation prototype; WBS: Work breakdown structure.

Table 11. Stage 5 and Gate 5 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 5- Design readiness	Objectives – 1. Detailed design readiness		
Gate 5 – Design verification review	Objectives – Based on the EP build and test experience, drawings are released for procurement of new parts under the PPRF number, and the final BOM is prepared, thus confirming the designs for further activity of facility and process planning		
NPD activities	Lead role	Inputs	Deliverables
(1) Updating technical specifications and features w. r. t. current competition and new market scenario (Review QFD) (2) Driveline performance analysis (3) Updating drawings to conform to design guidelines (4) Identify SC/CC (5) Final Engineering bill of materials ready in metaphase (6) Develop Appearance including colour and trim (7) Design Review DR-2 (Process Design) (8) Update CAE Analysis – strength, durability, and kinematics (9) Price negotiation/Finalize price (10) Release of PO (11) Receive aggregates/components for EP (12) Product launch manager identified for M2, M4 projects (13) EP build as per prototype control plan (14) Initiation of Process design and Process FMEAs (15) Identify resource requirement	(1) Product design (2) Product development (3) Program Management	(1) Design Goals (2) Reliability and Quality Goals (3) Preliminary Bill of Material (4) Preliminary Process Flow Chart (5) Preliminary Listing of Special Product and Process Characteristics (6) Product Assurance Plan (7) Management support (8) Design Failure Mode and Effects Analysis (DFMEA) (9) Design for Manufacturability and Assembly (10) Design Verification (11) Design Reviews (12) Prototype Build – Control Plan (13) Engineering Drawings (Including Math Data)	(1) Availability of preliminary BOM/"Add-Delete" list (2) Cascading of VDS - SDS – CDS (3) Develop Customer Input Requirements (4) MP preparation (if required) (5) Design Review (Product)-DR1 and recommendations available (6) DFMEAs Initiated and Completed with Recommended actions (7) QFD – system level completed for identified system (8) Vehicle-level Design ready (9) Systems and subsystem level layouts ready (10) SDS, CDS level designs ready for release (11) System Level DVP (12) Component Level DVP ready

Abbreviations : AIAG: Automotive industry action group; APQP: Advance product quality planning; BOM: Bill of material; CAE: Computer-aided engineering; CC: Critical characteristics; CDS: Component design specification; CIR: Customer input requirements; CP: Clay prototype; DFMEA: Design failure mode effect analysis; DR: Design review; DVP: Design verification plan; DVP&R: Design verification plan and report; EBOM: e-Bill of material; ELV: End of life of a vehicle; EP: Engineering prototype; FMEA: Failure mode effect analysis; FTG: Facility, tooling, gauge; M2, M4 Projects: Minor and major modification projects; MP: Mechanical prototype; PFMEA: Process failure mode effect analysis; PO: Purchase order; QFD: Quality function deployment; SAP: System applications and products in data processing; SC: Significant characteristics; SDS: System design specification; TCE: Team center engineering; VDS: Vehicle design specification; VP: Validation prototype; Ys YCS: Failure modes.

Table 12. Stage 6 and Gate 6 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 6 – Design confirmation	Objectives- 1. Design analysis and simulation 2. Manufacturing planning (FTG planning)		
Gate 6 – Program confirmation review	Objectives- 1. Final design release 2. Long lead tooling signs off		
NPD activities	Lead role	Inputs	Deliverables
(1) Marketing Launch team in place (2) EP concern resolution and changes incorporated in drawings (3) EN drawings released for VP build (4) Preparation of Installation drawings (5) Identified vendors' commitment preparedness for J1 (All vendors signed off) (6) Select/Design, Develop, and Procure new Facilities, Tools and Gauges (7) Prepare Process Flow Chart based on FTG, Floor Plan Layout (8) Process FMEA completed (9) Review resource requirement (10) Manufacturing Launch Team in place (11) Review and confirm Business case (12) Go ahead/Hold decision for manufacturing and marketing (13) Lessons learnt recorded (14) Gateway review meeting	(1) Product planning/Program Management (2) Supplier (3) Manufacturing engineering/ Facility creation (4) Design and development (5) HR (6) Marketing (7) Launch planning/service and training	(1) Design goals (2) Reliability and Quality goals (3) Preliminary Bill of Material (4) Preliminary Process Flow Chart (5) Preliminary Listing of Special Product and Process Characteristics (6) Product Assurance Plan (7) Management support (8) Design Failure Mode and Effects Analysis (DFMEA) (9) Design for Manufacturability and Assembly (10) Design Verification (11) Design Reviews (12) Prototype Build – Control Plan (13) Engineering Drawings (Including Math Data) (14) Engineering Specifications (15) Material Specifications (16) Drawing and Specification Changes (17) New Equipment, Tooling, and Facilities Requirements (18) Special Product and Process Characteristics (19) Gages/Testing Equipment Requirements (20) Team Feasibility commitment and management support	(1) Revalidate Business Case assumptions and confirm Business Case (2) Go/Kill decision given (3) Identification and Deployment of additional and requisite resources (4) EN released for VP build (5) Team Feasibility Commitment (6) Preliminary Process BOM (7) All vendors signed off (8) Vehicle available for customer validation (9) Lessons Learnt Recorded

Abbreviations: AIAG: Automotive industry action group; APQP: Advance product quality planning; BOM: Bill of material; CMVR: Central motor vehicle regulation; DFMEA: Design failure mode effect analysis; DVP and R: Design verification plan and report; EN: Engineering notice; EP: Engineering prototype; FMEA: Failure mode effect analysis; FTG: Facility, tooling, gauge; J1: Job1; SAP: System applications and products in data processing; VP: Validation prototype.

Table 13. Stage 7 and Gate 7 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 7 – Design validation	Objectives- 1. Tooled up parts trial 2. Crash test 3. Aggregate test 4. Detail costing		
Gate 7 – Design release review	Objectives – 1. Design release 2. Long lead tooling signs off 3. Component reliability plan 4. Business case validation		
NPD activities	Lead role	Inputs	Deliverables
(1) Marketing launch budget prepared and strategy finalized (2) Dealer selection completed (3) Concern resolution for VP build trial (4) Design Review (DR-3) for Validation completed (5) Tuning of performance, Internal jury evaluation (6) Review the supplier PSW status (7) Tooled-up parts received (8) Application for CMVR Certification (and Homologation) prepared (9) Reliability, durability tests, and key life testing of VP vehicles start as per DVP (Vehicle and Rig) (10) Emission tuning trials (regulation) (11) Install new tools/facilities (12) Operator process instructions (SOP) prepared (13) Operator training initiated (14) Process (SAP) BOM available (15) Manufacturing process sheets initiated (16) VP build control plan completed (17) VP and/or static build ready from Make-like production (including timing verification, BOM Validation) (18) Production Validation Test plan ready (19) NOVA-C evaluation of EP vehicles with scores completed (20) Gateway review meeting conducted	(1) Product planning/Program Management (2) Supplier (3) Manufacturing engineering/ Facility creation (4) Design and development (5) HR (6) Marketing (7) Launch planning/service, and training	(1) Design Goals (2) Reliability and Quality Goals (3) Preliminary Bill of Materials (4) Preliminary Process Flow Chart (5) Preliminary Listing of Special Product and Process Characteristics (6) Product Assurance Plan (7) Management support (8) Design Failure Mode and Effects Analysis (DFMEA) (9) Design for Manufacturability and Assembly (10) Design Verification (11) Design Reviews (12) Prototype Build –Control Plan (13) Engineering Drawings (Including Math Data) (14) Engineering specifications (15) Material Specifications (16) Drawing and Specification Changes (17) New Equipment, Tooling, and Facilities Requirements (18) Special Product and Process Characteristics (19) Gages/Testing Equipment Requirements (20) Team Feasibility, commitment, and management support	(1) VP tooling and parts development (2) Review Supplier PSW status (3) Operator Training initiated (4) BOM Validation (5) VP and/or Static Build ready from "Make-like" Production Process (6) NOVA-C evaluation and scores for EP (7) Dealer selection done (8) Tooled-up parts received (9) Functional Image Targets met through Internal Jury (10) Design Review (Validation)-DR 3 (11) DVP for VP (12) Testing of VP vehicles commences (13) SAP BOM available (14) Application for homologation. (15) PV Test plans ready

Abbreviations: APQP: Advance product quality planning; BOM: Bill of material; CMVR: Central motor vehicle regulation; DFMEA: Design failure mode effects analysis; DR: Design review; DVP: Design verification plan; EP: Engineering prototype; FEU: Field evaluation unit; MSA: Measurement system analysis; NOVA-C: New overall vehicle audit – complete; PFMEA: Process failure mode effects analysis; PDB: Product data book; PO: Purchase order; PP: Production prove out; PSW: Part submission warrant; PTR: Production trail run; PV: Production validation; SAP: System applications and products in data processing; SOP: Standard operating process; VR: Validation prototype.

Table 14. Stage 8 and Gate 8 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 8 – Transition to production	Objectives- 1. To validate the manufacturing process 2. To ensure that customer requirements will be met		
Gate 8 – Sign off review	Objectives – Formal acknowledgment that all aspects of product design are completed as per intent and verified		
NPD activities	Lead role	Inputs	Deliverables
(1) Plan product profile (variant/configuration) (2) Plan for media launch (Selection of advertising agency) (3) Identify and initiate the development of a dealership (4) Concern resolution for VP test trial reports and incorporate changes (5) Update the CAE model based on durability results (6) Verify the part quality (7) Reliability, Durability tests, Key life testing, VP testing completed (Vehicle and Rig) (8) Tool and facility try-out and Production Trial Run completed (9) PP build initiated (10) Review Team feasibility commitment (11) Measurement system evaluation starts (12) Preliminary process capability study starts (13) Dealer development initiated (14) Gateway review meetings (15) Jury and Management evaluation for Functional Image (16) Production of promotional material (17) Production of sales aids (18) Preparation of sales person/centers (19) Reconfirm the CAE results (20) Re-affirm performance targets based on Jury evaluation (21) Verify specifications on performance (22) Engineering sign-off (Design and performance targets (23) Concern resolution for field trails (24) Facilitate supplier PPAP (25) Production trial run at supplier (26) Final component price available (27) Verify receipt of PSW for supplier provided parts	(1) Product planning/ Program Management (2) Supplier (3) Manufacturing engineering/ Facility creation (4) Design and development (5) HR (6) Marketing (7) Launch planning/ service, and training (8) Finance and Budgeting	(1) Packaging Standards (2) Product/Process Quality System Review (3) Process Flow Chart (4) Floor Plan Layout (5) Characteristics Matrix (6) Process Failure Mode and Effects Analysis (PFMEA) • Pre-Launch Control Plan (7) Process Instructions (8) Measurement Systems Analysis Plan (9) Preliminary Process Capability Study Plan (10) Packaging Specifications (11) Management Support	(1) Review Supplier PSW status (2) Dealer development initiated (3) Tool Try-out (TTO) completed (4) Process validation completed (5) Production Trial Run (PP) completed (6) Manufacturing Facility ready, including logistics (7) Results of Reliability, VP Tests available (8) Reliability analysis complete (9) Product Cost/Time performance details available. (10) CAE updated based on VP results. (11) Review Team Feasibility Commitment (12) Field Evaluation Units/Pilot lot build commencement for early user feedback (13) PVT involvement begins (14) Production Validation Testing initiated (15) External Jury Evaluation (16) Final component price available (17) Certification for homologation (18) Performance Targets re-affirmed through External Jury (19) Operator training completed (20) Engineering Sign off by:

(Cont'd...)

Table 14. *(Continued)*

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
(28) CMVR/Homologation testing complete (Present vehicle to authorities and obtain certificate) (29) Initiation of FEU/Pilot builds for early user feedback using Prelaunch control plan (30) Process FMEA Complete (31) High priority manufacturing issues sorted out (32) Facilities, tools, and gages completion (33) Logistics in place (34) Finalize process (M/C Availability/ capacity/capability) (35) Operator training completed and resources deployed (36) Review and update in-plant indicators (37) Perform production validation testing (38) Verify process capability (PPk) (39) PVT involvement starts (40) Field trials (41) Decide warranty policy (42) Identify and develop a service network (43) Conduct weekly launch meetings (44) Product price (BNDP) available (45) Capture and collate Lessons learned (46) Go ahead/hold the decision for full-scale production (47) Gateway review meetings			Customer Management, Product Development, Vendor Management (21) Mfg. Issues (High Priority) sorted out (22) Design and Performance Signed Off and frozen (23) Product Price (BNDP) Available (24) Logistics in place (25) Lessons Learnt Recorded (26) Supplier PSW complete

Abbreviations: APQP: Advance product quality planning; BNDP: Net dealer price before taxation; BOM: Bill of material; CAE: Computer-aided engineering; CMVR: Central motor vehicle regulation; FEU: Field evaluation units; FMEA: Failure mode effects analysis; FTR: Fitment trial run report; IPR: Intellectual property rights; J1: Job1; PFMEA: Process failure mode effects analysis; PIPC: Percentage of indices process capable; PIST: Percentage of inspection points that satisfies specified tolerance; PP: Production prove out; PPAP: Production part approval process; PPk: Process capability index; PSW: Part submission warrant; PTR: Production trail run; PVT: Plant vehicle team; SAM: Serviceability, accessibility, maintainability; TAKT: Total activity completion time; TTO: Tooling try outs; VP: Validation prototype.

Table 15. Stage 9 and Gate 9 new product development activities

Stage-Gate NPD Activities	**Stage-Gate Objectives, Lead Roles, Inputs and Deliverables**		
Stage 9 – Start of production	Objectives- 1. To validate the manufacturing process 2. To ensure that customer requirements will be met		
Gate 9 – Launch readiness review	Objectives – Start production of new vehicle, at the manufacturing location, at the TAKT time by the production operators		
NPD activities	Lead role	Inputs	Deliverables
(1) Product Pricing Strategy finalized (2) Advertisement, Exhibitions and demonstrations material ready (3) Distribution network identified (4) Marketing launch plan ready (5) Concern monitoring and resolution on FEUs initiated (6) Design Review (DR-4) for Production (7) Complete the Suppliers' PSW and start material procurement (8) Evaluation of packaging and logistics completed (9) Initiate and finalize ramp-up plan (10) Final variant production plan ready (11) Manufacturing sign off completed (12) Final process (SAP) BOM completed (13) Manufacturing Process Sheets completed (14) Manufacturing PSW complete (15) Complete Production Validation Tests (16) Quality Planning Sign-off (17) Production build control plan prepared (18) Review NOVA-C audits and Update NOVA C targets (19) Field trials continued (20) Develop special service tools (21) Field support plan preparation (22) Dealer development completed (23) Service engineer and Dealer training completed	(1) Product planning/ Program Management (2) Supplier (3) Manufacturing engineering/Facility creation (4) Design and development (5) HR (6) Marketing (7) Launch planning/ service, and training	(1) Production Trial Run (2) Measurement Systems Evaluation (3) Preliminary Process Capability Study (4)Production Part Approval (5) Production Validation Testing (6) Packaging Evaluation (7) Production Control Plan (8) Quality planning sign-off and management support	(1) Service and Dealer Training completed (2) Parts catalog/PAB and operator manual available (3) Quality Planning Sign off completed (4) Ramp-up planning initiated and finalized (5) Spares' Kit and Costing available with dealers (6) Production Validation Testing completed (7) Manufacturing Sign-off completed (8) Packaging and logistics evaluation complete (9) Marketing Launch Plan (Media etc.) ready (10) Final variant production plan ready (11) Action plan on feedback from FEU initiated (12) Manufacturing PSW Complete

(Cont'd...)

Table 15. *(Continued)*

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
(24) Preparation of Parts and Accessories Bulletin (25) Preparation of Parts Catalogue and Owner's Manual (26) Spare kit, spare pricing, preparation, and distribution (27) Compute product costing (28) Conduct regular launch meetings (29) Gateway review meeting			(13) Distribution network identified (14) Final SAP BOM completed (15) Dealer Development completed (16) Design Review-DR4 completed

Abbreviations: APQP: Advance product quality planning; BOM: Bill of material; DR: Design review; FEU: Field evaluation unit; IPR: Intellectual property rights; MOST: Meynord operation sequence technique; NOVA C: New overall vehicle audit - complete; PSW: Part submission warrant; PV: Production validation; PVT: Plant vehicle team; SAP: System applications and products in data processing; SOP: Start of production; TAKT: Total activity completion time.

Table 16. Stage 10 and Gate 10 new product development activities

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
Stage 10 – Launch and implementation	Objectives- Introduction of the product into the market		
Gate 10 – Manufacturing and market review	Objectives- Manufacturing and Market launch		
NPD activities	Lead role	Inputs	Deliverables
(1) Market launch events (Event Management at identified markets) (2) Resolution of "A" concerns reported from FEUs completed (3) Availability of components for production quantity as per the ramp-up plan (4) Component ownership transferred to respective supply module(s) (5) Durability tracking study on FEUs continues (6) First run process capability established (Pp, PPk targets) (7) Line balancing and TAKT on production batch achieved	(1) Product planning/Program Management (2) Supplier (3) Manufacturing engineering/ Facility creation (4) Design and development (5) HR (6) Marketing	(1) Production Trial Run (2) Measurement Systems Evaluation (3) Preliminary Process Capability Study (4) Production Part Approval (5) Production Validation Testing	(1) All "A" concerns resolved (2) NOVA-C target achieved (3) Production on line stabilized (4) Ramp-up planning and deployment (5) Concern Management (6) Early field concern resolution (7) First lot of Saleable products

(Cont'd...)

Table 16. *(Continued)*

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
(8) Man61ufacture and roll down production batch as per Production Control Plan (9) Initiate Launch Sampling Plan and long-duration vehicle road tests (10) Verify achievement of NOVA-C targets (11) Verify ongoing process capability (CPk) (12) Conduct weekly launch meetings (13) Gateway review meetings	(7) Launch planning/service and training	(6) Packaging Evaluation (7) Production Control Plan (8) Quality Planning Sign-off and management support	(8) Distribution of product (9) Launch form filled (10) Build quantity for nationwide launch (11) Event management for Launch

Abbreviations: APQP: Advanced product quality planning; CDS: Component design specification; DFMEA: Design failure mode effect analysis; DVP: Design verification plan; FEU: Field evaluation unit; J1: Job1; MOST: Meynord operation sequence technique; NOVA-C: New overall vehicle audit - complete; PFMEA: Process failure mode effect analysis; PSW: Part submission warrant; SDS: System design specification; TAKT: Total activity completion time.

Table 17. Stage 11 and Gate 11 new product development activities

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables		
Stage 11 – Ramp up	Objectives- Ramp up of production as per the marketing schedule		
Gate 11 – Final Project review	Objectives- Formal closing of the project and the team disbanded.		
NPD activities	Lead role	Inputs	Deliverables
(1) Event Management (Market Launch in different territories) (2) Record and update variance to targets (Q, C, F, T, etc) (3) Update the generic SDS in the reference library (4) Continuous improvements on field concerns (5) Ramp up builds (6) Capture and analyze early customer feedback (USVR) (7) Based on the ramp-up plan and feedback from the field, alternate source development was initiated	(1) Product planning/Program Management (2) Supplier (3) Manufacturing engineering/ Facility creation (4) Design and development	(1) Production Trial Run (2) Measurement Systems Evaluation (3) Preliminary Process Capability Study (4) Production Part Approval (5) Production Validation Testing	(1) Capturing early Customer Feedback (2) Durability tracking study (3) Second Source development (4) Lessons learnt recorded in the format (5) Project variance recorded with real time cost, quality and time

(Cont'd...)

Table 17. *(Continued)*

Stage-Gate NPD Activities	Stage-Gate Objectives, Lead Roles, Inputs and Deliverables	
(8) Durability tracking study completed (9) Monitor NOVA-C (10) PVT owns the responsibility for the model (11) Capture and collate Lessons Learnt, Update knowledge base (12) Project Closing Report (13) Final Project closing review meeting (14) Evaluate team performance	(6) Packaging Evaluation (7) Production Control Plan (8) Quality Planning sign-off and management support	(6) CAE updated based on Production Validation Test results (7) Customer Survey Process initiated [IQS] (8) Project inventory disposed off (9) CFT disbanded, PVT takes over

Abbreviations: APQP: Advance product quality planning; CAE: Computer-aided engineering; CFT: Cross-functional team; Q, C, F, T: Quality, cost, function, time; IQS: Initial quality study; J1: Job 1; MM: Materials management; NOVA C: New overall vehicle audit – complete; PSW: Part submission warrant; PVT: Plant vehicle team; SDS: System design specification; USVR: User security verification routine.

Program confirmation review, after gate five (Design Verification Review) approval, where the final design is released and long lead tooling is signed off.

(7) Stage seven, named design validation, includes the lead role of the design and development department, supported by other departments such as product planning, HR, supplier, marketing, launch planning, and manufacturing engineering. The objective of this stage is a trial for tooled-up parts, crash tests, aggregate tests, and detailed costing. The objective of stage seven requires completing 20 activities. The deliverables of stage seven are reviewed in gate seven, named as design release review, after gate six (Program Confirmation Review) approval, where the final design is released, long lead tooling is signed off, the component reliability plan is ready, and the business case is validated.

(8) Stage eight, named transition to production, includes the lead role of the design, development, and manufacturing department, supported by other departments such as product planning, HR, supplier, launch planning, finance, budgeting, and manufacturing engineering. The objective of this stage is to validate the manufacturing process and to ensure that customer requirements will be met. The objective of stage eight requires completing 47 activities. The deliverables of stage eight are reviewed in gate eight, named Sign-off review, after gate seven (Design Release Review) approval, where formal acknowledgment is made that all aspects of product design are completed as per intent and verified.

(9) Stage nine, named the start of production, includes the lead role of the manufacturing department supported by other departments such as product planning, design and development, HR, supplier, launch planning, marketing, and manufacturing engineering. The objective of this stage is to validate the manufacturing process and to ensure that customer requirements will be met. The objective of stage nine requires completing 29 activities. The deliverables of stage nine are reviewed in gate nine, named launch readiness review, after gate eight (Sign Off Review) approval, where production of the new vehicle starts at the manufacturing location at the total activity completion time by the production operators.

(10) Stage 10, named launch and implementation, includes the lead role of the launch planning and marketing department, supported by other departments such as product planning, design and development, HR, supplier, and manufacturing engineering. The objective of this stage is the

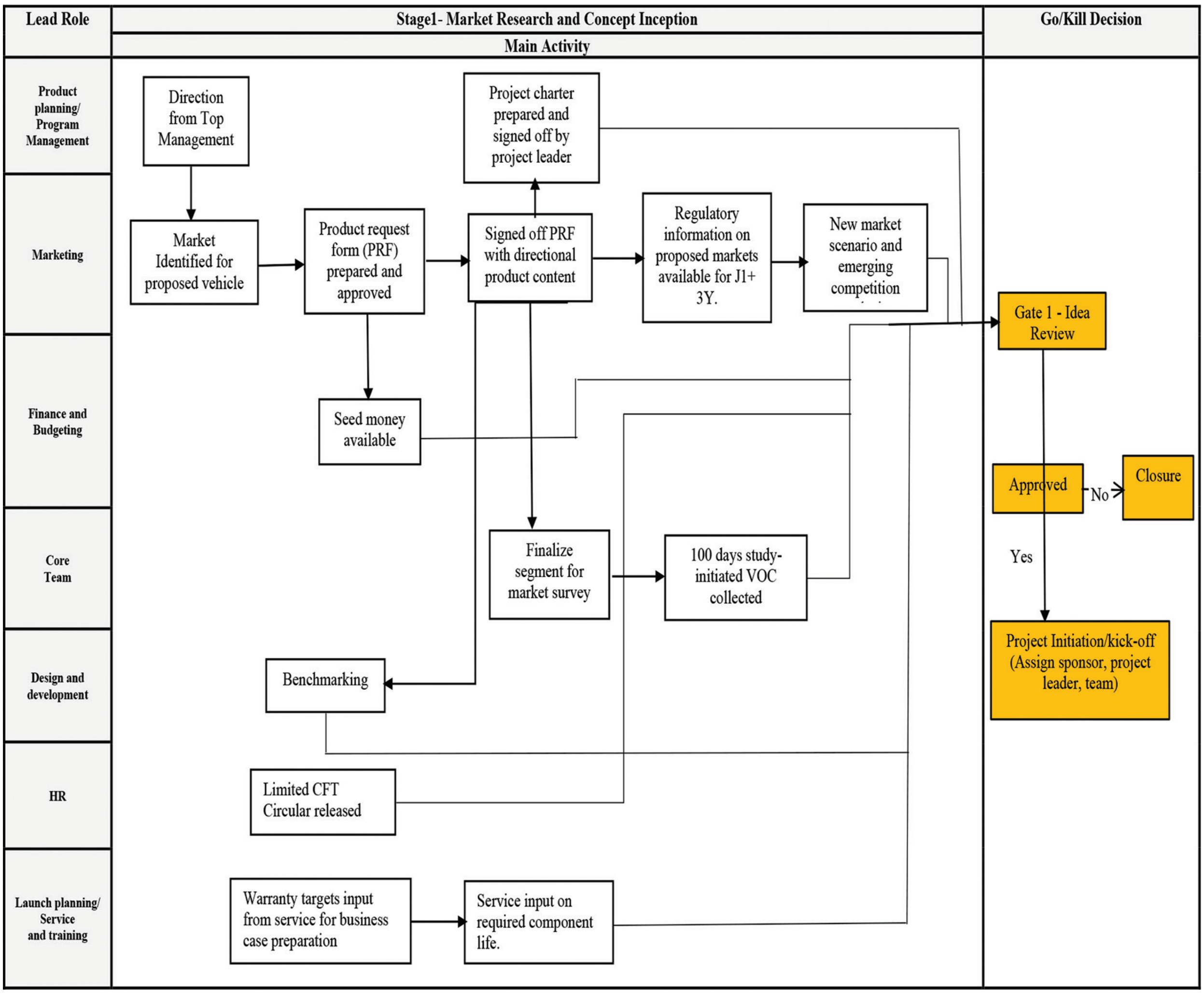


Fig. 4. Proposed new product development process framework for Stage 1 and Gate 1 new product development activities

introduction of the product into the market. The objective of stage ten requires completing 13 activities. The deliverables of stage 10 are reviewed in gate 10, named manufacturing and market review, after gate nine (Launch Readiness Review) approval, where the product is launched.

(11) Stage 11, named ramp-up, includes the lead role of manufacturing, launch planning, and marketing department, supported by other departments such as product planning, design and development, supplier, and manufacturing engineering. The objective of this stage is to ramp up production as per the marketing schedule. The objective of Stage 11 requires completing 14 activities. The deliverables of stage eleven are reviewed in Gate 11, named final project review, after gate 10 (Manufacturing and Market Review) approval, where the formal closing of the project and team is disbanded.

The proposed NPD framework has the following hallmarks as compared to the existing frameworks:

(1) The common stages across all available frameworks include marketing, design, development, manufacturing, and product launch. To simplify and enhance the effectiveness of this framework, each main stage of the NPD process is further divided into sub-stages.
(2) The number and titles of stages and gates are modified to align with specific objectives and deliverables.
(3) Each activity is defined in a structured format, specifying the lead role and its integration within the framework.
(4) This framework is applicable to projects that are new to design and development.
(5) Each stage emphasizes the involvement of both external and internal customers, enhancing cross-functional interaction.
(6) A more disciplined approach to cross-functional meetings and communication in NPD activities involves clear agendas that focus on specific deliverables and criteria, keeping discussions on

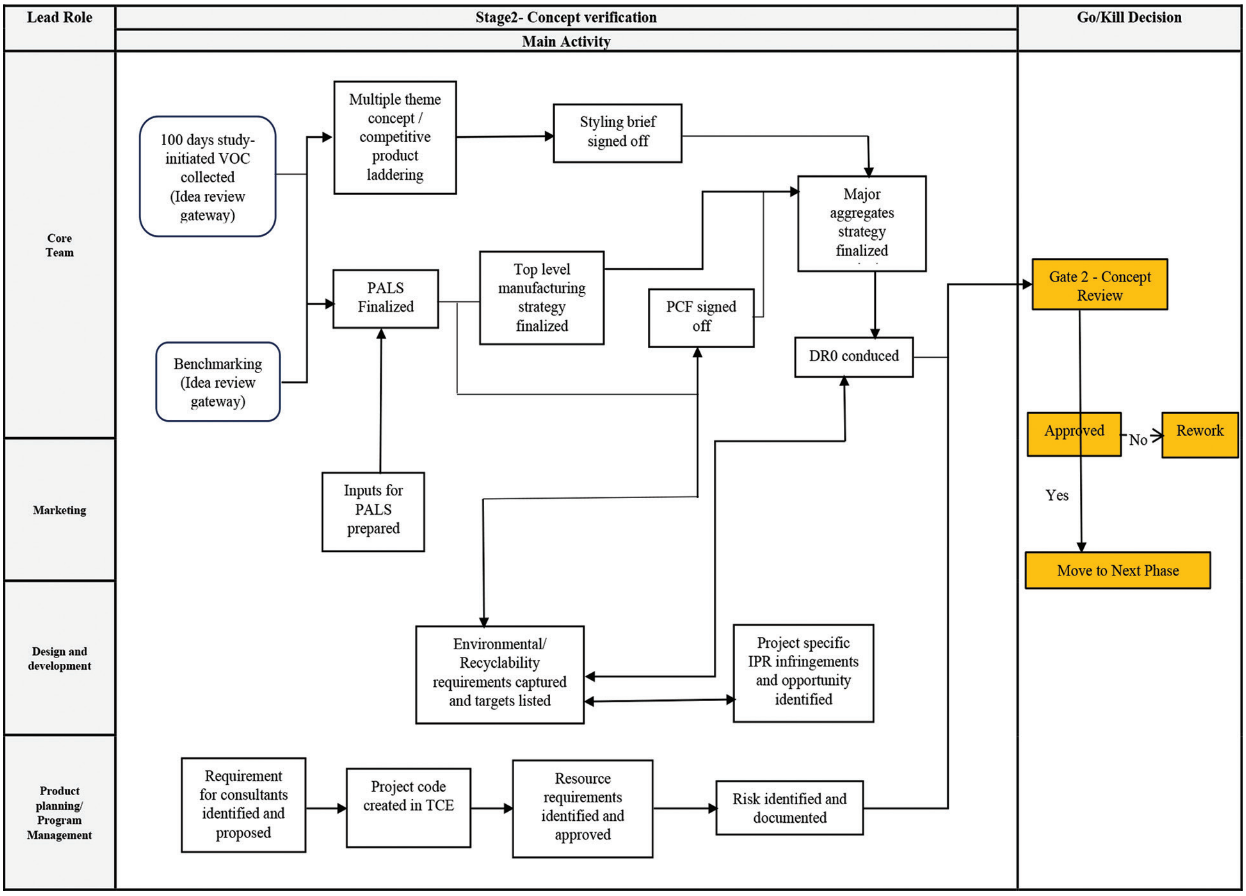


Fig. 5. Proposed new product development process framework for Stage 2 and Gate 2 new product development activities

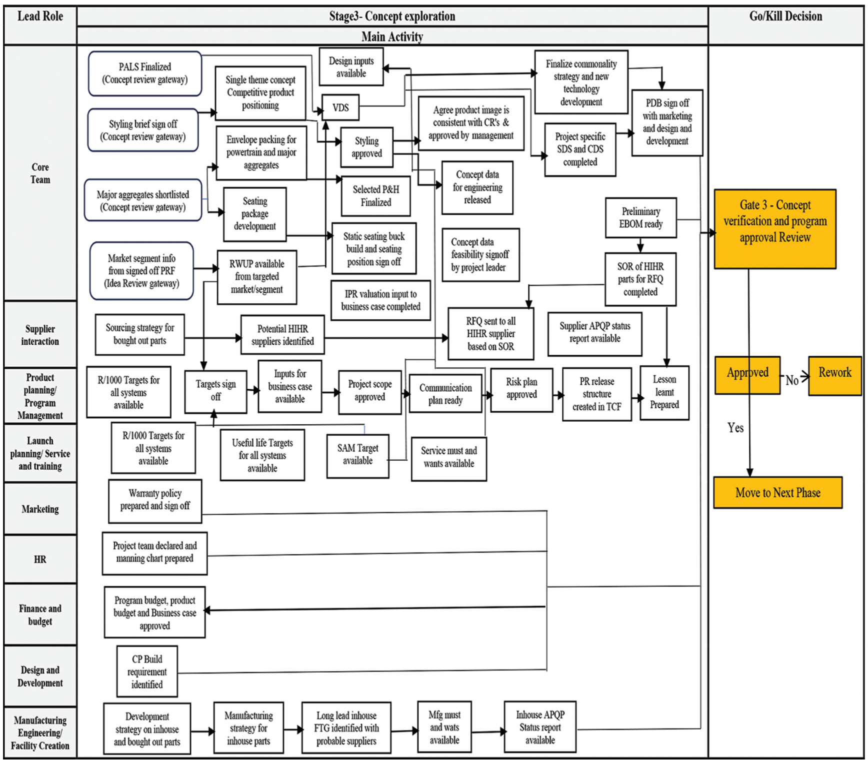


Fig. 6. Proposed new product development process framework for Stage 3 and Gate 3 new product development activities

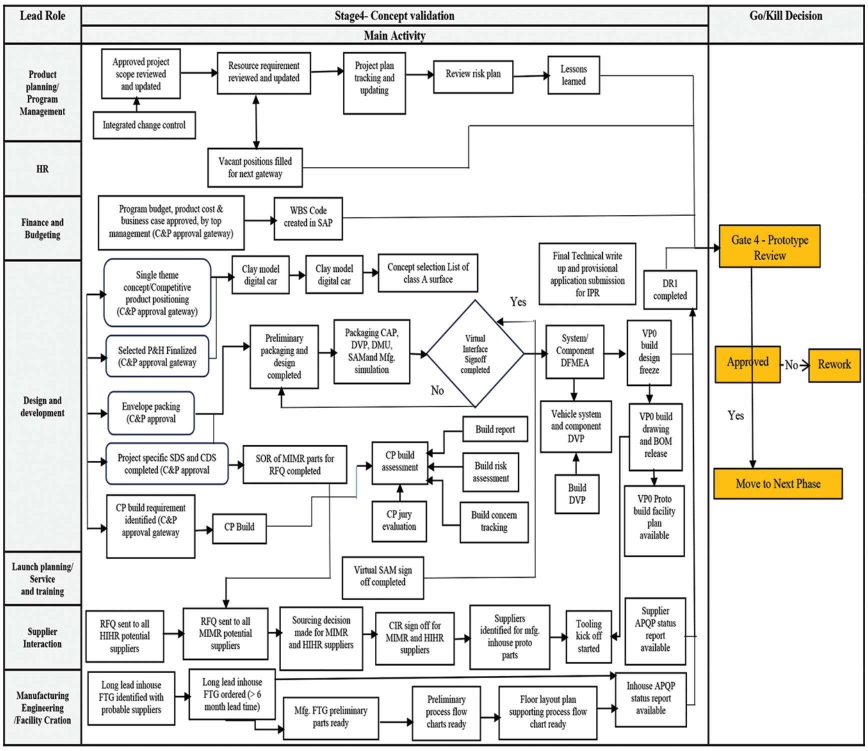


Fig. 7. Proposed new product development process framework for Stage 4 and Gate 4 new product development activities

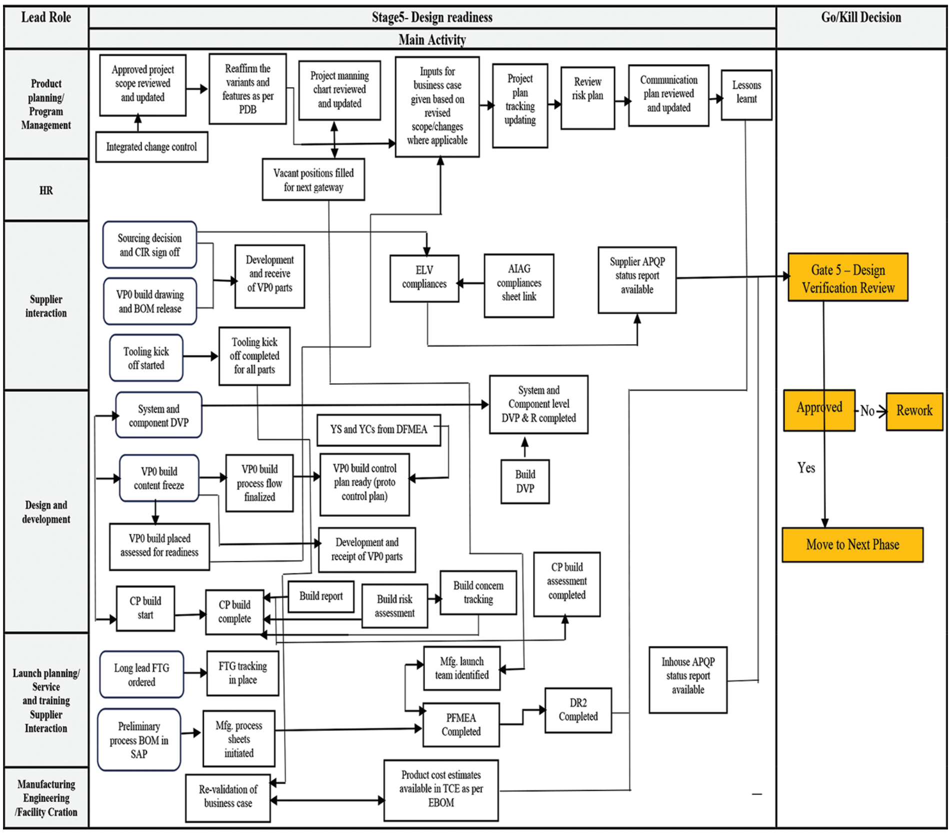


Fig. 8. Proposed new product development process framework for Stage 5 and Gate 5 new product development activities

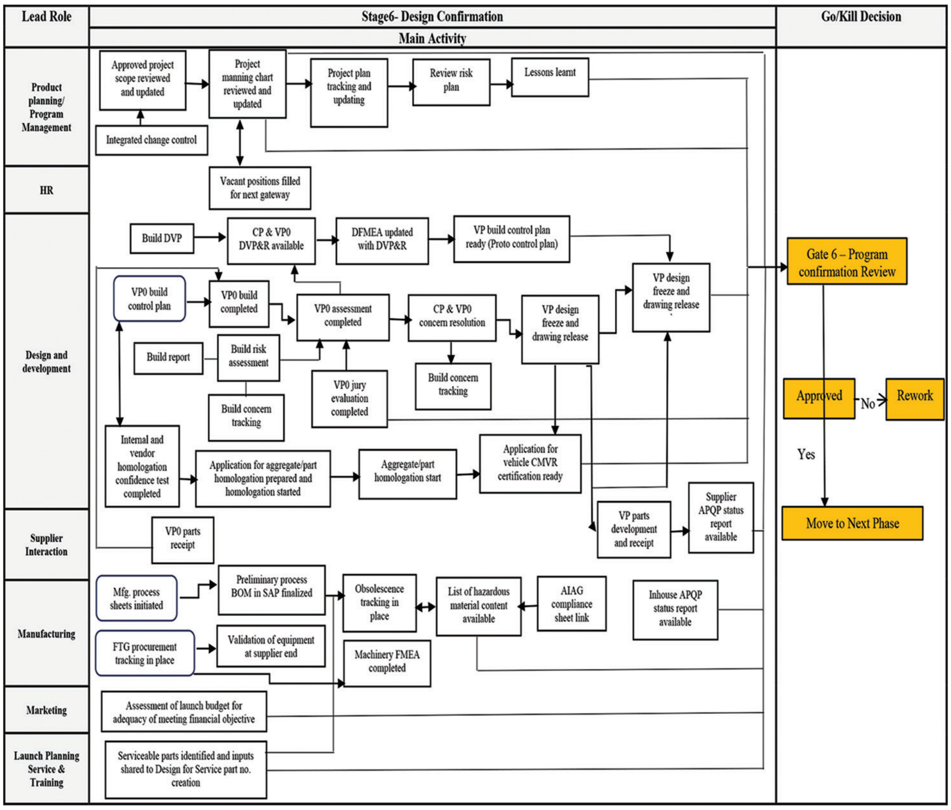


Fig. 9. Proposed new product development process framework for Stage 6 and Gate 6 new product development activities

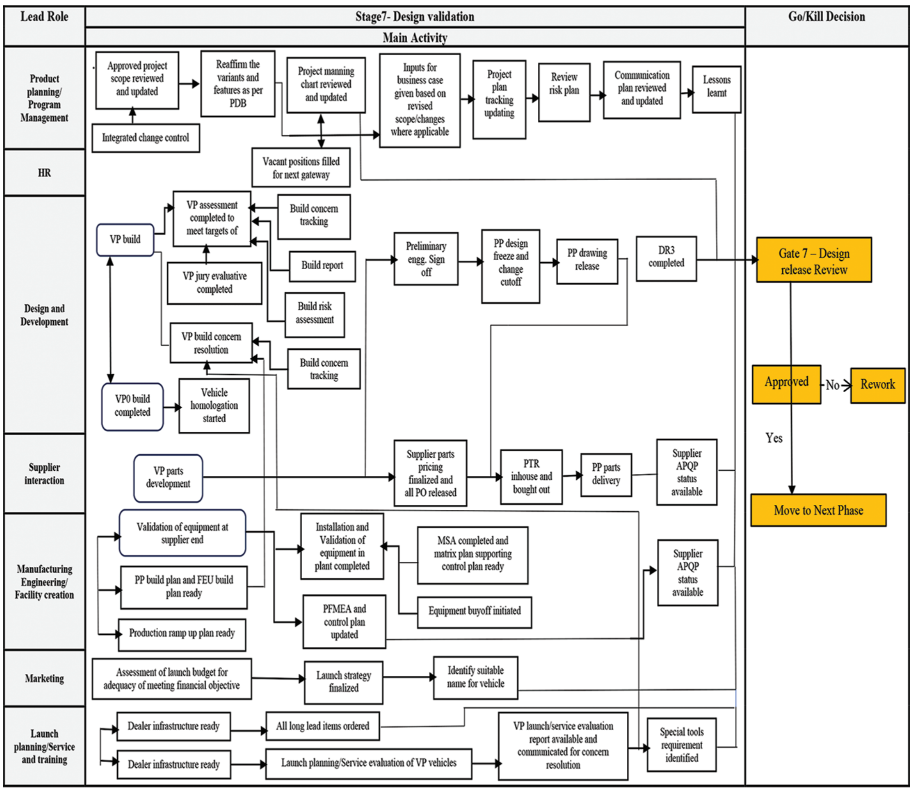


Fig. 10. Proposed new product development process framework for Stage 7 and Gate 7 new product development activities

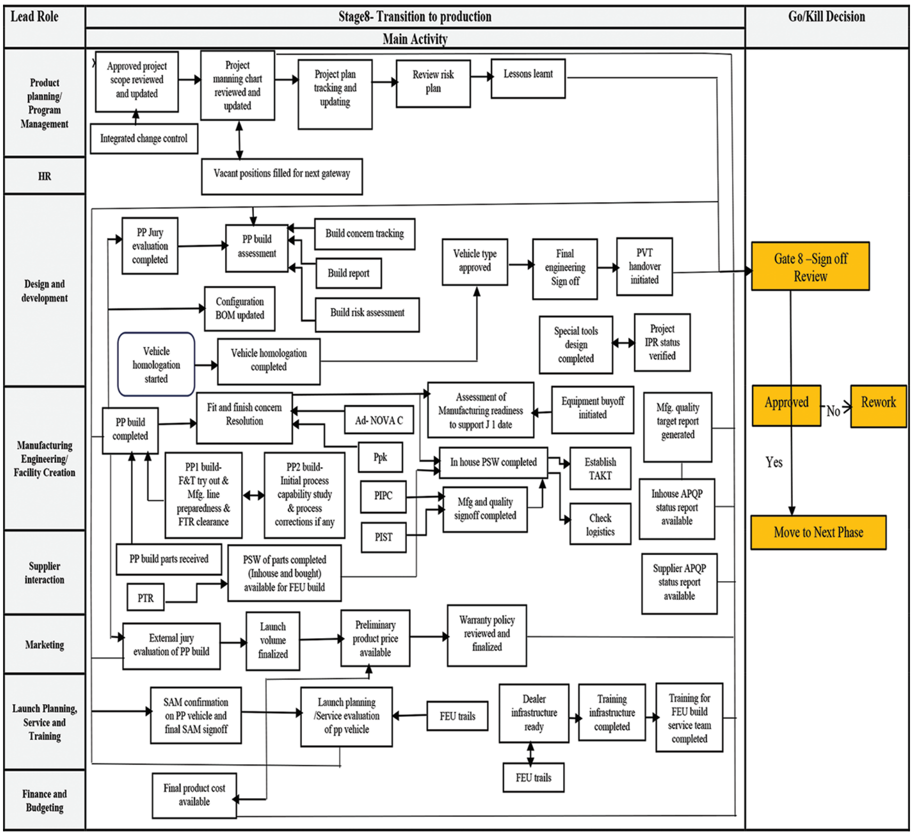


Fig. 11. Proposed new product development process framework for Stage 8 and Gate 8 new product development activities

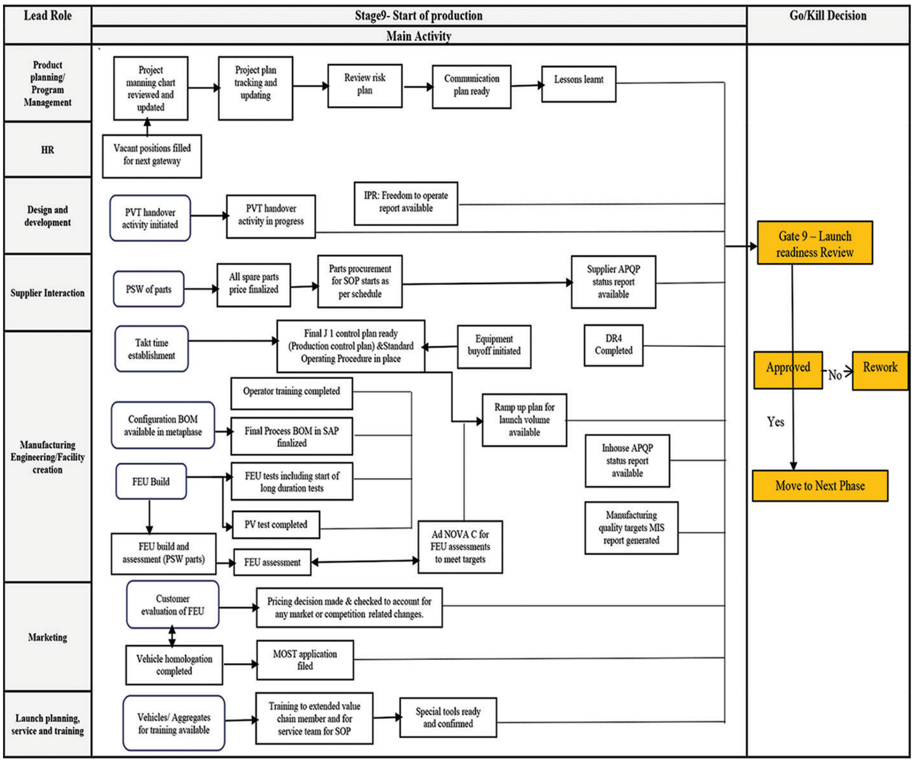


Fig. 12. Proposed new product development process framework for Stage 9 and Gate 9 new product development activities

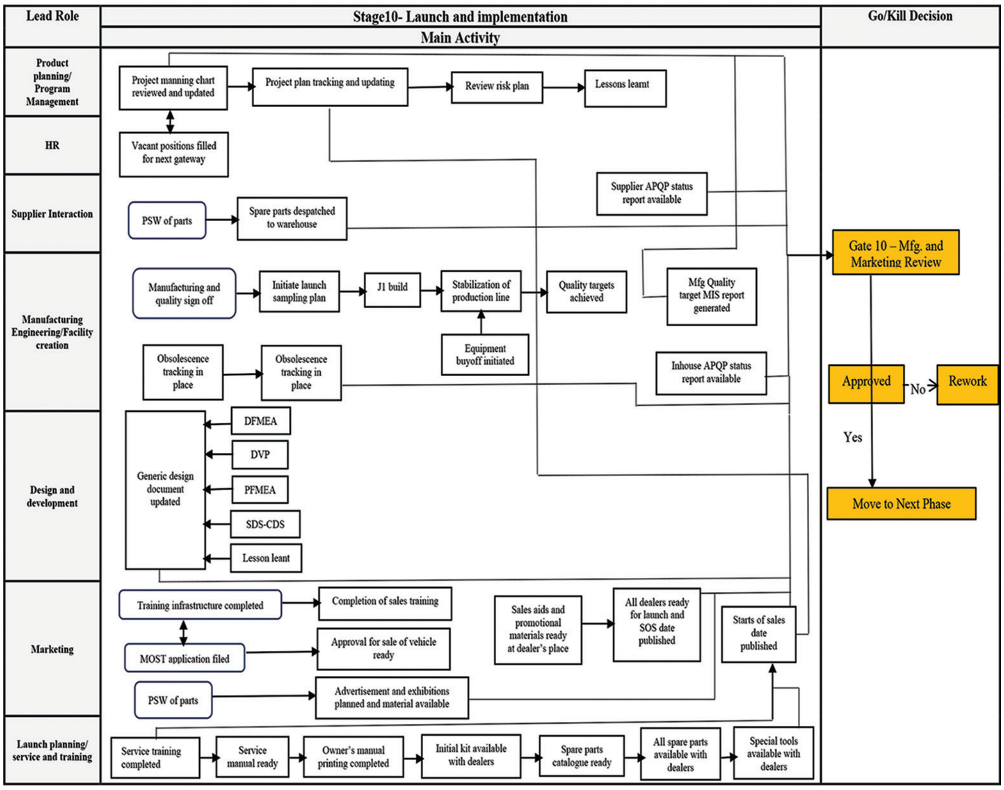


Fig. 13. Proposed new product development process framework for Stage 10 and Gate 10 new product development activities

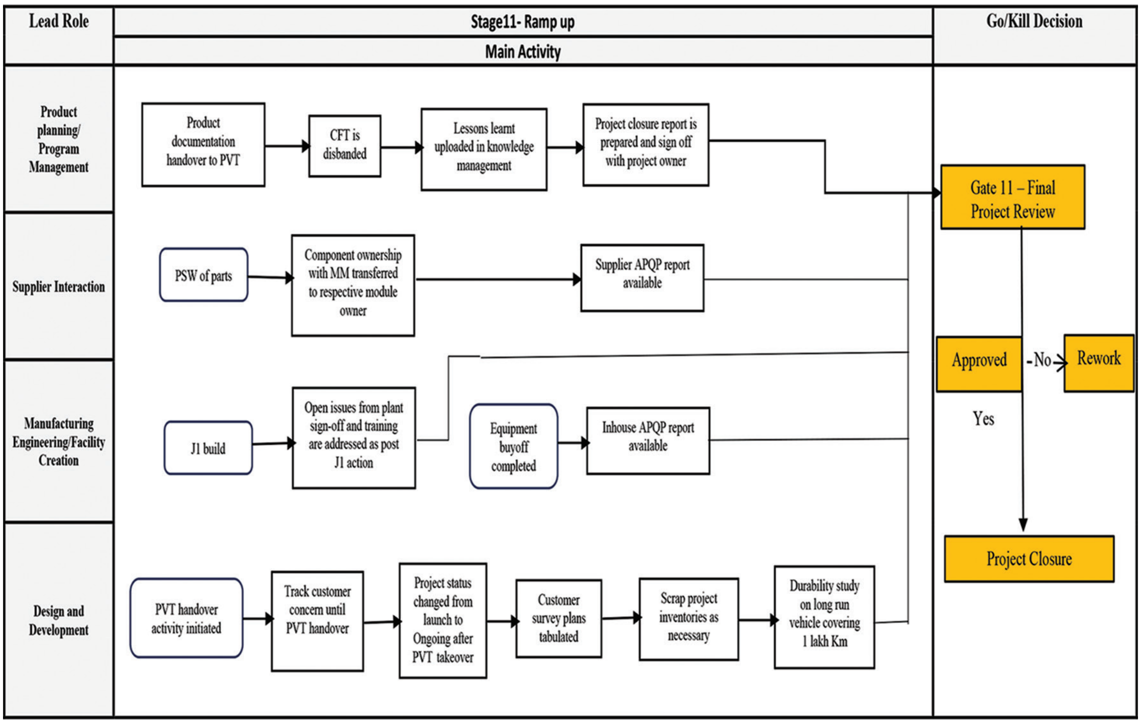


Fig. 14. Proposed new product development process framework for Stage 11 and Gate 11 new product development activities

track and efficient.

(7) Cross-functional teams frequently incorporate diverse perspectives and expertise, promoting proper teamwork and active participation from

all stakeholders throughout the project.

(8) A proper feedback mechanism is maintained by clear communication, making the framework efficient and a straightforward change control procedure. This structured framework ensures that changes are evaluated, approved, and implemented systematically, minimizing disruptions and maintaining NPD activities' alignment with objectives.

(9) The framework includes robust cost control mechanisms, such as periodic review activities, which enhance cost planning, tracking, and control.

(10) This framework allows for incremental progress and avoids overwhelming the team with unrealistic expectations. Criteria-based evaluation, cross-functional inputs, continuous feedback, adequate resources, and a realistic timeframe make the milestones achievable and realistic.

(11) Customized to meet automobile needs.

(12) Cross-functional team collaboration enhances interdepartmental communication and knowledge sharing.

(13) It is highly effective in data management for input and delivery processes due to its structured and systematic approach. It ensures comprehensive data collection through meticulous requirements gathering, integrates data from various functional areas, and maintains high data quality and consistency through standardization and validation.

(14) Checks such as the use of standardized templates and guidelines, training for data collection and documentation, and regular audits are incorporated for data compatibility and consistency.

(15) Structured review and performance metrics of the framework. Improve system feedback on deliverables (met/not met).

(16) Standardization across frameworks is achieved through the use of consistent processes, ensuring uniformity and efficiency. This standardization helps in aligning main activities and deliverables with the objectives, facilitating clear planning and execution.

(17) Non-value-added activities are minimized, which shifts the work culture from reactive to proactive mode.

(18) This framework is notably flexible in handling various projects due to its iterative and adaptable nature. The framework's cross-functional collaboration encourages diverse perspectives and expertise, enhancing its ability to address unique challenges across different projects. In addition, the process often includes stages for regular review and adjustment, which supports its adaptability to various project demands and uncertainties.

7. Conclusion

The literature study has identified several common challenges in NPD, including market uncertainty, resource allocation, time management, technical challenges, regulatory compliance, cross-functional collaboration, risk management, innovation management, customer feedback, supply chain coordination, cost control, quality assurance, and post-launch issues. The review of the NPD process in the automobile industry emphasized its critical role in overcoming these challenges and concluded that there is a need to redesign the NPD process framework. In alignment with existing literature, particularly regarding the Indian automobile sector, it has been confirmed that the Stage-Gate framework is the most suitable for industries that prioritize meticulous planning and control. The identified limitations of benchmarked NPD process frameworks, followed by the Stage-Gate approach implemented by automobile and supporting companies, and specifications and requirements by a survey among the NPD professionals, were considered in the design and development of the proposed NPD process framework, followed by some hallmarks as compared to the existing frameworks.

8. Future Scope

It is advisable to verify the proposed NPD process framework by conducting case studies to evaluate its NPD performance. With the imminent industrial revolution and the rise of new technologies and digital transformations, the mapping and integration of the proposed NPD process framework with emerging technologies such as AI/DS, PDM, PLM, IoT, Machine Learning, and TRIZ has the potential to boost innovation, shorten development cycles, and enhance product quality.

References

Adams-Bigelow, M. (2004). First results from the 2003 comparative performance assessment study (CPAS). In: *The PDMA Handbook of New Product Development*. Wiley, United States, p546–566. https://doi.org/10.1002/9780470172483

Agrawal, A., & Bhuiyan, N. (2014). Achieving success in NPD projects. *International Journal of Industrial and Manufacturing Engineering*. 8(2), 476–481.

Almeida, O., Figueiredo, P.S., Estevao-Beal, V., & Uchoa-Passo, F. (2020). Critical success factors of product development projects in the automotive industry. *Journal of Technology Management and Innovation*, 15(2), 56–70. https://doi.org/10.4067/S0718-27242020000200056

Anand, D.U., Janakiraman, S.R., & Sethukannan, V.S. (2009). *Lean Product Development-Redefining the Indian Automotive Product Development Process Using Lean Framework*. SAE Technical Paper 2009-01-0117. https://doi.org/10.4271/2009-01-0117

Anand, G., & Kodali, R. (2008). Development of a conceptual framework for lean new product development process. *International Journal of Product Development*, 6(2), 190-224. https://doi.org/10.1504/IJPD.2008.01924

Arnold, G.W., & Floyd, M.C. (1997). Reengineering the new product introduction process. *International Journal of Production Economics*, 52(1-2), 179–183. https://doi.org/10.1016/S0925-5273(97)00052-2

Barczak, G., Griffin, A., & Kahn, K.B. (2009). Trends and drivers of success in NPD practices. *Journal of Product Innovation Management,* 26(1), 3–23. https://doi.org/10.1111/j.1540-5885.2009.00331.x

Barringer, B.B., & Gresock, A.R. (2008). Formalizing the front-end of the entrepreneurial process using the stage-gate model as a guide: An opportunity to improve entrepreneurship education and practice. *Journal of Small Business and Enterprise Development*, 15(2), 289–303. https://doi.org/10.1108/14626000810871682

Boejang, H., Hambali, A., Hassan, M.Z, Esa, S., & Rauterberg, G.W.M., (2017). An exploration on new product development process of Malaysian small-sized automaker. *Journal of Advanced Manufacturing Technology*, 11(2), 33–46.

Brem, A., & Kurzdorfer, T. (2016). New product development, R&D and culture: Results from a multiple case study of German and Chinese innovation processes. *International Journal of Product Development*, 21(2–3), 144–174. https://doi.org/10.1504/IJPD.2016.078862

Broum, T., Kopecky, M., & Kleinova, J. (2011). Enhancement of stage-gate process by value analysis. *Annals of DAAAM and Proceedings*, 22(1), 755–756.

Brown, S.L., & Eisenhardt, K.M. (1995). Product development: Past research, present findings, and future directions. *Academy of Management Review*. 20(2), 343–378. https://doi.org/10.5465/amr.1995.9507312922

Cano, E.L., Garcia-Camus, J.M., Garzas, J., Moguerza, J.M., & Sánchez, N.N. (2021). A Scrum based framework for new product development in the non-software industry. *Journal of Engineering and Technology Management*, 61, 101634. https://doi.org/10.1016/j.jengtecman.2021.101634

Chao, L.P., & Ishii, K. (2005). Design process error-proofing: benchmarking gate and phased review life-cycle models. In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 47411. p301–310. https://doi.org/10.1115/DETC2005-84235

Chirumalla, K. (2017). Challenges in Managing New Product Introduction Projects: An Explorative Case Study. In: *DS 87-2 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 2: Design Processes, Design Organisation and Management, Vancouver, Canada.* p259-268. Available from: https://www.designsociety.org/publication/39580/challenges

Clark, K.B., & Fujimoto, T. (1991). *Product Development Performance: Strategy, Organization, And Management in the World Auto Industry*. HBS Press, Boston. Available from: https://tinyurl.com/mrx945ep

Connell, J., Edgar, G.C., Olex, B., Scholl, R., Shulman, T., & Tietjen R. (2001). Troubling successes and good failures: Successful new product development requires five critical factors. *Engineering Management Journal*, 13(4), 35–39. https://doi.org/10.1080/10429247.2001.11415136

Cooper, L.P. (2003). A research agenda to reduce risk in new Product Development through knowledge management, a practitioner perspective. *Journal of Engineering and Technology Management*, 20(1–2), 117–140. https://doi.org/10.1016/S0923-4748(03)00007-9

Cooper, R.G. (1990). Stage-gate systems: A new tool for managing new products. *Business Horizons*, 33(3), 44–54. https://doi:10.1016/0007-6813(90)90040-I

Cooper, R.G. (2008). What leading companies are doing to re-invent their NPD processes. *PDMA Visions Magazine*, 32(3), 6–10.

Cooper, R.G. (2022). The 5th generation stage-gate idea-to-launch process. *IEEE Engineering Management Review*, 50(4), 43–55. https://doi.org/10.1109/EMR.2022.3222937

Cooper, R.G., & Scott E.J. (2003). *Benchmarking Best Practices Performance Results and the Role of Senior Management*. Vol. 1. Product Development Institute Inc., United States. Available from: https://acrobat.adobe.com/id/urn:aaid:sc:AP:4367a4ca-4931-467c-adaa-cdf988e14c0c

Cooper, R.G., Edgett, S.J., & Kleinschmidt E.J. (2002). *Improving New Product Development Performance and Practices*. American Productivity and Quality Center, United States.

Available from: https://surli.cc/fajugt

Cooper, R.G., Edgett, S.J., & Kleinschmidt, E.J. (2003). *Best Practices in Product Innovation: What Distinguishes Top Performers*. Product Development Institute, United States.

Cooper, R.G., Edgett, S.J., & Kleinschmidt, E.J. (2004a). Benchmarking best NPD practices-I. *Research-Technology Management*, 47(1), 31–43. https://doi.org/10.1080/08956308.2004.11671606

Cooper, R.G., Edgett, S.J., & Kleinschmidt, E.J. (2004b). Benchmarking best NPD practices -II. *Research-Technology Management*, 47(3), 50–59. https://doi.org/10.1080/08956308.2004.11671630

Cooper, R.G., Edgett, S.J., & Kleinschmidt E.J. (2004c). Benchmarking best NPD practices-III. *Research-Technology Management,* 47(6), 43-55. https://doi.org/10.1080/08956308.2004.11671662

Davoodi, A., & Aslanzadeh Y. (2014). *Stage-Gate Modified for Lean PD*. Available from: https://odr.chalmers.se/server/api/core/bitstreams/f60a3e0c-4d4a-4843-a8c0-7d2310c148e5/content

Eljayar, A., & Busch, J.S. (2021). Agile-Stage-Gate approach: Exploratory research on the structure, roles, and responsibilities. *Athens Journal of Technology and Engineering*, 8(1), 39–90.

Ettlie, J.E., & Elsenbach, J.M. (2007). Modified Stage-Gate regimes in new product development. *Journal of Product Innovation Management*, 24(1), 20–33. https://doi.org/10.1111/j.1540-5885.2006.00230.x

Fernandes, R.Q.K., Cavalcanti, V., & de Andrade, A.M. (2017). Design management in the toy industry: Case studies on design insertion for the development process in Brazilian toy companies. *Strategic Design Research Journal*, 10(3), 230–240.

Gerhard, D., Brem, A., & Voigt, K.I. (2008). Product development in the automotive industry: Crucial success drivers for technological innovations. *International Journal of Technology Marketing*, 3(3), 203–222. https://doi.org/10.1504/IJTMKT.2008.019922

Griffin, A. (1997). Modelling and measuring product development cycle time across industries. *Journal of Engineering and Technology Management*. 14, 1–24. https://doi.org/10.1016/S0923-4748(97)00004-0

Gronlund, J., Sjodin, D.R., & Frishammar, J. (2010). Open innovation and the stage-gate process: A revised model for new product development. *California Management Review*. 52(3), 106–131. https://doi.org/10.1525/cmr.2010.52.3.106

Gurbuz, E. (2018). Theory of new product development and its applications. In: *Marketing*. IntechOpen, London, p57–75. https://doi.org/10.5772/intechopen.74527

Hamidizadeh, M., Hoseini, M., Akhavan, M., & Hojaeefard, H. (2018). A New consideration on new product development models. *International Journal of Automotive Engineering*, 8(1), 2666–2671.

Harmancioglu, N., McNally, R.C., Calantone, R.J., & Durmusoglu, S.S. (2007). Your new product development (NPD) is only as good as your process: An exploratory analysis of new NPD process design and implementation. *R&D Management*, 37(5), 399–424. https://doi.org/10.1111/j.1467-9310.2007.00486.x

Jaksic, M.L., & Petkovic, M.J.J. (2014). *Sustainable Technology Management and Development-State University and Industry Performance Model*. Sustainable Technology Management [Project Work].

Kadwe, K., Khan, A.A. & Babela, V. (2017). Survey analysis of critical success factors for NPD in Indian manufacturing industries. *International Journal of Current Research in Multidisciplinary (IJCRM)*, 4, 2661–2665.

Kahn, K.B., Barczak, G., Nicholas, J., Ledwith, A., & Perks, H. (2012). An examination of new product development best practice. *Journal of Product Innovation Management*, 29(2), 180–192. https://doi.org/10.1111/j.1540-5885.2011.00888.x

Kazimierska, M., & Grębosz-Krawczyk, M. (2017). New product development (NPD) process–an example of industrial sector. *Management Systems in Production Engineering*, 25(4), 246–250.

Kelley, T. (2001). *The Art of Innovation: Lessons in Creativity from IDEO. America's Leading Design firm. 10.* Crown Currency. Available from: https://tinyurl.com/yc2fz3ta

Khannan, M.S.A., Tontowi, A.E., Herliansyah, M.K., & Sudiarso, A. (2021). New product development method trends and future research: A systematic literature review. *Jurnal Teknik Industri: Jurnal Keilmuan dan Aplikasi Teknik Industri*. 23(1), 11–24. https://doi.org/10.9744/jti.23.1.11-24

Koranne, S., & Shende, V. (2017). Leveraging Project Performance analysis to achieve sustainable Business Goal. In: *Project Management National Conference, India*. Available from: https://www.pmi.org.in/conference2017/pdfs/papers-pdfs/theme-accelerate-economic-growth/6-leveraging-project-performance-analysis.pdf

Kumar, S., & Wellbrock, J. (2009). Improved new product development through enhanced design architecture for engineer-to-order companies. *International Journal of Production Research*, 47(15), 4235–4254.

https://doi.org/10.1080/00207540801939030

Lapunka, I., Barton-Pudlik, J., & Jagoda-Sobalak, D. (2023). *Overview and Prioritization of Critical Success Factors in NPD Models for the Chemical Industry.* Scientific Papers of Silesian University of Technology, Organization and Management / Zeszyty Naukowe Politechniki Slaskiej, Seria Organizacji i Zarzadzanie, p186. Available from: https://managementpapers.polsl.pl/wp-content/uploads/2024/02/186-%c5%81apunka-barton-pudlik-jagoda-sobalak.pdf

Liang, W.H. (2010). *Coordination Mechanisms for New Product Introduction*. Available from: https://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/mr66157.pdf?is_thesis=1&oclc_number=1007700786

Loganathan, Y.D., & Jayakrishnan, M. (2014). *Converging Product Development Processes: A Case of Indian Automobile Industry*. SAE Technical Paper (No. 2014-01-0285). https://doi.org/10.4271/2014-01-0285

Masyhuri, M. (2022). Applying a multiple convergent process in achieving a successful of new product development. *ADPEBI International Journal of Business and Social Science*, 2(1), 1–12. https://doi.org/10.54099/aijbs.v2i1.48

Mayer, H.M. (2020). *Innovation Due to Covid: Yes, But How*. Available from: https://www.forbes.com/sites/hannahmayer

McDermott, C.M., & O'Connor, G.C. (2002). Managing radical innovation: An overview of emergent strategy issues. *Journal of Product Innovation Management*, 19(6), 424–438. https://doi.org/10.1111/1540-5885.1960424

Munoli, N. (2017). *3D Hybrid Model for New Product Development*. United States: Rochester Institute of Technology. Available from: https://repository.rit.edu/cgi/viewcontent.cgi?article=10866&context=theses

Nieuwenhuis, P., & Wells, P. (2015). *The Global Automotive Industry*. Wiley, United States. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118802366

O'Connor, P. (1994). Implementing a stage-gate process: A multi-company perspective. *Journal of Product Innovation Management*, 11(3), 183–200. https://doi.org/10.1016/0737-6782(94)90002-7

Owens, J., & Cooper, R. (2001). The importance of a structured new product development (NPD) process: A methodology. In: *IEE International Symposium on Engineering Education: Innovations in Teaching, Learning and Assessment*. https://doi.org/10.1049/ic:20010040

Owens, J.D., & Davies, J. (2000). The importance of a new product development (NPD) process: Getting started. In: *1st European Conference on KM, Bled School of Management, Bled, Slovenia*. Available from: https://acrobat.adobe.com/id/urn:aaid:sc:AP:bf4b0f6f-7a3b-42b2-b2dc-a6febd0340ee

Phillips, R., Neailey, K., & Broughton, T.A. (1999). Comparative study of six stage-gate approaches to product development. *Integrated Manufacturing Systems*, 10(5), 289–297. https://doi.org/10.1108/09576069910371106

Pietzsch, J.B., Shluzas, L.A., Pate-Cornell, M.E., Yock, P.G., & Linehan, J.H. (2009). Stage-gate Process for the Development of Medical Devices. *The Journal of Medical Devices*, 32, 021004. https://doi.org/10.1115/1.3148836

Poolton, J., & Barclay, I. (1998). New product development from past research to future applications. *Industrial Marketing Management*, 27(3), 197–212. https://doi.org/10.1016/S0019-8501(97)00047-3

Rahim, A.R.A., & Baksh, M.S.N. (2003). Case study method for new product development in engineer-to-order organizations. *Work Study*, 52(1), 25–36. https://doi.org/10.1108/00438020310458705

Rajeshwari, K. (2017). New product development in an entrepreneurial company. *International Journal of Product Development*, 22(2), 81–103. https://doi.org/10.1504/IJPD.2017.086447

Ranjan, A. (2014). *Process Reengineering for New Product Introduction at an Analytical Instrument Manufacturing Firm*, (Doctoral Dissertation, Massachusetts Institute of Technology). Available from: http://hdl.handle.net/1721.1/93850

Ries, E. (2011). *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. Crown Currency*. Available from: https://www.tolobranca.nl/wp-content/uploads/2020/04/global-overview-lean-startup-book-eric-ries.pdf

Ringen, G., & Welo T. (2013). Towards a More Event-Driven NPD Process: First Experiences with Attempts of Implementation in the Front-end Phase. In: *DS 75-1: Proceedings of the 19th International Conference on Engineering Design (ICED13) Design for Harmonies*. Design Processes, Seoul, Korea. Vol. 1, p99–108. Available from: https://surl.lu/arkynt

Salavati, M., Tuyserkani, M., Mousavi, S.A., Falahi, N., & Abdi, F. (2016). Improving new product development performance by risk management. *Journal of Business and Industrial Marketing*, 31(3), 418–425. https://doi.org/10.1108/JBIM-04-2013-0090

Shepherd, C., & Ahmed, P.K. (2000). NPD frameworks:

A holistic examination. *European Journal of Innovation Management*, 3(3), 160–173. https://doi.org/10.1108/14601060010341166

Silva, G.C.D., & Kaminski, P.C. (2017). Proposal of framework to managing the automotive product development process. *Cogent Engineering*, 4, 1–25. https://doi.org/10.1080/23311916.2017.1317318

Singh, A.K., & Garg, A. (2015). Impact of information integration on decision-making in a supply chain network. *Production Planning and Control*, 26(12), 994–1010. https://doi.org/10.1080/09537287.2014.1002024

Smolnik, T., & Bergmann, T. (2020). Structuring and managing the new product development process review on the evolution of the Stage-Gate® process. *Journal of Business Chemistry,* 2, 41–57. https://doi.org/10.17879/22139478907

Stosic, B., & Milutinovic, R. (2014). Possibilities of opening up the stage-gate model. *Romanian Statistical Review*, 62(4), 41–53.

Sumantran, V. (2004). Accelerating product development in the automobile industry. *International Journal of Manufacturing Technology and Management*, 6(3-4), 361–371. https://doi.org/10.1504/IJMTM.2004.005424

Tennant, C., & Roberts, P. (2003). The creation and application of a self-assessment process for new product introduction. *International Journal of Project Management*, 21(2), 77–87. https://doi.org/10.1016/S0263-7863(02)00019-4

Thomke, S., & Feinberg, B (2022). *Experimentation Works: The Surprising Power of Business Experiments*. Harvard Business Press. Available from: https://sloanreview.mit.edu/article/capturing-the-real-value-of-innovation-tools

Ulrich, K., & Eppinger, S. (2020). *Product Design and Development*. 7th ed. McGraw-Hill, United States. Available from: https://tinyurl.com/2bavs3c8

Ulrich, K.T., Eppinger, S.D., & Goyal, A. (2009). *Product Design and Development*. McGraw-Hill Education, United States.

Wang, L., & Kourouklis, A. (2012). On the Development of a Theoretical Framework for New Product Development. In: *International Conference on Knowledge Management and Information Sharing, SCITEPRESS*. Vol. 2, p249–259. Available from: https://www.scitepress.org/papers/2012/41364/41364.pdf

Wijewardhana, G.E.H., Weerabahu, S.K., Nanayakkara, J.L.D., & Samaranayake, P. (2021). New product development process in apparel industry using Industry 4.0 technologies. *International Journal of Productivity and Performance Management*, 70(8), 2352–2373. https://doi.org/10.1108/IJPPM-02-2020-0058

Williams, O.S. (2008). *New Product Introduction Process Improvements in An Automotive Company*. Available from: https://dspace.lib.cranfield.ac.uk/handle/1826/6580

Wind, J., & Mahajan, V. (1997). Issues and opportunities in new product development: An introduction to the special issue. *Journal of Marketing Research*, 34(1), 1–12. https://doi.org/10.1177/002224379703400101

Wuest, T., Liu, A., Lu, S.C.Y., & Thoben, K.D. (2014). Application of the stage gate model in production supporting quality management. *Procedia Cirp.* 17, 32–37. https://doi.org/10.1016/j.procir.2014.01.071

Wynn, D.C., & Clarkson, P.J. (2018). Process models in design and development. *Research in Engineering Design*. 29, 161–202. https://doi.org/10.1007/s00163-017-0262-7

Yadav, O.P., Nepal, B.P., & Jain, R. (2007). Managing product development process complexity and challenges: A state-of-the art review. *Journal of Design Research*, 6(4), 487–508. https://doi.org/10.1504/JDR.2007.016853

Yang, C.L., Huang, K.P., Huang, C.L., & Huang, R.H. (2016). Integration of corporate social responsibility into new product development. *International Journal of Management Economics and Social Sciences*, 5(1), 14–34. https://doi.org/10.2139/ssrn.2764236

Yin, C., & Zhang, W. (2021). New product development process models. In: *International Conference on E-Commerce and E-Management (ICECEM) Dalian, China*. p240–243. https://doi.org/10.1109/ICECEM54757.2021.00054

Zhu, X., Xiao, Z., Dong, M.C., & Gu, J. (2019). The fit between firms' open innovation and business model for new product development speed: A contingent perspective. *Technovation*, 86, 75–85. https://doi.org/10.1016/j.technovation.2019.05.005

DOI: 10.6977/IJoSI.202506_9(3).0004
B.G. Shinde, S.B. Sanap, et al./Int. J. Systematic Innovation, 9(3), 31-74 (2025)

AUTHOR BIOGRAPHIES

Balasaheb Shinde is Ph.D scholar at MIT-ADT University, Pune, India. He has completed his post-graduation in CAD/CAM and graduated in Mechanical Engineering from Savitribai Phule Pune University, Pune. He has served Mahindra and Mahindra for 11 years and was responsible for part development and materials management. Currently, he is working as an Assistant. Prof. at Amrutvahini College of Engineering, Sangamner (Maharashtra, India) since August 2014. His research interest includes product design and development, materials, and quality management. Karnataka, INDIA in 2014.

Dr. Sudarshan Sanap is currently working as Dean of the Faculty of Engineering and Sciences at MIT Art, Design and Technology University, Pune, India. He completed his postgraduate studies in Design Engineering and a PhD in Mechanical Engineering. During his 25 years of service, he has published more than 50 research papers in national and international peer-reviewed conferences and journals. Quality in education at the Higher Education Institute is the ultimate vision.

Dr. Sachin Pawar is currently working as a Professor in the Mechanical Engineering Department at MIT Art, Design and Technology University, Pune, India. He completed his Ph.D. from North Carolina State University, USA, and worked at Tomsk Polytechnic University, Russia, for his postdoctoral research. During his 22 years of academic and research work experience, he has published research papers on Mechanical and Aerospace Engineering topics in peer-reviewed conferences and journals.

Dr. Vishnu Wakchuare is working as a Professor and Head of the Department of Mechanical Engineering at Amrutvahini College of Engineering, Sangamner, [Savitribai Phule Pune University, Pune]. He has completed his post-graduation in CAD/CAM from SGGS CoE&T Nanded, MBA from SPPU Pune, and PhD from SPPU Pune. He has 25 years of experience in engineering. He has published more than 22 journal papers and more than 30 papers in reputed international conferences, including AIMTDR, NAMRAC, IEEE, and ASME MSEC. His research interests span various domains, including competitive manufacturing, product design and development, additive manufacturing, and quality management.

DOI: 10.6977/IJoSI.202506_9(3).0005
H. Panigrahi, S. Naidu, etc./Int. J. Systematic Innovation, 9(3), 75-95 (2025)

Handwriting match and AI content detection

Hrishikesh Panigrahi, Siddhanth Naidu, Ambuj Pandey, Phiroj Shaikh*, Amiya Kumar Tripathy

Department of Computer Engineering, Don Bosco Institute of Technology, Mumbai, India

*Corresponding author E-mail: phiroj@dbit.in

(Received 22 November 2024; Final version received 12 February 2025; Accepted 09 April 2025)

Abstract

Machine-generated text presents a potential threat not only to the public sphere but also to education, where the authenticity of genuine students is compromised by the presence of convincing, synthetic text. There are also concerns about the spread of academic misconduct, particularly direct replication among students. In response to these challenges, this paper introduces the Handwriting Match and Artificial Intelligence (AI) Content Detection System (HMAC). HMAC utilizes optical character recognition (OCR) mechanisms to convert handwritten and typed content from a single-page portable document format into machine-readable text, thus enabling further analysis. Drawing on recent advances in natural language understanding, HMAC aims to preserve the educational value of assignments by effectively detecting AI-generated content. In addition, HMAC has a strong plagiarism detection system that uses a comparative analysis of student submissions in a particular academic field. This paper describes HMAC's architecture, methodology, and results, emphasizing its key contributions: improved handwritten content extraction with OCR and improved identification of AI-generated content. The study addresses the research question of how HMAC improves the identification of AI-generated content and supports academic integrity compared to other methodologies.

Keywords: Academic Assessment, Artificial Intelligence Content Detection, Document Analysis, Similarity Detection, Transformer-Based Models

1. Introduction

A new era of convenience and accessibility has been brought about by the widespread deployment of generative models, as demonstrated by ChatGPT, which has completely changed the landscape of academic assignments in recent years. However, there have been challenges along the way with this evolution, including a concerning trend of students using artificial intelligence (AI)-generated content directly in their assignments because of the ease of accessing such tools in obtaining the content required for the assignments. When students choose easily accessible generative content above the learning process, assignments lose their inherent value, which is a danger to the educational process.

In response to this critical issue, our research introduces the Handwriting Match and AI Content Detection System (HMAC) as a robust and multifaceted solution. HMAC is meticulously designed to recognize and address the inappropriate incorporation of AI-generated content in student assignments, particularly those submitted in handwritten or typed form, often in Portable Document Format (PDF) Romero et al. (2012). By leveraging transformer-based models fine-tuned on the GPT-wiki-intro dataset, HMAC employs state-of-the-art technology to discern the nuanced differences between human-generated and AI-generated content.

As AI systems become increasingly integrated into our daily lives, understanding the factors influencing their acceptance and interaction becomes paramount. Pelau et al. (2021) explored the intricate dynamics of human-AI interaction, examining the roles of interaction quality, empathy, and perceived anthropomorphic characteristics in shaping the acceptance of AI within the service industry. Daniel et al. (2019) delved into the delicate balance between AI and human behavior, emphasizing the need to comprehend and prevent potential harm that may arise from AI systems acting like humans. Uzun (2023) investigated academic integrity concerns related to ChatGPT, shedding light on methodologies for detecting AI-generated content in educational settings. In addition, Sadasivan et al. (2023) probed the reliability of detecting AI-generated text, questioning the effectiveness of existing methodologies.

In the field of language understanding, the advent of Bidirectional Encoder Representations from Transformers (BERT) marked a significant paradigm shift. Kenton and Toutanova (2019) introduced BERT as a pre-training model for deep bidirectional transformers, showcasing its prowess in language understanding. These endeavors collectively constitute the backdrop against which we explore the evolving landscape of AI and its intricate intersections with human dynamics.

Beyond content identification, HMAC's primary objective is to retain and impart the educational value that comes with academic assignments. HMAC positions generative models as supplemental tools, encouraging students to participate in the learning process before incorporating AI-generated content, as opposed to using it as a quick fix. This paradigm change fosters a deeper comprehension of the subject matter, which is essential for the holistic development of students.

At its core, HMAC employs an optical character recognition (OCR) mechanism to convert both handwritten and typed content into machine-readable text. This process involves several steps, including word detection through platforms such as Roboflow, an ordering algorithm to arrange words into coherent lines, and integration with OCR models such as Microsoft's transformer-based OCR (trocr)-base-handwritten, for precise word recognition and sentence formation. The system simultaneously applies an AI content recognition model to the OCR-processed text, revealing information about the proportions of AI and human content. The second aspect of our investigation delves into the world of OCR, a technology critical to deciphering and extracting information from visual data. Wu et al. (1997) laid a foundational groundwork for text extraction from images, a critical step in the OCR process. Subsequently, Li and Doermann (1998) pioneered automatic text tracking in digital videos, propelling OCR capabilities into dynamic visual contexts. Kim (1999) introduced local color quantization for automatic text location in complex color images, advancing OCR techniques in handling intricate visual scenarios. Jain and Yu (1998) further expanded OCR capabilities into images and video frames, contributing to the development of comprehensive OCR methodologies.

One of HMAC's unique features is its integration of a plagiarism check, which involves comparing uploaded assignments to a database of assignments related to the same subject. By taking a comprehensive approach, instances of academic dishonesty among students are reported, and AI-generated content is identified. The system generates detailed reports for users, including percentages of AI-generated content and plagiarism detection results.

HMAC provides a complete and proactive solution that essentially guards against the improper application of generative models in educational settings. Assignments are meant to be instructive. This platform provides educators with the resources they need to keep a strict evaluation environment in place. Examining the design, process, and outcomes of HMAC, this work offers a viable answer to the changing problems associated with academic integrity in the digital era of research.

In this work, we address the following research question: (i) How does HMAC significantly improve the detection of AI-generated work, especially in terms of academic integrity, compared to current approaches? (ii) How does the OCR aspect of our technology enhance the process of extracting and interpreting handwritten text from photographs, especially for assignments and academic documents?

2. Related Work

Several recent studies have focused on detecting and analyzing AI-generated content, particularly in terms of academic integrity and the responsible use of AI. Rodriguez et al. (2022) investigated the cross-domain detection of GPT-2-generated technical text, shedding light on the challenges and implications of identifying machine-generated content across different domains. Mitrović et al. (2023) explored decision explanations for machine learning models, specifically in the context of distinguishing between ChatGPT and human-generated text, emphasizing the importance of interpretability in AI-based detection systems.

Addressing academic integrity concerns, Uzun (2023) conducted an investigation into ChatGPT's impact on academic settings, proposing methods for detecting AI-generated content to preserve academic integrity. Mindner et al. (2023) delved into the classification of human- and AI-generated texts, exploring features for effectively identifying content generated by models such as ChatGPT.

Wahle et al. (2023) introduced the concept of AI Usage Cards, aiming to facilitate the responsible reporting of AI-generated content by providing transparent and informative details about the AI models involved. In the realm of evasion strategies, Lu et al. (2024) discussed the challenges of large language models being guided to evade AI-generated text detection, highlighting the need for robust detection mechanisms Jauhiainen et al. (2016), Lindén et al. (2012).

Several studies have made significant contributions to the field of OCR, with each focusing on a different aspect of accuracy, efficiency, and post-correction methodologies Lund et al (2013), Lund et al (2011), Reynaert (2010).

Dong and Smith (2018) proposed a multi-input attention mechanism for unsupervised OCR correction, demonstrating its effectiveness in enhancing OCR accuracy Wick et al. (2018). The similar type work has been done by Evershed and Fitch (2014), Hämäläinen and Hengchen (2019), Kauppinen (2016), Koistinen et al. (2017), Kettunen et al. (2018), Llobet et al. (2010), Silfverberg et al. (2016), Springmann et al. (2014), Vobl et al. (2014). Guha et al. (2019) presented Devnet, an efficient convolutional neural network architecture for handwritten Devanagari character recognition, contributing to the advancement of OCR techniques for specific script recognition. The effective use of the neural networks in such areas has been done by Graves et al. (2006) and Srivastava (2014).

Kettunen and Koistinen (2019) explored the utilization of the open-source OCR engine Tesseract for re-OCR of Finnish Fraktur, providing insights into quality improvement strategies. Li et al. (2021) introduced TrOCR, a transformer-based OCR model with pre-trained models, showcasing the potential of transformer architectures in OCR tasks.

Sabu and Das (2018) conference paper, "*A Survey on Various Optical Character Recognition Techniques*," was presented at the Conference on Emerging Devices and Smart Systems 2018 in India. This manuscript sheds light on various OCR methods and provides useful insights into their applications. It also recommends future research avenues into advanced OCR techniques and their potential applications in document analysis.

3. Related Work

3.1. Design

The HMAC proposed requires users to input their assignments in PDF format through a specific web interface. The system then uses an OCR process to convert typed or written text into machine-readable text. To provide users with visibility into both human and AI content percentages, the system simultaneously runs the OCR-processed text through an AI content detection model to ascertain the likelihood that the material was generated by AI. By comparing the submitted work with a database of assignments in the same subject, HMAC also includes a plagiarism check feature. The system then generates a thorough report for users, including the percentage of material generated by AI and the findings of any plagiarism detection.

3.2. Architecture

A single-page PDF document is first fed into the process and transformed into machine-readable text. To find duplicate material and establish whether the content is AI generated, this text is processed and put through match detection and AI content detection. The processed file is kept in a database together with its metadata and analysis findings. The system verifies that each file has been examined and compared. If not, a different file is retrieved from the database for processing. The final comparison findings and AI content detection results are displayed after all files have been analyzed.

The architecture depicted in Fig. 1 ensures a comprehensive and systematic approach to content analysis in PDF files. It seamlessly integrates PDF-to-text conversion, match detection, AI content detection, and database management, offering a robust solution for identifying duplicate and AI-generated content while maintaining the integrity of the analyzed material.

3.3. Method of Data Collection

A thorough literature search was conducted to identify relevant research articles on AI-generated content detection techniques. Databases such as IEEE Xplore, ACM Digital Library, arXiv, and Google Scholar were searched using relevant keywords, such as "Generative Models Analysis," "AI Content Detection," and "Optical Character Recognition (OCR)." Following our literature search, we carried out a two-step quality assessment process. In the first step, we screened the identified studies for relevance to the scope of our study. Manuscripts that did not directly address the topic or were unrelated to our HMAC project were not further considered.

Then, we rigorously evaluated their scientific quality. This evaluation included an assessment of each manuscript's methodology, including the research design, data collection techniques, and experimental setup. The articles were chosen based on their relevance to the topic and contribution to the existing body of knowledge in the field. Only peer-reviewed articles and conference papers were considered for the review. Duplicate articles, non-English articles, and articles irrelevant to the scope of HMAC were excluded.

4. OCR in Handwriting Match and AI Content Detection System

OCR is a critical stage in the complex framework of HMAC. OCR is the cornerstone in our effort to convert complex handwritten characters into text that can be read by a computer when integrating handwritten assignments into our system. Handwritten text has intricacies that must be extracted, processed, and understood. This intricate process is broken down into multiple steps, each with a distinct objective.

The first step is converting handwritten assignments in PDF format into image files. This conversion creates the foundation for further OCR processes using the pdf2image library. Then, using Roboflow, the system explores word detection, exposing the minute nuances of each word in the handwritten document. The challenges posed by variations in handwriting styles and irregular word sequencing are addressed using an innovative algorithm called the Word Sequencing Algorithm, ensuring accurate and ordered detection. Once the words are identified and ordered, the system uses OCR models, with a focus on Microsoft's trocr-base handwritten model. This specialized OCR model excels at recognizing and transcribing each word in handwritten content, bridging the analog and digital dimensions of education. The stages are outlined below and explained in detail.

The flowchart in Fig. 2 depicts the step-by-step process used in the OCR system.

4.1. Document Image Generation

The first step of the proposed software is to process a single-page PDF input in real-time. After uploading the PDF, the backend code quickly converts it to a JPG image using the pdf2image Python library. This library requires an external utility called Poppler to render PDFs. It is worth noting that while Poppler is commonly found on Linux systems, it may require additional steps to install on Windows. After saving the image, the system seamlessly proceeds to the next stage: semantic segmentation with Roboflow.

4.2. Semantic Segmentation with Roboflow

Following the successful conversion of the PDF to an image, the Word Detection stage focuses on precisely detecting each word within the converted image. For this task, we used Roboflow. Roboflow

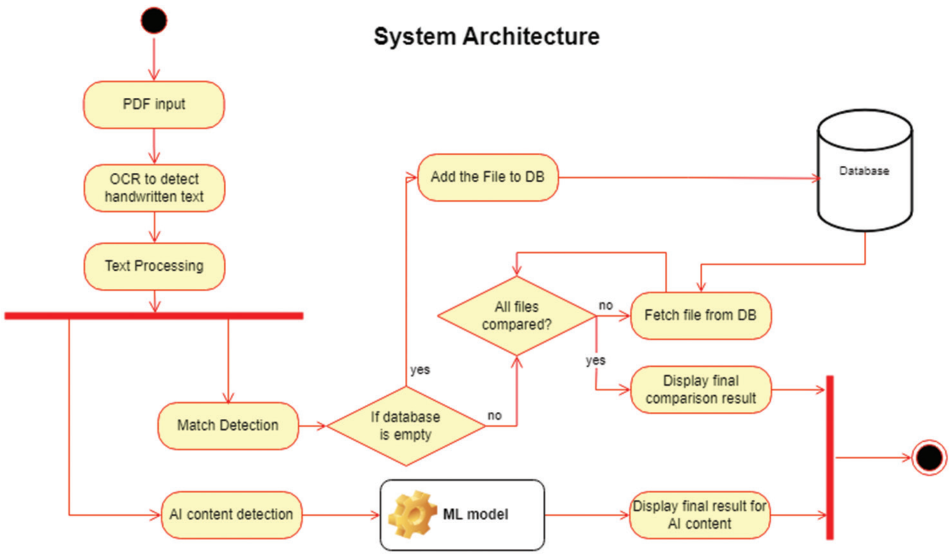


Fig. 1. Handwriting Match and Artificial Intelligence Content Detection System Architecture
Abbreviations: AI: Artificial intelligence; DB: Database; ML: Machine learning; OCR: Optical character recognition; PDF: Portable document format.

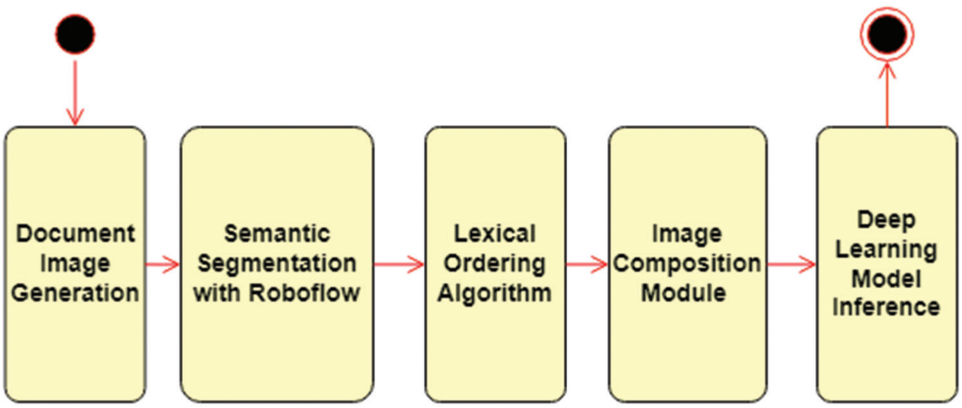


Fig. 2. Overview of the optical character recognition workflow

is the preferred option due to its tailored capabilities for detecting words in handwritten documents. The platform's robust capabilities are used to identify each word in the handwritten document.

The Roboflow model's results are stored in an array called "box_dimension," with each element representing the properties of a detected word: "*x*" for the *x* coordinate, "*y*" for the *y* coordinate, "width" for the word's width, and "height" for its height. Simultaneously, another array is created to hold the images for each detected word. These images are created using the data stored in the "box_dimension" array, with each property facilitating the extraction of a specific portion of the image. Each extracted word image is then saved to a designated folder.
For instance, the "box_dimensions" array looks like this.

[{'index': 1, 'x': 602, 'y': 360, 'width': 102, 'height': 39},
{ 'index': 2, 'x': 33, 'y': 144, 'width': 86, 'height': 35},
{ 'index': 3, 'x': 691, 'y': 706, 'width': 79, 'height': 34},
{ 'index': 4, 'x': 404, 'y': 613, 'width': 100, 'height': 36},
{ 'index': 5, 'x': 214, 'y': 707, 'width': 89, 'height': 28}]

However, a problem arises due to the potential disorder in the detection order, resulting in a mismatch between the sequence of words in the original PDF and the order in which they are detected. Although it might seem reasonable to sort the "box_dimension" array first by *x* values and then by *y* values, the second sorting step overrides the first. As a result, a more sophisticated algorithm is required to ensure proper word sequencing while preserving the original content's integrity.

4.3. Lexical Ordering Algorithm

The ordering algorithm 1 is critical for organizing the detected words and reconstructing the original flow of sentences and paragraphs from an unordered set of word images. This algorithm takes a systematic approach to grouping words into lines of text based on their vertical *y* positions in the document and then further groups them based on their horizontal *x* positions.

This algorithm is intended to arrange words into lines according to their spatial coordinates, specifically their vertical *y* and horizontal *x* positions within a given "box_dimensions" list. Initially, the list of word bounding boxes is sorted by *y* coordinates to group words on the same line. The algorithm iterates through each word, keeping track of which words appear on the same horizontal line. This is maintained by the current_line list. The algorithm uses a threshold value (20 in this case) to determine if a word belongs to the current line by comparing the *y* distance of the current word with the previous word in the current_line. If the difference in the y distance falls within the threshold, the word is added to the current_line; otherwise, the current_line is sorted by *x* coordinates (left-to-right order), appended to the lines list, and reset to start a new line with the current word. This process continues until all words are processed. Finally, the final current_line is sorted to the *x* coordinates and added to the lines list to ensure the algorithm correctly sequences words into lines as they appear in the text layout.

This algorithm is crucial for structuring the text, as it efficiently groups words into lines based on their vertical positions. It is critical for establishing the correct word sequence within each line, resulting in a structured representation of the original content.

4.4. Image Composition Module

The ordering algorithm is critical for organizing the detected words and reconstructing the original flow of sentences and paragraphs from an unordered set of word images. This algorithm takes a systematic approach to grouping words into lines of text based on their vertical *y* positions in the document and then groups them based on their horizontal *x* positions.

In the following stage, using the correct word sequence obtained in the previous step, an OCR model was used to determine the textual content of

Algorithm 1: Lexical Ordering

```
function WordSequencing (box_dimensions)
  box_dimensions ← Sorted (box_dimensions,
  key=lambda box: (box["y"]))
  lines ← []
  current_line ← [box_dimensions[0]]
  threshold← 20

  for i in range (1, len (box_dimensions)) do
      if box_dimensions[i]["y"]−current_line[−1]["y"]
      < threshold then
         ▷ Add the word to the current line if it is on
         the same line
            Append (current_line, box_dimensions[i])
         else ▷ Sort the current line by x-value and add
         it to the lines list
               current_line← Sorted (current_line,
  key=lambda box: (box["x"]))
            Extend (lines, current line)
            current_line ← [box_dimensions[i]] ▷
            Reset the current line to the current word
        end if
     end for

  ▷ Sort the last line by x-value and add it to the
lines list
   current_line←Sorted (current_line, key=lambda
   box: (box["x"]))
  Extend (lines, current_line)
  return lines
end function
```

each detected word. The traditional method involves processing each word individually through the OCR model, resulting in corresponding text outputs. However, this method can be time-consuming, particularly in cases where the document contains a large number of words.

To address this issue, the proposed software uses a strategic technique known as image concatenation. Image concatenation is the process of combining two or more images to create a single composite image. Rather than processing each word individually, the software combines 10 images into a single line using concatenation. This concatenated image is then fed into the OCR model for analysis. By concatenating multiple images into a single line, the software not only speeds up the OCR process but also increases overall efficiency. This technique is especially useful when dealing with a large volume of text, resulting in a more streamlined and resource-efficient workflow.

One notable challenge in implementing the image concatenation technique is the potential lack of discernible padding between two concatenated images. OCR models may find it challenging to recognize and distinguish individual words inside the combined image due to the absence of spacing. A calculated approach is used to solve this problem: putting white space between two concatenated photos. White padding improves the accuracy of OCR. This ensures that the OCR process can effectively interpret each word in the concatenated line, allowing these words to be seamlessly combined into a coherent sentence.

4.5. Deep Learning Model Inference

Following the concatenation of lines, each composite line is processed using Microsoft's trocr-base handwritten model. This OCR model is designed specifically for the recognition of handwritten text, using advanced deep-learning techniques to ensure precise identification and transcription of each word in the handwritten content. The model's emphasis on handwritten text recognition makes it an appropriate choice for accurately interpreting the complexities of handwritten assignments.

4.5.1. Model architecture

The TrOCR model is architecturally designed around the transformer framework, which includes both an image transformer and a text transformer. The dual-transformer architecture depicted in Fig. 3 is fundamental to TrOCR's ability to accurately extract visual features from images and perform language modeling for OCR (Li et al., 2021).

The transformer architecture is implemented in a standard encoder-decoder configuration within TrOCR. The encoder component is specifically engineered to capture representations of image patches, leveraging the visual information inherent in the input. The decoder, on the other hand, is responsible for generating a workpiece sequence, guided not only by the visual features extracted from the image but also by the predictions made in the preceding steps. TrOCR utilizes the conventional transformer encoder-decoder structure, which emphasizes its adaptability and effectiveness in dealing with both image-based and language-related tasks. This architectural choice highlights the transformer model's versatility, allowing it to seamlessly integrate image processing and language generation components within a unified framework. The above diagram presents the architectural design of the model.

- *Encoder:* The encoder part processes the input image and extracts high-level features that represent the content of the image. In the context of OCR, this could involve understanding the shapes and patterns of characters.
- *Decoder*: The decoder part interprets the features generated by the encoder and produces the final output, which is the recognized text. The decoder considers the context of characters and their relationships to improve accuracy. The decoder uses self-attention masking to prevent gaining more information during training than prediction. The attention mask ensures that the output for position *i* only considers the previous output and input for positions less than *i* (Li et al., 2021).

$$h_i = Proj(Emb(Token_i)) \tag{1}$$

$$\sigma\left(h_{ij}\right) = \frac{e^{h_{ij}}}{\sum_{k=1}^{v}\left(e^{h_{ik}}\right)} \qquad for\ j = 1, 2, ..., V \tag{2}$$

A linear layer projects the decoder's hidden states from the model dimension to the dimension of the vocabulary size *V*, and the softmax function calculates the probabilities over that vocabulary. We used beam search to obtain the final result.

5. AI Text Detection

5.1. Dataset

Rapid advancements in large language models such as GPT have brought enormous potential and unexpected challenges. One such challenge is the widespread use of GPT-generated text, which raises questions about its authenticity and potential for misuse. To meet this critical need, a dataset was used specifically for detecting text generated by GPT (Bhat, 2023). Table 1 provides a detailed breakdown of the dataset's columns, which include identifiers, Wikipedia URLs, titles, Wikipedia introduction paragraphs, and corresponding content generated by the GPT (Curie). It

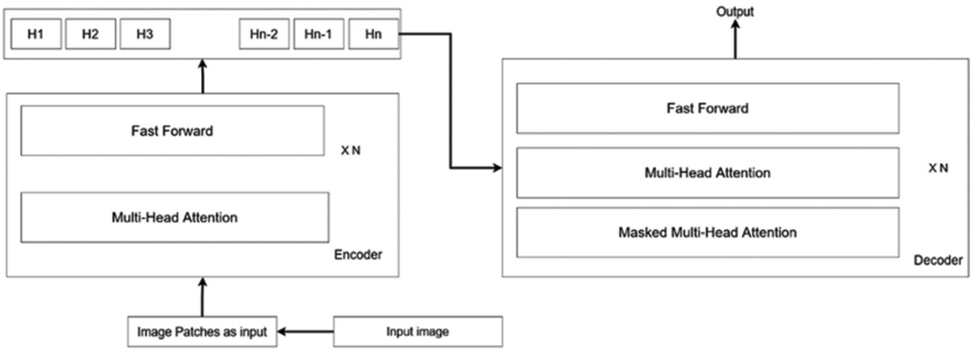


Fig. 3. Architecture of the transformer-based optical character recognition model

Table 1. Summary of dataset columns and their descriptions

Column	Data type	Description
Id	int64	ID
url	string	Wikipedia URL
title	string	Title
wiki_intro	string	Introduction paragraph from Wikipedia
generated_intro	string	Introduction generated by GPT (Curie) model
title_len	int64	Number of words in title
wiki_intro_len	int64	Number of words in wiki_intro
generated_intro_len	int64	Number of words in generated_intro
prompt	string	Prompt used to generate intro
generated_text	string	Text continued after the prompt
prompt_tokens	int64	Number of tokens in the prompt
generated_text_tokens	int64	Number of tokens in generated text

also includes length metrics and token counts for titles, Wikipedia introductions, generated introductions, prompts, and the text that follows the prompt, providing a comprehensive overview of the dataset structure.

5.1.1. Feature engineering

To optimize model performance and focus on the most salient features, we applied a strategic feature engineering process. This involved carefully selecting the following essential features:

(i) ID: A unique identifier assigned to each text introduction, enabling efficient data management and tracking

(ii) Text: The raw text content of the introduction, serves as the primary input for the GPT-detection model

(iii) Label: A binary classification label, explicitly indicating whether the text was generated by a human (0) or GPT (1). This label serves as the ground truth for model training and evaluation.

After modification, the dataset structure has been simplified for specific analysis. The “ID” column now displays the identifier in string format, while the “text” column combines data from both “wiki_intro” and “generated_intro.” The “Label” column, which is important for classification, is introduced, with 0 representing human-generated content and 1 representing AI-generated content. Table 2 displays the results of feature engineering applied to the dataset, which resulted in a consolidated “Text” column by combining content from “wiki_intro” and “generated_intro.” The “Label” column represents the classification label, with “0” indicating human-generated content and “1” indicating AI-generated content. These engineered features form the foundation for training and evaluating classification models that distinguish between human-generated and AI-generated text.

5.2. Model

The selection of an appropriate model was critical for the AI text detection process. Transformer-based models, particularly DistilBERT (Bidirectional Encoder Representations from Transformers), well-known for their ability to understand natural language, were evaluated and chosen for their adaptability to the task at hand. The chosen model was then meticulously trained using the annotated dataset. The dataset was divided into training, validation, and test sets, and the model’s performance was optimized based on continuous evaluation and validation results (Sanh et al., 2019).

To determine whether the input text was created by AI or by humans, the AI text detection model generates a probabilistic evaluation. The possibility that the text was produced by AI is represented by one percentage, whereas the likelihood that it was authored by humans is represented by another.

5.2.1. Model architecture

The architecture of the DistilBERT model, a condensed form of BERT, for AI text identification, is displayed in Fig. 4. It begins with an input layer that handles text sequences up to a certain maximum length (512 tokens, for example). The pre-trained DistilBERT model receives these text sequences as input after they have been transformed into token embedding.

The pre-trained model (DistilBERT) acts as the core component, which uses its bidirectional attention mechanism to comprehend contextual relationships in the input text. DistilBERT is effective for tasks such as text classification as it minimizes the number of parameters while maintaining the majority of BERT's representational capacity. After passing through the pre-trained model, the text representations are routed to a fine-tuned layer. This layer represents the specific modifications and task-specific training that were applied to the pre-trained model to help it adapt to the AI text detection task. The fine-tuned model's output is passed through a linear layer pre-classifier, reducing dimensionality and preparing the data for further processing. The representations are fed into a fully connected layer, which maps the features to the final classification task. The final stage is the topic tagging layer, which does the classification. It generates probabilistic scores that indicate whether the input text is human-generated (class 0) or AI-generated (class 1). These probabilities provide interpretable information about the text's likely source.

Table 2. Engineered features for classification

Column	Datatype	Description
ID	int64	ID
Text	string	Text taken from wiki_intro & generated_intro
Label	int64	0 for human, 1 for AI

In short, during the forward pass, the input IDs and attention mask are processed by the pre-trained DistilBERT layer, resulting in a series of hidden states. These hidden states are sent through a fine-tuned layer, a linear pre-classifier, and fully connected layers. The final output layer creates a two-dimensional vector that represents the probabilities for both human-generated and AI-generated text.

6. Similarity

A parallelized comparison methodology has been used in the test of uploaded assignments for similarity, a critical component of the HMAC. By utilizing concurrent futures and a ThreadPoolExecutor, it effectively compares the contents of an uploaded file with several entries in the database, saving the findings for additional examination. A ThreadPoolExecutor was used to start the parallel processing mechanism. ThreadPoolExecutor in Python facilitates concurrent task execution via threads, which is ideal for independent tasks such as file similarity comparisons. It optimizes resource use by reusing threads and scales tasks efficiently by adjusting thread pool size. Asynchronous execution enables tasks to run independently, enhancing responsiveness. Compared to multiprocessing, which uses separate processes with higher memory overhead, and manual threading, which requires more complex thread management, ThreadPoolExecutor offers a simpler, more efficient solution. It submits tasks in parallel to compare the content of the uploaded file with each file model in the list. The comparisons list stores the outcomes for later review and documentation.

6.1. Stages for Detecting Similarity

For similarity detection, first, HMAC calculates the TF-IDF for the documents and then uses cosine similarity to compare the number of similar terms present in the currently uploaded document to the documents present in the database.

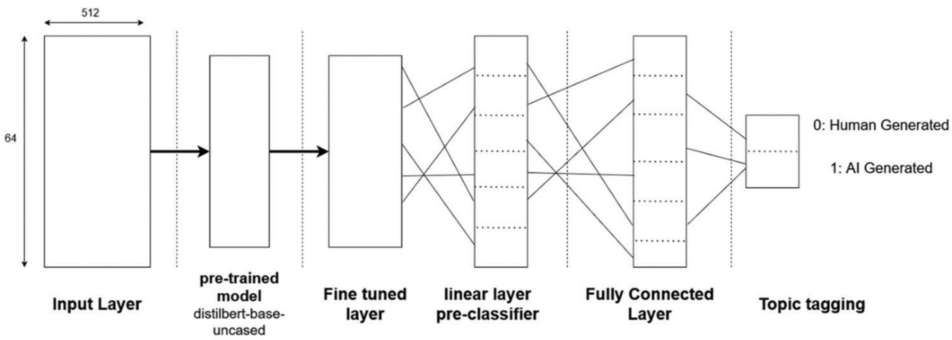


Fig. 4. Architecture of the artificial intelligence (AI) text detection model

6.1.1. Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF is a statistical measure that evaluates the importance of a term in a document relative to its occurrence in a collection of documents. The TF-IDF score for a term t in document d is calculated as follows:

TF−IDF(t,d,D)

$$TF\text{–}IDF(t, d, D) = TF(t, d) \times IDF\,(t, D) \tag{3}$$

$$IDF\left(t,d\right) = \log\left(\frac{Total\,no.of\;documents\,in\,corpus\,d}{No.of\;documents\;with\,term\,t+1}\right)+$$

$$1TF\left(t,d\right) = \frac{Number\;of\;terms\,t\;appears\,in\,document\;d}{Number\;of\;terms\,in\,document\;d} \tag{4}$$

where:

(i) TF (t, d) is the term frequency of term t in document d

(ii) IDF (t, D) is the inverse document frequency of term t in the document corpus d.

In HMAC, the TF-IDF matrix is utilized to represent the importance of each term in the uploaded assignment and other submissions, forming the basis for similarity calculations. Below is an excerpt from the output using a sample set of documents:

(0, 129) 0.1857772922586975
(0, 106) 0.1857772922586975
(0, 99) 0.1857772922586975
(0, 127) 0.1857772922586975
(0, 53) 0.1857772922586975
(0, 119) 0.1857772922586975
(0, 43) 0.1857772922586975
(0, 3) 0.1857772922586975…

In the example above, the line (0, 129) 0.1857772922586975 indicates that in the first document, the TF-IDF score for the term represented by the word at index 2 is approximately 0.1857772922586975.

The TF-IDF matrix is presented in a compressed sparse row format, an efficient representation for sparse matrices. Each line in the output corresponds to a non-zero entry in the matrix, with the following components:

(i) (i, j) value: Represents a non-zero entry in the matrix

(ii) i: Row index

(iii) j: Column index

(iv) value: TF-IDF score for the term in the document.

6.1.2. Document similarity

Document similarity scores are computed using the TF-IDF matrix as a base. The matrix's non-zero entries show how crucial particular terms are for differentiating between papers. Comprehending the TF-IDF score distribution offers valuable perspectives on the distinct attributes of every document. Document similarities can be interpreted more nuancedly when terms with higher TF-IDF scores are identified as having substantially contributed to the document's content.

In the context of HMAC, this metric serves as a robust indicator of similarity between the uploaded assignment and other submissions. A high cosine similarity implies a closer resemblance between the two documents. The cosine similarity scores are stored in a list of tuples, where each tuple is structured as (currentfile_index, db_file_index, similarity_score).

where,

(i) currentfile_index: Refers to the index of the currently uploaded file

(ii) db_file_index: Refers to the index of a document stored in the database

(iii) similarity_score: Represents the calculated similarity between the current file and the database file.

Each tuple has three fields, with the 0^{th} index being the currently uploaded document, the first index being the document present in the database, and the second index being the similarity score. Below is an example of this list with two documents present in the database and uploading a third.

For better visualization, these scores can also be represented as a similarity matrix, where each cell at position (currentfile_index, db_file_index) contains the similarity score. The sample data is presented in Table 3. [(0, 1, 0.025970408434077174), (0, 2, 1.0), (1, 2, 0.025970408434077174)].

The diagonal entries are marked as "-" since a document's similarity with itself was not computed. This matrix representation complements the tuple format, making it easier to identify relationships between documents.

Table 3. Cosine similarity matrix representing the similarity scores between the current file and the database file

Current file/ Database file	File 1	File 2	File 3
File 1	-	0.025970	1.0
File 2	0.025970	-	0.025970
File 3	1.0	0.025970	-

6.1.3. Parallel comparison

ThreadPoolExecutor in Python facilitates concurrent task execution through threads, which is ideal for independent tasks like file similarity comparisons. It optimizes resource use by reusing threads and scales tasks efficiently by adjusting thread pool size. Asynchronous execution enables tasks to run independently, enhancing responsiveness. Compared to multiprocessing, which uses separate processes with higher memory overhead, and manual threading, which requires more complex thread management, ThreadPoolExecutor offers a simpler, more efficient solution. It is particularly advantageous for central processing unit-bound or I/O-bound tasks, balancing ease of implementation with performance gains in Python applications. Choosing ThreadPoolExecutor depends on task characteristics and integration needs within existing codebases and libraries. Given below is the algorithm for file comparison.

Here are the definitions for the variables and terms used in Algorithm 2:

- executor: The ThreadPoolExecutor instance used to manage concurrent task execution
- file_list: A list containing the file models to be compared against the uploaded file
- future_to_filename: A dictionary that maps each future task to its corresponding file model's filename
- comparisons: A list to store the comparison data, including the uploaded file, other files, and similarity results
- file: The current file being processed in the loop
- future: Represents a submitted task
- filename: Corresponds to a future task
- result: The result retrieved from a future task
- comparison_data: A dictionary that stores the comparison data for each file comparison
- Exception as e: The exception caught during error handling

Algorithm 2. File similarity comparison

```
function InitializeExecutor
    Create a ThreadPoolExecutor as executor
    return executor
end function

function SubmitTasks (executor, file_list,
  future_to_filename)
   for each file in file list do
      future ← executor.submit (compare_file_
       similarity, file['content'], file['model'])
      future_to_filename[future] ← file['model']
   end for
end function

function ProcessCompletedTasks (futures, future_to_
    filename, comparisons)
   for each future in futures that are completed do
    filename ← future_to_filename[future]
    result ← future.result()
    if result is valid then
      comparison_data ← {'uploaded file':uploaded_
      file_data, 'other_file': filname, 'similarity
      result':result}
      comparisons.append (comparison_data)
    else
      Error message for being unable to calculate
      similarity
    end if Exception as e
          Error message with exception details: e
  end for
end function

function Main (uploaded_file_data, file_list)
    executor ← InitializeExecutor
    future_to_filename ← {}
    SubmitTasks (executor, file_list,
    future_to_filename)
   comparisons ← []
    ProcessCompletedTasks (futures, future_to_
    filename, comparisons)
    Return comparisons
end function
```

7. Results

7.1. OCR

7.1.1. Document image generation

The primary purpose of this stage was to provide a standardized format for further analysis and to ensure uniformity in the subsequent stages of the project. The input at this stage is a single-page PDF file, as shown in Fig. 5.

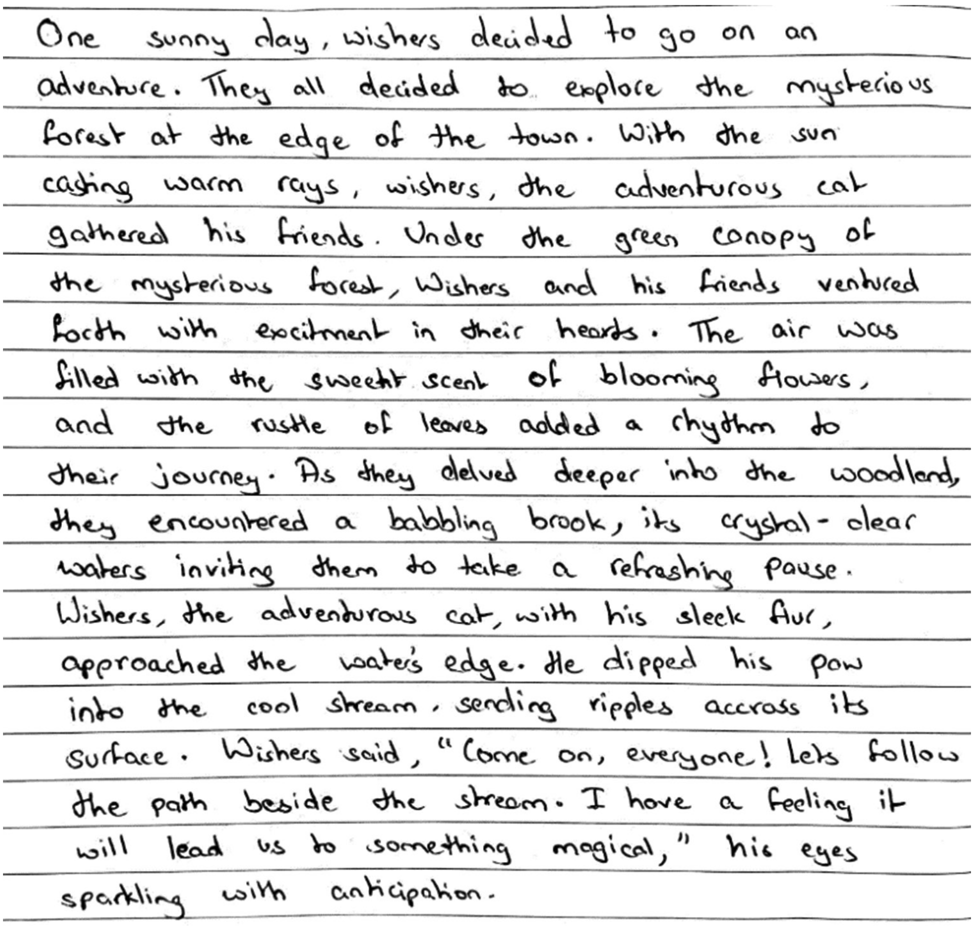
7.1.2. Semantic segmentation with roboflow

This critical phase required the precise identification of individual words within the converted document images. This stage was critical in preparing the data for further processing, ensuring that the subsequent OCR phase focused on accurately segmented regions, allowing for more granular and precise extraction of text content from document images.

The properties of the segmented images, such as their x and y coordinates, width, and height, were systematically stored in the "box_dimensions" array. This array was critical in maintaining the spatial information of each word in the document.

For Fig. 5, the following shows its "box_ dimension" array.

[{'index': 1, 'x': 602, 'y': 360, 'width': 102, 'height': 39},



One sunny day, wishers decided to go on an
adventure. They all decided to explore the mysterious
forest at the edge of the town. With the sun
casting warm rays, wishers, the adventurous cat
gathered his friends. Under the green canopy of
the mysterious forest, Wishers and his friends ventured
forth with excitment in their hearts. The air was
filled with the sweeht scent of blooming flowers,
and the rustle of leaves added a rhythm to
their journey. As they delved deeper into the woodland,
they encountered a babbling brook, its crystal-clear
waters inviting them to take a refrashing pause.
Wishers, the adventurous cat, with his sleek fur,
approached the water's edge. He dipped his paw
into the cool stream, sending ripples accross its
surface. Wishers said, "Come on, everyone! Lets follow
the path beside the stream. I have a feeling it
will lead us to something magical," his eyes
sparkling with anticipation.

Fig. 5. Document image generation: standardizing input for subsequent stages. This image represents the outcome of the first stage in the optical character recognition workflow, which focuses on document image generation

{'index': 2, 'x': 33, 'y': 144, 'width': 86, 'height': 35},
{'index': 3, 'x': 691, 'y': 706, 'width': 79, 'height': 34},
….
{'index': 154, 'x': 559, 'y': 372, 'width': 23, 'height': 14},
{'index': 155, 'x': 390, 'y': 660, 'width': 17, 'height': 14},
{'index': 156, 'x': 608, 'y': 747, 'width': 15, 'height': 13}]

The processed words were systematically organized and saved in a specific folder. Each word was given a unique title, denoted as "words_Index," with the index corresponding to the word's position in the "box_dimensions" array. The table 4 depicts one such example Roboflow model's output. This indexing system provided a seamless way to access the associated images by directly referencing the "box_dimensions" array.

7.1.3. Lexical ordering algorithm

This stage addressed the challenge of ensuring the correct sequence of words within the document, particularly when dealing with unordered or randomly detected words. The algorithm aimed to arrange the words in a sequence that accurately reflected the original order in the document.

After applying this algorithm, the "box_dimension" array for Fig. 6 is updated as follows:

[{'index': 18, 'x': 25, 'y': 9, 'width': 65, 'height': 32},
{'index': 117, 'x': 126, 'y': 19, 'width': 80, 'height': 27},
{'index': 74, 'x': 223, 'y': 7, 'width': 74, 'height': 37},
...
{'index': 78, 'x': 52, 'y': 793, 'width': 113, 'height': 35},
{'index': 82, 'x': 195, 'y': 790, 'width': 64, 'height': 23},
{'index': 126, 'x': 293, 'y': 787, 'width': 155, 'height': 34}]

The system's systematic word indexing allows for quick access to individual words by index number. For example, if a word is given the index number 10, it indicates that it is the first word in the corresponding PDF section. Based on the array after applying the algorithm to Fig. 5, the ordered words are produced as depicted in Table 5.

DOI: 10.6977/IJoSI.202506_9(3).0005
H. Panigrahi, S. Naidu, etc./Int. J. Systematic Innovation, 9(3), 75-95 (2025)

Table 4. Word detection and indexing results. This table displays the Roboflow model's output, demonstrating the detection of words in random order across various images

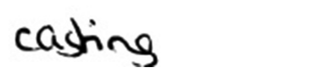
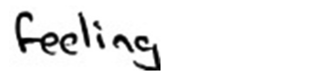
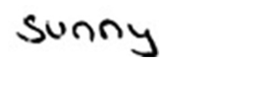
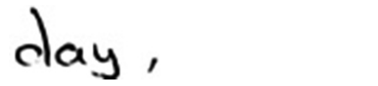
Image			
Title (words_Index)	words_1.jpg	words_2.jpg	words_3.jpg

Table 5. Ordered words extracted after applying the lexical ordering algorithm. This table presents the first three words extracted in order after applying the lexical ordering algorithm (Section 4.3)

Image			
Title (words_Index)	words_18.jpg	words_117.jpg	words_74.jpg

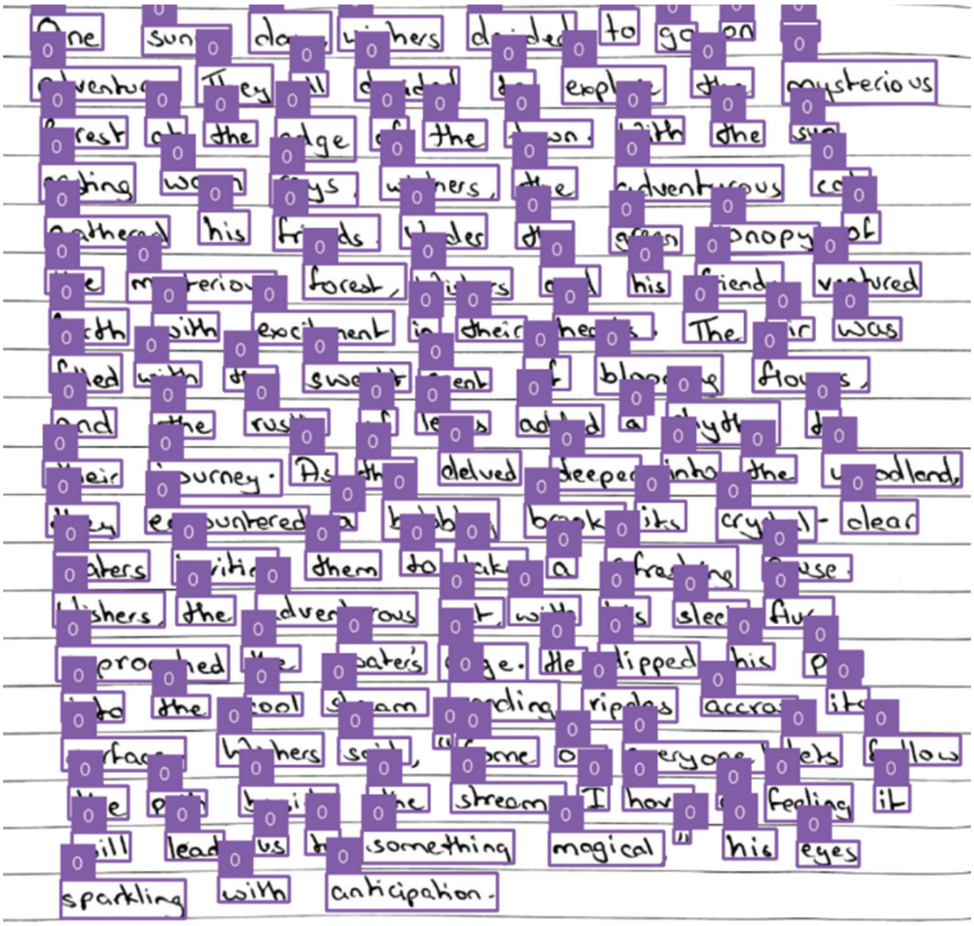


Fig. 6. Semantic segmentation with Roboflow—refined output. This image shows the results of the semantic segmentation phase using Roboflow

7.1.4. Image composition module

Rather than processing each word image individually, this method involved concatenating groups of 10-word images into a single, continuous line.

For Fig. 5, this stage produced the following output (Fig. 7), based on a batch size of 11.

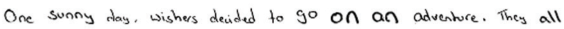


Fig. 7. Image composition module output: optimized concatenation. This image depicts the output of the image composition module (Section 7.1.4). It displays the first 11 words, as the specified batch_size was 11, demonstrating the effectiveness of the optimized concatenation approach

DOI: 10.6977/IJoSI.202506_9(3).0005
H. Panigrahi, S. Naidu, etc./Int. J. Systematic Innovation, 9(3), 75-95 (2025)

7.1.5. Deep learning model inference

The image composition module generated concatenated lines, which were then fed into the OCR model for recognition and transcription. Using the transformer's encoder-decoder structure, the model successfully converted visual features extracted from images into a coherent sequence of word pieces, effectively reconstructing the original handwritten content.

For Fig. 6, the model generates the following output:

One sunny day, wishes decided to go on ACT Adventure. They all decided to explore the mysterious forest at the edge of the town with the SWR. Casting warm rays. Wishers, the adventurous cat gathered his friends. Under the Green Canopy of the mysterious forest, Wishers and his friends ventured forth with excitement in their hearts. The air was filled with the sweet scent of blooming flowers, and the rustle of leaves added a rhythm to their journey. As they delved deeper into another woodland, they encountered a balding brook, its crystal clear waters inviting them to take A refastening pause. wishers, the adventurous cat, with his sleek fur, approached the water's edge. He dipped his POW into the cool stream sending rifles accrabs its surface Wishers said "Come On, everyone! Let us follow the path beside the stream. I have A feeling it will lead the US to something magical, It his eyes sparkling with anticipation.

7.1.6. Evaluation

To evaluate the results, we used three error measurements: character error rate (CER), word error rate (WER), and word recognition accuracy (WRA). These evaluation metrics is replicating the performance indices as mentioned in Appendix 1.

The CER, or percentage of erroneous characters in the system output, is a common metric in OCR tasks. It can be computed by dividing the number of incorrect characters by the sum of correct characters and errors in the system output. Similarly, the WER represents the percentage of incorrect words in the system output. It is computed by dividing the number of incorrect words by the sum of correct words and system errors.

$$\text{CER, WER}=\frac{\text{Errors}}{\text{Correct}+\text{Errors}} \tag{5}$$

Similar to WER, WRA measures the accuracy of whole-word recognition. It is calculated as the ratio of correctly recognized words to the total number of words.

$$\text{WRA}=\frac{\text{Correctly recognized words}}{\text{Total words}} \tag{6}$$

To determine the number of errors, we first aligned the ground truth sentence and OCR prediction lines at the character level (both CER and WER). We then calculated the overall Levenshtein distance between the system output and the ground truth, considering deletions and insertions (Li et al., 2021).

While calculating the CER is relatively simple, different evaluation systems employ different alignment approaches when calculating WER. Fig. 8 illustrates one alignment of a misspelled word example. The alignment is character-level, so the missing letters "e" and "x" will be paired with the empty string. If this alignment is used to calculate WER, the word example will be paired with the entire word exam, resulting in one error.

7.1.6.1. Handwriting quality assessment

It is important to note that the experiment used the same type of handwriting for all qualities. For instance, to maintain consistency in the evaluation, "Good" handwriting only includes text from one individual.

7.1.6.2. Factors affecting handwriting quality

To assess any handwriting, it is essential to know the elements that affect handwriting quality. Every element, from word spacing to writing style, influences our system's overall efficacy in a variety of handwriting attributes. Table 6 presents the factors that are taken into consideration for handwriting quality.

7.1.6.3. Word-level error analysis

Word-level error analysis is a thorough examination of our handwriting recognition system's

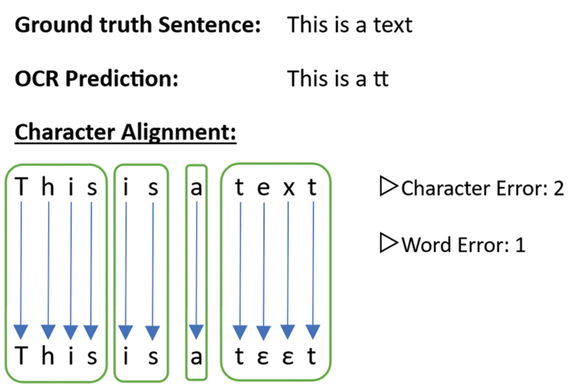


Fig. 8. An example of how the number of word errors varies according to alignment. After aligning the lines at the character level, aligning them at the word level results in a word error count of 1 (e.g., the word example is aligned with t ε ε t, the empty string)
Abbreviation: OCR: Optical character recognition

performance at the level of individual words. We evaluated the recognition accuracy of each sample by comparing the actual text to the system's output. The CER and WER were calculated to assess the system's ability to correctly identify individual characters and entire words. In Table 7, the CER and WER percentages are shown.

7.1.6.4. Line-level error analysis

By evaluating the recognition accuracy of complete lines of text, line-level error analysis provides a more comprehensive view than just analyzing individual words. Each sample's real lines and those identified by the system were compared. We evaluated the whole line recognition accuracy using line CER and line WER metrics.

An error analysis was performed at the line level. We examined how many lines were recognized completely correctly and found that most of the lines did not have any errors. With an average of over 70%, every 12 lines were recognized correctly. In the set of incorrect lines, most contained only minor errors, typically due to common OCR confusion as described by Kissos and Dershowitz (2016), Levenshtein, (1966).

7.1.6.5. Evaluation table

In this comprehensive performance evaluation, we present the results of our HMAC system across a range of handwriting qualities. We designed a structured table with important indicators to provide a more detailed picture of our system's functioning. The table includes an analysis of the number of sentences, handwriting quality (best, good, and worst), total lines processed, processing time, average speed per sentence, and accuracy score.

Table 7 shows a notable relationship between handwriting quality and OCR performance metrics. As handwriting declines from "Best" to "Worst," processing time, average speed per sentence, and error rates (both CER and WER) increase significantly, while WRA decreases dramatically. Specifically, "Best" handwriting quality has the lowest CER and WER rates, 2.35% and 8.11%, respectively, and the highest WRA at 95.30%. In contrast, "worst" handwriting quality results in a significantly higher CER (33.22%) and WER (66.44%), with a drastic reduction in WRA to as low as 4.03%. This degradation is reflected in the average speed per sentence, which rises from around 22 seconds for "Best" handwriting to more than 39 seconds for "Worst." These patterns indicate that as

Table 6. Summary of key criteria for evaluating handwriting quality

Handwriting	Consistency in Y-coordinate	Spacing between words	Character legibility	Pen pressure and stroke consistency	Aesthetic appeal
Best	Same line	Consistent	Clearly legible characters	Uniform throughout the writing	Pleasing
Average	Some variation, within a defined threshold	Minor variations, within an acceptable range	Some variations not hinder overall legibility	Minor variations	Adequate with room for improvement
Worst	Significant variation	Irregular or excessive	Inconsistencies	Significant irregularities	Unattractive or messy

Table 7. Handwriting Match and Artificial Intelligence Content System Performance Evaluation

Number of sentences	3 (43 words, including punctuation)			6 (115 words, including punctuation)			11 (181 words, including punctuation)			Average (6.6 words)
Handwriting qualities	Best	Good	Worst	Best	Good	Worst	Best	Good	Worst	Average
Total lines	3	5	7	9	12	15	14	19	22	11.7
Processing time (seconds)	65.07	71.8	80.6	134.76	286.85	216.90	414.4	414.9	430.7	244.97
Average speed per sentence (seconds)	21.69	23.93	26.8	22.46	47.8	36.15	37.67	37.71	39.15	37.15
CER (%)	2.35	4.69	17.92	1.18	4.04	28.74	2.03	7.14	33.22	8.56
WER (%)	8.11	21.62	40.54	5	12	68.00	8.97	19.87	66.44	24.22
WRA (%)	95.30	83.70	10.81	95.60	83	2.00	78.21	53.85	4.03	57.76

Abbreviations: CER: Character error rate; WER: Word error rate; WRA: Word recognition accuracy.

handwriting clarity deteriorates, OCR systems struggle more, taking longer to process text and producing lower accuracy. This emphasizes the importance of handwriting quality in achieving efficient and accurate OCR, demonstrating that even minor deviations from optimal handwriting can have a significant impact on recognition performance.

The graphical representation of Table 7 offers a clear visual insight into our OCR system's performance across various handwriting qualities. This graphical representation facilitates an intuitive understanding of the system's capabilities and limitations by highlighting trends and disparities across various categories of handwriting quality. Fig. 9 shows distinct trends by plotting a weighted sum of key metrics: total lines, processing time (seconds), average speed per sentence (seconds), CER, WER, and WRA. The bar colors represent different sentence counts—3 sentences (blue), 6 sentences (green), and 11 sentences (red), while the black line represents the average sentence count for each metric.

7.1.7. Limitations

Handwriting recognition through OCR encounters several challenges, impacting its effectiveness in capturing and interpreting diverse styles of handwritten text. It struggles with text written directly on lines, as this deviates from the standard way of writing. In contrast, it performs very well when text is placed between lines. Furthermore,

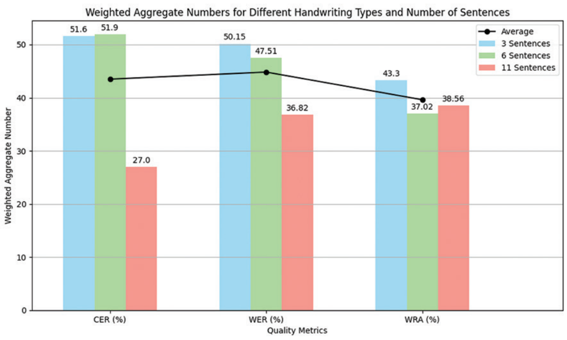


Fig. 9. Graphical representation of Table 7, offering clear visual insight for the untrained eye

Table 8. Performance measures before and after fine-tuning. This table compares key performance measures, such as accuracy and F1 score, before and after the fine-tuning process

Performance measures	Before fine-tuning	After fine-tuning
Accuracy	93.07	98.3
F1	0.68	0.84

OCR's reliance on perfectly aligned images makes it vulnerable to inaccuracies with tilted or rotated inputs. Page curvature complicates recognition because OCR is designed for flat pages and can struggle with curved surfaces. Inconsistencies in lined pages, as well as deviations from expected straight-line formation, have an impact on accuracy. A fixed threshold for recognizing handwriting does not account for the variability in individual styles, and poor handwriting can result in time-consuming processing, reducing real-time efficiency. It is obvious that additional developments in OCR technology are required considering these constraints. It is imperative to tackle these challenges to maximize the practical uses of OCR and enhance its flexibility to the ever-changing scenarios posed by handwritten documents in the real world.

7.2. AI Content Detection

The trained AI model achieved commendable accuracy in distinguishing between human-generated and AI-generated text, as shown in Table 7. After a meticulous training process on a dataset tailored to the study's objectives, the model demonstrated a strong ability to flag content with the signature characteristics of AI-generated language.

The OCR stage extracts textual content from images, which is then used as input for the AI content detection model within the HMAC system. The AI content detection model, which is primarily based on advanced transformer-based architecture, examines the provided text to determine whether it was generated by a human or by AI. The findings of this detection process are complex and offer insightful information about the text in question. The model calculates the percentage of content that is attributed to human authorship and the percentage that has the characteristics of AI-generated language. The output's dual nature serves as a quantifiable breakdown, clearly showing the proportion of the submitted assignments generated by AI versus those written by humans. The text shown in Fig. 7 was generated using ChatGPT, and when processed by the model, it produced the following result.

- Class 0: 01.20%
- Class 1: 98.80%.

7.3. Performance Metrics

The performance of the AI content detection model was evaluated at two stages: before and after fine-tuning. The model was refined, allowing it to distinguish between human-generated and AI-generated content more accurately. The model's remarkable accuracy of 93.07% before fine-tuning demonstrates its capacity to accurately classify the source of content in each

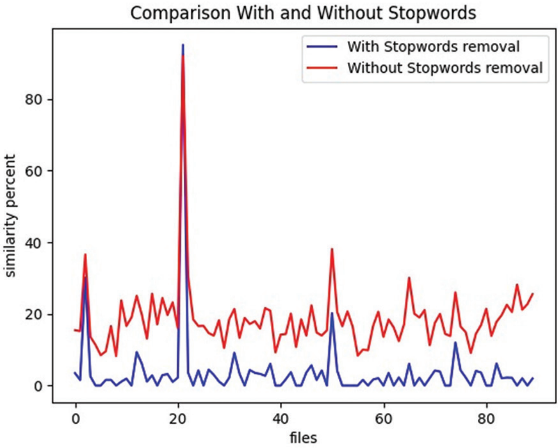


Fig. 10. Comparison of similarity percentage with and without stopwords

assignment. After applying fine-tuning techniques, its accuracy dramatically increased to 98.3%. Moreover, significant improvements were observed in the F1 score, a measure that strikes a balance between recall and precision. Before fine-tuning, the F1 score was 0.68, indicating a satisfactory performance. After fine-tuning, it rose dramatically to a remarkable 0.84. The result are well presented in Table 8.

7.4. Similarity Check

For similarity detection, the text generated from the handwriting detection step was compared to the previously submitted assignments stored in the database.

Fig. 10 depicts the change in similarity detection results before and after removing stopwords.

For this graph, a total of 21 files were used for comparison. The contents of 10 files were completely unique, while the remaining files contained similar to the first set of 10 files. The graph clearly shows that removing stopwords significantly decreased the similarity percentage. The mean difference in similarity was calculated to be around 14.905 using the specified formula.

8. Conclusion and Recommendations

This multiple-layered strategy demonstrates that HMAC is a useful platform that not only detects the use of AI-generated content but also actively addresses issues on academic integrity. The technology provides students with an easy platform to submit assignments, streamlining the assessment process for professors. By doing so, HMAC provides teachers with the resources they need to properly detect duplicate and AI-generated content, ensuring an impartial and rigorous assessment environment. HMAC serves as a preventive measure against the inappropriate use of generative models in educational contexts and is crucial in preserving the learning objectives of assignments by providing teachers with a useful tool for assessment. It helps to maintain the educational value of assignments while actively discouraging the misuse of generative models. HMAC acts as a vital safety net as education evolves in the digital age, ensuring that assignments fulfill their purpose of fostering authentic learning experiences.

The contributions of this paper include the following:

(i) Improved handwritten content extraction with OCR: This paper describes how our OCR component improved the extraction and interpretation of handwritten content from images, particularly in the context of academic documents and assignments. This includes improvements in accuracy and efficiency, which add to the overall landscape of document digitization.

(ii) Enhanced identification of AI-generated content: Compared to other approaches, this research demonstrates how HMAC considerably enhances the recognition of AI-generated content while upholding academic integrity.

Expanding the system's support for multi-page PDFs is a critical priority, necessitating optimizations in the OCR, content detection, and plagiarism-checking modules to enable seamless navigation of assignments spanning multiple pages. The applicability of HMAC to a wider range of academic resources would be significantly increased through this modification. The future holds promise for improving HMAC's understanding of semantic context and enabling more nuanced contextual analysis. Incorporating sophisticated natural language processing techniques would allow the system to not only identify content but also comprehend its meaning, fostering a better understanding of assignments and their legitimacy. Implementing a strong user feedback mechanism is an important part of future development. Allowing users to provide feedback on AI content detection accuracy and plagiarism checks promotes iterative refinement. A more user-centric and efficient system can be created with the assistance of user-generated suggestions for handling particular assignment types or enhancing the user interface. Furthermore, the integration of HMAC with Learning Management Systems shows promise for streamlining assignment submission and analysis processes in educational institutions. Developing plugins or application programming interfaces for seamless integration with Moodle, Canvas, and Blackboard, among other educational platforms, improves HMAC's accessibility and usability in educational ecosystems.

In addition, incorporating features that provide users with insights into the decision-making processes of AI content detection models, thereby increasing transparency and trust, is an important consideration for future iterations of HMAC. This comprehensive vision for future development establishes HMAC as an evolving, adaptable, and user-friendly system at the forefront of content analysis and plagiarism detection in educational settings.

References

Bhat, A. (2023). *GPT-Wiki-Intro (Revision 0e458f5). Hugging Face*. Available from: https://huggingface.co/datasets/aadityaubhat/gpt-wiki-intro [Last accessed on 2024 May 24].

Daniel, F., Cappiello, C., & Benatallah, B. (2019). *Bots Acting Like Humans: Understanding and Preventing Harm*. Available from: https://www.floriandaniel.it/papers/danielic2019.pdf [Last accessed on 2024 May 24].

Dong, R., & Smith, D.A. (2018). Multi-input attention for unsupervised OCR correction. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*. Vol. 1. p2363–2372.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*: Human Language Technologies, Volume 1 (Long and Short Papers), 4171–4186.

Englmeier, T., Fink, F., & Schulz, K.U. (2019). AI-PoCoTo-combining automated and interactive OCR postcorrection. In: *Proceedings of the Third International Conference on Digital Access to Textual Cultural Heritage*. ACM.

Evershed, J., & Fitch, K. (2014). Correcting noisy OCR: Context beats confusion. In: *Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage*. ACM, p45–51.

Graves, A., Fernández, S., Gomez, F., & Schmidhuber, J. (2006). Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In: *Proceedings of the 23rd International Conference on Machine Learning*. ACM, p369–376.

Guha, R., Das, N., Kundu, M., Nasipuri, M., & Santosh, K. (2019). Devnet: An efficient cnn architecture for handwritten Devanagari character recognition. In: *International Journal of Pattern Recognition and Artificial Intelliegence*. World Scientific, Singapore.

Hämäläinen, M., & Hengchen, S. (2019). From the paft to the fiiture: A fully automatic NMT and word embeddings method for OCR post-correction. In: *Proceeding of International Conference on Recent Advances in Natural Language Processing*. INCOMA, p432–437.

Jain, A.K., & Yu, B. (1998). Automatic text location in images and video frames. In: *Proceeding of International Conference of Pattern Recognition.* ICPR, Brisbane, p1497–1499.

Jauhiainen, T.S., Linden, B.K.J., & Jauhiainen, H.A. (2016). Heli, a word-based backoff method for language identification. In: *Proceedings of the Third Workshop on NLP for Similar Languages Varieties and Dialects VarDial3*. Osaka, Japan, p12.

Kauppinen, P. (2016). *OCR Post-Processing by Parallel Replace Rules Implemented as Weighted Finite-State Transducers*. University of Helsinki, Finland.

Kettunen, K., & Koistinen, M. (2019). *Open Source Tesseract in re-OCR of Finnish Fraktur from 19th and Early 20th Century Newspapers and Journals-Collected Notes on Quality Improvement*. Digital Humanitarian Network, Virtual, p270–282.

Kettunen, K., Kervinen, J., & Koistinen, M. (2018). Creating and using ground truth OCR sample data for Finnish historical newspapers and journals. In: *Proceeding of DHN 2018 Digital Humanities in the Nordic Countries 3rd Conference*. Helsinki.

Kim, P.K. (1999). *Automatic Text Location in Complex Color Images Using Local Color Quantization*. Vol. 1. IEEE TENCON, p629-632.

Kissos, I., & Dershowitz, N. (2016). OCR error correction using character correction and feature-based word classification. In: *2016 12th IAPR Workshop on Document Analysis Systems (DAS)*. IEEE, p198–203.

Koistinen, M., Kettunen, K., & Kervinen, J. (2017). How to improve optical character recognition of historical finnish newspapers using open source tesseract OCR engine? In: *Proceedings of the LTC*. p279–283.

Koistinen, M., Kettunen, K., & Pääkkönen, T. (2017). Improving optical character recognition of finnish historical newspapers with a combination of fraktur and antiqua models and image preprocessing. In: *Proceedings of the 21st Nordic Conference on Computational Linguistics*. p277–283.

Levenshtein, V.I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10, 707–710.

Li, H., & Doermann, D. (1998). Automatic text tracking in digital videos. In: *Proceeding of IEEE 1998*

Workshop on Multimedia Signal Processing, Redondo Beach, California, USA, p21–26.

Li, M., Lv, T., Cui, L., Lu, Y., Florencio, D., Zhang, C., Li, Z., & Wei, F. (2021). *TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models*. arXiv, 2109, 10282.

Lindén, K., Silfverberg, M., Pirinen, T., Hardwick, S., Drobac, S., & Axelson, E. (2012). *HFST-An Environment for Creating Language Technology Applications*. *Studies in Computational Intelligence*. Springer, Berlin.

Llobet, R., Cerdan-Navarro, J.R., Perez-Cortes, J.C., & Arlandis, J. (2010). OCR post-processing using weighted finite-state transducers. In: *2010 20th International Conference on Pattern Recognition*. p2021–2024.

Lu, N., Liu, S., He, R., Wang, Q., Ong, Y.S., & Tang, K. (2024). *Large Language Models can be Guided to Evade AI-Generated Text Detection*. https://doi.org/10.48550/arXiv.2305.10847

Lund, W.B., Kennard, D.J., & Ringger, E.K. (2013). Combining multiple thresholding binarization values to improve OCR output. In: *Document Recognition and Retrieval XX*, Vol. 8658. International Society for Optics and Photonics, p86580R.

Lund, W.B., Walker, D.D., & Ringger, E.K. (2011). Progressive alignment and discriminative error correction for multiple OCR engines. In: *2011 International Conference on Document Analysis and Recognition*. IEEE, p764–768.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. https://doi.org/10.48550/arXiv.1301.3781

Mindner, L., Schlippe, T., & Schaaff, K. (2023). *Classification of Human- and AI-Generated Texts: Investigating Features for ChatGPT*. arXiv, 2308, 05341.

Mitrović, S., Andreoletti, D., & Ayoub, O. (2023). *ChatGPT or Human? Detect and Explain. Explaining Decisions of Machine Learning Model for Detecting Short ChatGPT-generated Text*. https://doi.org/10.48550/arXiv.2301.13852

Pelau, C., Dabija, D.-C., & Ene, I. (2021). *What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. Computers in Human Behavior*, 122, 106855.

Reul, C., Christ, D., Hartelt, A., Balbach, N., Wehner, M., Springmann, U., Wick, C., Grundig, C., Büttner, A., & Puppe, F. (2019). *Ocr4all-an Open-Source Tool Providing a (Semi-) Automatic OCR Workflow for Historical Printings*. https://doi.org/10.48550/arXiv.1909.04032

Reul, C., Springmann, U., Wick, C., & Puppe, F. (2018). *State of the art optical character recognition of 19th century Fraktur scripts using open source engines*. https://doi.org/10.48550/arXiv.1810.03436

Reynaert, M.W. (2010). Character confusion versus focus word-based correction of spelling and OCR variants in corpora. *International Journal of Documents Analysis and Recognition (IJDAR),* 14(2), 173–187.

Rodriguez, J.D., Hay, T., Gros, D., Shamsi, Z., & Srinivasan, R. (2022). *Cross-Domain Detection of GPT-2-Generated Technical Text*. Available from: https://aclanthology.org/2022.naacl-main.88 [Last accessed on 2024 May 24].

Romero, V., Toselli, A.H., & Vidal, E. (2012). *Multimodal Interactive Handwritten Text Transcription*. Vol. 80. World Scientific, Singapore.

Sabu, A. M., & Das, A. S. (2018). A survey on various optical character recognition techniques. *In Proceedings of the 2018 International Conference on Emerging Devices and Smart Systems* (ICEDSS) (pp. 1–5). IEEE. https://doi.org/10.1109/ICEDSS.2018.8544323

Sadasivan V.S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). *Can AI-Generated Text be Reliably Detected?* https://doi.org/10.48550/arXiv.2303.11156

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). *DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter*. https://doi.org/10.48550/arXiv.1910.01108

Silfverberg, M., Kauppinen, P., & Lindén, K. (2016). Data-driven spelling correction using weighted finite-state methods. In: *Proceedings of the SIGFSM Workshop on Statistical NLP and Weighted Automata*. Association for Computational Linguistics, Berlin, p51–59.

Springmann, U., & Lüdeling, A. (2016). *OCR of Historical Printings with an Application to Building Diachronic Corpora: A Case Study Using the RIDGES Herbal Corpus*. https://doi.org/10.48550/arXiv.1608.02153

Springmann, U., Najock, D., Morgenroth, H., Schmid, H., Gotscharek, A., & Fink, F.(2014). OCR of historical printings of latin texts: Problems, prospects, progress. In: *Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage*. ACM, p71–75.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929–1958.

Uzun, L. (2023). *ChatGPT and Academic Integrity Concerns: Detecting Artificial Intelligence Generated Content*. Available from: https://www.researchgate.net/publication/370299956-chatgpt-and-academic-integrity-concerns-detecting-artificial-intelligence-generated-content [Last accessed on 2024 May 24].

Vobl, T., Gotscharek, A., Reffle, U., Ringlstetter, C., & Schulz, K.U. (2014). Pocoto-an open source system for efficient interactive postcorrection of OCRed historical texts. In: *Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage*. ACM, p57–61.

Wahle, J. P., Ruas, T., Mohammad, S. M., Meuschke, N., & Gipp, B. (2023). *AI Usage Cards: Responsibly reporting AI-generated content* [Conference poster]. 2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Santa Fe, NM, USA. https://doi.org/10.1109/JCDL57899.2023.00060

Wick, C., Reul, C., & Puppe, F. (2018). *Calamari-a High-Performance Tensorflow-Based Deep Learning Package for Optical Character Recognition*. https://doi.org/10.48550/arXiv.1807.02004

Wick, C., Reul, C., & Puppe, F. (2018). Comparison of OCR accuracy on early printed books using the open source engines Calamari and OCRopus. *Journal for Language and Conputational Linguistics*, 33, 79–96.

Wu, V., Manmatha, R., & Riseman, E.M. (1997). Finding text in images. In: *Proceedings of Second ACM International Conference on Digital Libraries*. Philadelphia, PA,p23–26.

AUTHOR BIOGRAPHY

Hrishikesh Pramod Panigrahi is currently pursuing a Bachelor of Engineering in Computer Engineering at The Bombay Salesian Society's Don Bosco Institute of Technology, Mumbai University. He has demonstrated strong technical expertise through hands-on experience in backend development, machine learning, and web applications. Hrishikesh has led backend teams, contributed to AI-driven academic integrity tools, and developed innovative mobile and web solutions. His technical proficiency includes programming languages such as Java, GoLang, Python, and Kotlin, and frameworks such as Django, ReactJS, and TensorFlow. He has published a research paper titled "Smart Posture Analyzer For Exercise" at the Institute of Electrical and Electronics Engineering (IEEE). His areas of interest include artificial intelligence, system design, and full-stack development.

Siddhanth Naidu is a Computer Engineering student at The Bombay Salesian Society's Don Bosco Institute of Technology, Kurla (Mumbai University). He has worked on impactful projects such as FLMS (Flood Location, Management, and Solution), an IEEE-published flood management system, and Dawn of Survival, an Android game developed in Unreal Engine. His technical skills include Java, Spring Boot, Python, ReactJS, HTML/CSS, Unreal Engine, and Unity. He has presented research at the International Conference on IoT, Communication, and Automation Technology (ICACAT) 2023 and was part of the ACM (Aluminum Composite Material) Design Team at Don Bosco Institute of Technology, Kurla. Siddhanth's interests lie in AI, web development, system design, and game design.

Ambuj Pandey is a Computer Engineering graduate from The Bombay Salesian Society's Don Bosco Institute of Technology, Kurla (Mumbai University). He has worked on impactful projects such as EDT, Expiry (Expiry Date Tracker, an Android app to track the expiry of purchased groceries), FLMS (an IEEE-published flood management system), and Dawn of Survival (an Android game developed in Unreal Engine) during his college years. His technical skills include C++, Python, Docker, Django, ReactJS, HTML/CSS, and Unreal Engine. He has published a research paper at ICACAT 2023. Ambuj's interests lie in AI, web development, and game design.

Phiroj Shaikh is currently an Associate Professor at the Department of Computer Engineering, The Bombay Salesian Society's Don Bosco Institute of Technology, Mumbai, India, affiliated with the University of Mumbai. He earned a PhD degree in Computer Science & Technology (in the domain of Web Data Mining) from Nagpur University, Nagpur, India. He has academic experience of over two decades with more than 60 research publications in journals/conferences. His areas of interest include natural language processing (with a special focus on regional language development), computational linguistics, and algorithm analysis. He is currently focusing on Educational Technology research area. He has been invited as an expert for faculty development programs, as a session expert, and as a reviewer at various conferences.

Amiya Kumar Tripathy is currently a Professor at the Department of Computer Engineering, The Bombay Salesian Society's Don Bosco Institute of Technology, Mumbai, India, affiliated with the University of Mumbai. He earned a PhD degree in Computer Science & Engineering (in the domain of Data Mining & Wireless Sensor Networks) from the Indian Institute of Technology Bombay, Mumbai, India. He was an adjunct professor at the School of Science, Edith Cowan University (ECU), Australia. He was a visiting researcher at the Rajamangala University of Technology, Bangkok, Thailand, working on Internet of Things (IoT)-enabled remote monitoring for the Precision Agriculture Farming project. He has been in the software industry, research, and academia for more than two decades, having around 150 publications in journals/conference papers. His research focuses on data science, computer vision, remote sensing, and IoT for Precision Agriculture. He has contributed to numerous collaborative research and consultancy projects in the domain of data analytics in India and abroad. He has served on the technical program committees of several international conferences, been invited as a plenary speaker, and co-chaired sessions at various conferences.

Appendix

Appendix 1. Explanation of Precision, Recall, and F1-score

In this appendix, we provide an explanation of precision, recall, and F1-score, which are used as evaluation metrics for HMAC.

(i) *Precision:* A measure of the accuracy of positive predictions. It is defined as the ratio of true positives to the sum of true positives and false positives.

$$\text{Precision}=\frac{\text{No. of true positives}}{\text{No. of true positives + No. of false positives}}$$

A high precision indicates a small number of false positives, meaning that the model has a low tendency to classify negative instances as positive.

(ii) *Recall:* Also known as sensitivity or the true positive rate, it measures the proportion of actual positive instances that are correctly identified by the model. It is defined as the ratio of true positives to the sum of true positives and false negatives.

$$\text{Recall}=\frac{\text{No. of true positives}}{\text{No. of true positives + No. of false negatives}}$$

A high recall indicates a small number of false negatives, meaning that the model effectively captures most of the positive instances.

(iii) *F1-score:* The F1-score is the harmonic mean of precision and recall, providing a balanced measure of both metrics. It can be computed using the following formula:

$$F1-score=\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$$

The F1-score ranges between 0 and 1, where a value of 1 represents a perfect balance between precision and recall.

DOI: 10.6977/IJoSI.202506_9(3).0006
R.A. Awashreh/Int. J. Systematic Innovation, 9(3), 96-108 (2025)

Enhancing healthcare efficiency with artificial intelligence: Benefits, challenges, and the future of clinical practice

Raed A. Awashreh*

The Department of Government and Society, UAEU University, AL Ain, United Arab Emirates

*Corresponding author E-mail: raed.raya2020@gmail.com

(Received 06 January 2025; Final version received 21 February 2025; Accepted 10 March 2025)

Abstract

This study explores the integration of artificial intelligence (AI) tools in healthcare, focusing on their impact on cognitive workload, decision-making, and professional development. The findings indicate that AI tools significantly reduce cognitive load, enabling healthcare professionals to focus on higher-order tasks such as critical thinking and complex problem-solving. A majority of participants reported that AI positively influences their professional development, enhancing cognitive functions and empowering them in clinical decision-making. However, concerns were raised about AI's potential negative effects on hands-on clinical skills, particularly in areas such as physical examinations and surgeries, which require manual expertise. These concerns align with the theory of "skill degradation," where over-reliance on AI may hinder the development of essential practical skills. In addition, the study revealed that healthcare workers feared AI could reduce their autonomy in decision-making, emphasizing the need for maintaining human oversight in AI-driven processes. The findings suggest that a balanced approach to AI adoption is essential, where AI complements human expertise rather than replacing it. Training programs should be developed to ensure that healthcare professionals retain core competencies while utilizing AI effectively. Overall, while AI has the potential to improve healthcare delivery by enhancing efficiency and supporting decision-making, its integration must be managed carefully to preserve the essential role of healthcare professionals in providing high-quality care.

Keywords: Artificial Intelligence, Clinical Skills, Cognitive Load, Decision-Making, Healthcare, Professional Development

1. Introduction

The integration of artificial intelligence (AI) tools into healthcare has sparked a transformative shift, driving significant advancements in clinical decision-making, data analysis, and the overall efficiency of medical processes. AI's capacity to process vast amounts of data, detect patterns, and provide predictive insights holds the promise of improving diagnostic accuracy, personalizing treatment plans, and enhancing patient outcomes. In addition, AI tools can support healthcare professionals in high-pressure environments, reducing the burden of routine tasks and allowing them to focus more on complex patient care. However, as the adoption of AI continues to expand within healthcare settings, questions arise regarding its potential consequences on healthcare workers' professional practices, cognitive functions, and skill retention (Bekbolatova et al., 2024; Karalis, 2024).

Despite the numerous benefits AI offers, there is growing concern about its impact on the decision-making autonomy of healthcare professionals and their ability to maintain essential hands-on clinical skills. With the increased reliance on AI systems for diagnosing diseases, recommending treatment plans, and managing patient care, some worry that healthcare workers may gradually lose their ability to think critically and make independent decisions (Adegbesan et al., 2024). Furthermore, the reduction of direct involvement in routine procedures due to automation may lead to skill degradation, particularly in areas requiring manual dexterity or clinical judgment. This issue is especially pertinent given the fast-paced, high-stakes nature of healthcare, where the ability to respond quickly and accurately is crucial (Rashid & Rötting, 2021).

The potential erosion of healthcare professionals' cognitive engagement, alongside the possible decline in manual skills, raises concerns about the future of clinical practice (Lovett et al., 2023). This study aims to address these concerns by exploring the broader implications of AI integration on healthcare practice. Specifically, it focuses on how AI tools might reduce cognitive load, enabling more focused and effective decision-making. At the same time, it considers the risks of over-reliance on technology, including the potential for healthcare professionals to become less engaged in active problem-solving, leading to a gradual erosion of critical thinking and independent judgment.

In examining the role of AI in clinical settings, this research also investigates how these tools influence professional development (Fazakarley et al., 2023). While AI offers opportunities for healthcare workers to acquire new competencies, such as data interpretation and technology management, there is a need to ensure that the human element of care remains intact (Hazarika, 2020). Maintaining professional skills, including communication, empathy, and technical expertise, is fundamental to providing quality healthcare. Therefore, this study aims to strike a balance between leveraging AI's capabilities to augment medical practice and safeguarding the essential skills that healthcare professionals rely on (Pavuluri et al., 2024).

Ultimately, the research seeks to uncover the delicate balance between the benefits of AI tools and the potential risks they pose to healthcare workers' cognitive and professional growth. By understanding how AI affects cognitive load, decision-making autonomy, and skill retention, healthcare institutions can develop strategies that enhance the use of AI without compromising the essential elements of professional practice that is critical to delivering high-quality patient care. The research addresses the following questions:

(i) How does exposure to AI tools influence healthcare professionals' ability to make clinical decisions and retain key clinical skills?
(ii) What are healthcare professionals' perceptions of the impact of AI tools on their professional development and mental capabilities?
(iii) To what extent do healthcare professionals rely on AI tools for critical decisions and data processing?
(iv) What concerns exist regarding the long-term effects of AI tools on healthcare staff's autonomy and practical skills?
(v) Are there significant differences in the perceptions of AI tools based on gender, years of experience, or frequency of exposure?

In addition, the research aims to accomplish the following objectives:

(i) To assess the frequency of exposure to AI tools in healthcare settings and its correlation with professionals' clinical decision-making and cognitive workload.
(ii) To evaluate the impact of AI tools on healthcare staff's ability to process medical data efficiently and their perceived empowerment in clinical decision-making.
(iii) To explore concerns about the reduction of hands-on clinical skills and the potential negative impacts of AI on healthcare practice.
(iv) To identify demographic and experiential factors (e.g., gender, years of experience, exposure frequency) that influence healthcare professionals' perceptions of AI tools in clinical practice.
(v) To contribute to the understanding of AI's role in enhancing or limiting healthcare staff's professional development and the implications for the future of clinical practice.

This research aims to provide insights into the impact of AI on healthcare professionals' cognitive functions, decision-making skills, and professional autonomy. It also seeks to explore how AI can be integrated into healthcare without undermining the essential skills and judgment that are critical to the profession. In light of this, the research problem revolves around the increasing prevalence of AI tools in healthcare and the questions they raise about their impact on healthcare professionals' mental and professional abilities. As AI becomes more widespread, concerns have emerged regarding the potential for healthcare workers to become dependent on these technologies, which could diminish their cognitive involvement in decision-making and erode their hands-on skills. Despite the known benefits of AI, such as reducing human error and improving efficiency, it is crucial to further investigate its potential negative effects on professional autonomy and cognitive sharpness.

To address these concerns, this study aims to fill a critical gap in understanding the effects of AI tools on healthcare professionals. It is essential to examine how AI can complement, rather than replace, the expertise and judgment of medical professionals. By exploring the relationship between AI and healthcare workers' cognitive and professional abilities, the study seeks to inform policies and strategies that ensure AI enhances healthcare practices while safeguarding the skills and autonomy of medical staff. Therefore, this research not only contributes to understanding AI's role in healthcare but also helps guide its responsible implementation in clinical settings.

2. Theoretical Framework and Literature Review

The integration of AI into healthcare is reshaping the profession by offering potential advancements in diagnostics, decision support, and automation of routine tasks. However, its increasing prevalence raises significant questions about its impact on the cognitive and professional development of healthcare workers, particularly doctors and nurses (Khan Rony et al., 2024). While AI tools are praised for enhancing efficiency and improving clinical outcomes, concerns exist regarding their potential influence on cognitive load, professional autonomy, and skill retention. This literature review aims to examine the theoretical frameworks that address these complex dynamics between AI applications and healthcare professionals' abilities (Ahmad et al., 2023).

Cognitive load theory (CLT) is central to understanding the mental effort involved when healthcare workers engage with AI tools. AI systems, designed to automate tasks such as diagnostic imaging and data analysis, have the potential to reduce cognitive load by enabling professionals to focus on more complex clinical decisions (Fox & Rey, 2024). However, over-reliance on AI could result in diminished cognitive engagement, leading to a decline in critical thinking and problem-solving skills. This interplay between cognitive load and AI's influence on mental effort is crucial for understanding its impact on healthcare professionals' cognitive functions (Zhai et al., 2024). In the context of the study results, many healthcare professionals reported feeling more confident in their decision-making, suggesting that AI tools can indeed alleviate cognitive load. However, the long-term consequences of reduced engagement with critical thinking warrant further exploration (Çela et al., 2024).

The theory of professional autonomy provides insight into the potential impact of AI tools on healthcare workers' independence and judgment (Salvatore et al., 2018; Mrayyan et al., 2024). While AI systems offer efficiency and accuracy, they may inadvertently undermine professionals' sense of autonomy. By offering tailored recommendations and solutions, AI tools could reduce the space for independent decision-making, thus diminishing healthcare workers' confidence in their judgment. The results of this study echo this concern, as some participants expressed unease about the potential erosion of autonomy due to AI's growing influence on clinical decision-making (Zhai et al., 2024).

The integration of AI into healthcare influences healthcare workers' professional practices, with theories like human–technology interaction exploring how AI tools shape workflows and decision-making (Masudin et al., 2024). Healthcare workers show mixed reactions, appreciating AI's assistance with data while fearing the loss of hands-on skills (Fritsch et al., 2024). Skill acquisition theory suggests that while AI automates routine tasks, it also creates opportunities for developing new competencies in technology management and data interpretation (Taie, 2014; Mikalef et al., 2023).

The technology acceptance model and self-determination theory assess healthcare workers' engagement with AI, emphasizing factors such as ease of use, usefulness, and intrinsic motivation (Tao et al., 2023; Williamson & Prybutok, 2024). While AI can streamline administrative tasks, it must preserve professional autonomy and decision-making capabilities. AI's impact on decision-making and patient outcomes is mostly positive, but concerns about reduced cognitive engagement and skill erosion persist (Khalifa & Albadawy, 2024; Bekbolatova et al., 2024).

In Oman, AI integration in healthcare and medical education is in its early stages but shows great potential. Al Hadithy et al. (2023) found that although medical students supported AI in curricula, most had limited exposure and expressed concerns about ethics and employment impacts. Al Riyami (2024) highlighted barriers such as high costs and infrastructure gaps in Omani institutions, although AI tools such as DataRobot improved research efficiency. Varnosfaderani and Forouzanfar (2024) emphasized AI's role in clinical decision-making, medical imaging, and patient monitoring while addressing ethical issues like data privacy.

Ariffin et al. (2024) identified key areas for sustainable healthcare management in Oman, such as waste management and public health initiatives, which align with the country's Health Vision 2050. In parallel, El-Khoury & Albarashdi (2024) examined data protection laws in Oman and its neighboring countries, highlighting the strong privacy frameworks in place and the growing need for anonymization technologies. The successful integration of AI into Oman's healthcare system will rely on strengthening curricula, investing in infrastructure, and fostering interdisciplinary collaboration. As AI continues to advance healthcare by enhancing efficiency, it is crucial to address its impact on cognitive load, professional autonomy, and skill development (Pavuluri et al., 2024). Table 1 illustrates the relationship between AI applications and healthcare professionals' cognitive and professional abilities. The integration of theoretical frameworks helps provide a comprehensive understanding of how AI influences cognitive and professional capabilities, underscoring the importance of maintaining a balance between preserving skills and autonomy while embracing technological progress (Mashabab et al., 2024).

Table 1. Relationship between artificial intelligence (AI) applications and healthcare professionals' cognitive and professional abilities

Independent variable (AI applications/tools)	Dependent variable (cognitive and professional abilities)
AI tools for diagnostics: The use of AI applications in diagnostic processes, such as analyzing medical images, interpreting laboratory results, or predicting patient outcomes, which can assist or replace traditional diagnostic methods	Cognitive function and critical thinking: The ability of healthcare professionals (doctors and nurses) to engage in critical thinking, problem-solving, and independent judgment, particularly in complex or emergency situations. The potential reduction in cognitive engagement due to over-reliance on AI tools is a concern
AI-powered decision support systems: Systems that provide recommendations or support for decision-making in clinical settings, such as treatment plans, medication suggestions, or patient management strategies, based on AI algorithms	Professional autonomy: The extent to which healthcare professionals feel empowered to make decisions independently, without heavy reliance on AI tools. This includes confidence in decision-making and the freedom to practice autonomously. The integration of AI tools may reduce autonomy by providing pre-determined solutions, leading to concerns about diminished independence
Automation of routine tasks: The use of AI for automating repetitive tasks such as patient record management, scheduling, or routine medical procedures, which reduces the cognitive load on healthcare professionals	Skill retention and development: The maintenance and enhancement of professional skills, such as medical knowledge, manual dexterity, and patient interaction skills. While AI can reduce cognitive load, it may also lead to skill degradation if healthcare professionals become overly reliant on automation for routine tasks. In addition, AI could foster new competencies in technology management and data interpretation

Source: Developed by the author.

3. Methodology

3.1. Research Design

The study employed a descriptive-analytical methodology with a quantitative approach, utilizing a survey-based strategy to gather primary data from healthcare professionals, specifically doctors and nurses (Smith & Hasan, 2020). The survey incorporated closed-ended questions, Likert-scale items, and open-ended questions to capture diverse perspectives. The questionnaire, developed in consultation with experts in healthcare, AI, and research methodology, was divided into four sections: demographic information, AI exposure frequency, clinical decision-making impact, and professional development (Baburajan et al., 2022).

To investigate the effects of AI tools on clinical decision-making, skills retention, and professional development (Khosravi et al., 2024), the survey presented participants with a series of statements about their experiences with AI tools. Respondents rated their level of agreement on a five-point Likert scale, ranging from "Strongly disagree" to "Strongly agree," allowing for a comprehensive assessment of their interactions with AI in clinical contexts.

The survey was distributed through Google Forms and shared through the Ministry of Health networks on WhatsApp. Participation was voluntary, with no personal information requested. Participants provided informed consent and could withdraw at any time.

3.2. Sample Population

Given the unknown population in Oman, the district of North A'Sharqiyah and Muscat, the researchers opted to distribute the survey electronically. The sample comprised healthcare professionals with varying levels of experience and exposure to AI tools in clinical settings, including doctors, nurses, and medical technicians. Demographic data such as gender, years of experience, and the frequency of AI exposure were collected to facilitate a deeper analysis of how these factors influence perceptions of AI's role in healthcare. Among the 164 respondents, 27% were male, and 73% were female. All participants worked in the governmental healthcare sector. Regarding professional experience, 13% had 1–5 years of experience, 13.6% had 6–10 years, and the majority (63%) had over 10 years. The sample distribution by profession included 32.1% doctors, 58% nurses, and 9.9% medical technicians.

3.3. Data Analysis

Data were analyzed using both descriptive and inferential statistics. Descriptive statistics provided a summary of the frequency of responses for each question, such as the percentage of respondents in each category for the Likert scale items. Inferential statistics, including independent samples *t*-tests and Levene's test for homogeneity of variances, were employed to assess differences in responses based on factors such as sex, years of experience, and frequency of exposure to AI tools. The independent samples *t*-tests were used to compare perceptions of AI tools between different groups, identifying significant differences in areas such as confidence in clinical decision-making, AI's impact on clinical skills, and the role of AI in professional development. Levene's test for homogeneity of variances was conducted to check for equal variances across groups. A significant result ($p \leq 0.05$) indicated unequal variances, suggesting that different groups had inconsistent views on specific aspects of AI in healthcare.

3.4. Voluntary Participation

Participation in the study was voluntary, with the option to withdraw at any time without penalty. Participants were informed about the study's nature, the anonymity of their responses, and their right to withdraw. The research aims to explore the impact of AI tools on healthcare professionals' cognitive functions, decision-making, professional autonomy, and skill retention. Using a literature review and survey methodology, the study examined the relationship between AI tool exposure and professional development outcomes in healthcare.

4. Results

This section presents the results through findings from descriptive analysis and several major statistical tests. Table 2 demonstrates that the frequency of exposure to AI tools among healthcare staff varies, with a significant portion using them daily (49.4%). Smaller percentages reported using AI tools weekly (14.8%) or rarely (18.5%), while 9.9% had never used them. This variation in exposure reflects differing levels of engagement with AI in clinical practice. In terms of impact, most respondents agreed that AI tools reduce cognitive load in decision-making (59.2% agree or strongly agree), and 59.3% believed that AI enhances their ability to think critically and solve problems in complex situations. In addition, a substantial number of healthcare workers (58.6%) reported that they or their colleagues rely on AI tools for important clinical decisions.

Table 2. Frequency of exposure to artificial intelligence (AI) tools

Exposure to AI tools		**AI tools reduce the cognitive load associated with decision-making in clinical practice**	
Type	**Percentage**	**Type**	**Percentage**
Never	9.9	Strongly disagree	8.6
Rarely	18.5	Disagree	8.6
Daily	49.4	Neutral	23.5
Weekly	14.8	Agree	37.0
Monthly	7.4	Strongly agree	22.2
Total	100	Total	100.0
AI tools enhance healthcare staff's ability to think critically and solve problems in complex clinical situations		**My colleagues rely on AI tools to make important clinical decisions**	
Type	**Percentage**	**Type**	**Percentage**
Strongly disagree	7.4	Strongly disagree	4.9
Disagree	3.7	Disagree	13.6
Neutral	16.0	Neutral	22.2
Agree	54.3	Agree	45.7
Strongly agree	18.5	Strongly agree	13.6
Total	100.0	Total	100.0

Source: Developed by the author.

AI tools were perceived as valuable in helping healthcare staff process and interpret medical data more efficiently. Table 3 shows a significant portion of respondents (67.9%) either agreed or strongly agreed that AI tools enhance their ability to analyze medical data. However, 8.7% strongly disagreed or disagreed with this statement, indicating that some healthcare workers may not fully experience the advantages of AI in data processing. When it came to clinical decision-making, 19.8% of respondents agreed that AI empowers them to make better decisions, while a notable 45.7% remained neutral, suggesting that many healthcare professionals may be unsure of AI's impact on their decision-making.

Regarding clinical skills retention, 61.8% of respondents agreed or strongly agreed that AI tools support the retention of key clinical skills, such as diagnostic ability and patient management. However, 14.8% of healthcare workers felt neutral or disagreed with this statement, reflecting some skepticism about AI's role in preserving essential skills. On the other hand, 49.2% of healthcare workers believed that AI tools reduce hands-on clinical skills, such as performing physical examinations and surgeries. While 16.0% strongly agreed, 51.3% remained neutral or disagreed, highlighting the complexity of balancing AI use with maintaining practical clinical capabilities.

AI tools were noted as regarded as helping healthcare staff process and interpret medical data more efficiently. As shown in Table 3, a significant portion of respondents (67.9%) either agreed or strongly agreed that AI enhances their ability to analyze medical data. However, 8.7% strongly disagreed or disagreed with this statement, indicating that some healthcare workers may not fully experience the advantages of AI in data processing. When it came to clinical decision-making, 19.8% of respondents agreed that AI empowers them to make better decisions, while a notable 45.7% remained neutral, suggesting that many healthcare

Table 3. Perceived impact of artificial intelligence (AI) tools on clinical decision-making

AI tools help healthcare staff process and interpret medical data more efficiently		**I feel that AI tools empower me to make better clinical decisions**	
Type	**Percentage**	**Type**	**Percentage**
Strongly disagree	6.2	Strongly disagree	7.4
Disagree	2.5	Disagree	27.2
Neutral	23.5	Neutral	45.7
Agree	45.7	Agree	19.8
Strongly agree	22.2	Strongly agree	100.0
Total	100.0	Total	7.4
The use of AI tools supports the retention of key clinical skills, such as diagnostic ability and patient management		**The use of AI tools reduces healthcare staff's hands-on clinical skills (e.g., performing physical examinations and surgeries)**	
Type	**Percentage**	**Type**	**Percentage**
Strongly disagree	6.2	Strongly disagree	7.4
Disagree	2.5	Disagree	27.2
Neutral	23.5	Neutral	45.7
Agree	45.7	Agree	19.8
Strongly agree	22.2	Strongly agree	100.0
Total	100.0	Total	7.4

Source: Developed by the author.

professionals may be unsure of AI's impact on their decision-making.

Table 4 presents healthcare professionals' varied perspectives on the impact of AI tools in clinical practice. A significant portion of respondents (37%) agreed that AI tools reduce the need for traditional manual practices in healthcare, with 14.8% strongly agreed. However, concerns about the potential negative impact of AI on the development of hands-on skills remained, as 29.6% agreed and 22.2% strongly agreed that AI could hinder the development of skills related to patient care. Despite this, the majority of respondents (49.4%) believed that AI positively influences their professional development and cognitive capabilities, with only 3.7% strongly disagreed with this statement. These findings suggest that while AI was viewed as a helpful tool in improving efficiency, there are ongoing concerns regarding its long-term effects on skill retention and independent decision-making.

Table 5 presents responses regarding concerns about the long-term effects of AI tools on healthcare

Table 4. Perceived impact of artificial intelligence (AI) tools on medical practices

The use of AI tools results in a reduction of the need for traditional or manual medical practices in healthcare work		**I believe that AI tools can negatively impact the development of new skills, particularly those related to hands-on patient care**	
Type	**Percentage**	**Type**	**Percentage**
Strongly disagree	8.6	Strongly disagree	4.9
Disagree	12.3	Disagree	11.1
Neutral	27.2	Neutral	32.1
Agree	37.0	Agree	29.6
Strongly agree	14.8	Strongly agree	22.2
Total	100.0	Total	100
I have concerns about the long-term effects of using AI tools on healthcare staff's ability to make independent decisions and retain professional skills		**AI positively impacts my overall professional development and mental capabilities**	
Type	**Percentage**	**Type**	**Percentage**
Strongly disagree	3.7	Strongly disagree	3.7
Disagree	12.3	Disagree	12.3
Neutral	21.0	Neutral	21.0
Agree	49.4	Agree	49.4
Strongly agree	13.6	Strongly agree	13.6
Total	100.0	Total	100.0

Sources: Developed by author.

Table 5. Concerns and impact of artificial intelligence (AI) tools on healthcare practices

I have concerns about the long-term effects of using AI tools on healthcare staff's ability to make independent decisions and retain professional skill		**Use of AI tools negatively impacts the professional practices of healthcare**	
Type	**Percentage**	**Type**	**Percentage**
Disagree	0	Disagree	3.7
Strongly disagree	9.9	Strongly disagree	23.5
Neutral	25.9	Neutral	32.1
Agree	44.4	Agree	30.9
Strongly agree	19.8	Strongly agree	9.9
Total	100.0	Total	100.0

Source: Developed by the author.

staff's ability to make independent decisions and retain professional skills, as well as the impact of AI on professional practices. A majority of respondents (44.4%) agreed that they have concerns about the long-term effects of AI tools on healthcare staff's independent decision-making and skill retention, while 25.9% were neutral. On the other hand, when asked whether the use of AI tools negatively impacts professional practice, most respondents disagreed (23.5%), while a significant portion remained neutral (32.1%). These findings suggest that while healthcare professionals were concerned about the potential long-term effects of AI on decision-making and skills retention, the general view did not indicate a strong negative impact on professional practices.

For the statistical analysis, the results of significant differences are outlined as follows: when considering sex (male and female), the use of AI tools was associated with a reduction in the need for traditional or manual medical practices in healthcare work. The analysis revealed $F = 7.604$, $p=0.007$, indicating a significant difference between the groups. This suggests that opinions on this topic varied significantly across groups, requiring further investigation. In addition, the results showed significant differences ($p\leq0.05$) in several areas when examining years of experience in relation to AI tools in healthcare. A key finding is that AI tools helped healthcare staff process and interpret medical data more efficiently, with a *p*-value of 0.000, demonstrating a clear advantage in data handling. Furthermore, healthcare workers reported feeling more confident in making clinical decisions when AI tools were used, with a significant *p*-value of 0.001, underscoring the positive impact of these tools on decision-making confidence.

Moreover, AI tools provided recommendations that influenced clinical decision-making, as evidenced by a *p*-value of 0.000. This suggests that AI plays a crucial role in shaping decisions made by healthcare professionals. In addition, AI tools empowered healthcare workers to make better clinical decisions, further supported by a *p*-value of 0.000. This empowerment is particularly valuable in complex clinical situations.

Interestingly, the use of AI tools was also linked to the retention of key clinical skills, such as diagnostic ability and patient management ($p=0.000$), indicating that these tools can complement traditional skills. However, concerns were raised about AI's potential to reduce healthcare staff's hands-on clinical skills, such as performing physical examinations and surgeries, with a *p*-value of 0.001 suggesting a negative impact on practical skills. Furthermore, AI tools contributed to a reduction in the need for traditional or manual medical practices ($p=0.000$), reflecting the shift toward more automated approaches. However, some workers believed that AI tools could negatively affect the development of new skills, particularly those related to hands-on patient care, as shown by a *p*-value of 0.000.

Regarding professional development, AI tools had a positive impact on overall development and mental capabilities ($p=0.000$). However, concerns about the long-term effects of AI on the ability to make independent decisions and retain professional skills were noted ($p=0.004$). Some participants also expressed the belief that AI tools might negatively affect professional practice in healthcare ($p=0.000$), emphasizing the need for careful consideration of their role. Finally, AI tools enhanced healthcare staff's ability to think critically and solve problems in complex clinical situations, as indicated by a *p*-value of 0.000, highlighting the cognitive benefits of these technologies.

The results highlight significant differences ($p\leq0.05$) in several aspects of healthcare roles in relation to the use of AI tools. One area of importance is confidence in clinical decisions. Healthcare workers reported greater confidence in their decision-making abilities when AI tools were used, as indicated by a *p*-value of 0.003. This confidence is further supported by the fact that AI tools provide recommendations that influence clinical decisions ($p=0.028$), suggesting that these tools play an active role in shaping the decision-making process.

AI tools also appeared to empower healthcare workers by helping them make better clinical decisions, with a *p*-value of 0.011 reinforcing this positive impact. However, despite these benefits, concerns were raised that the use of AI tools could reduce healthcare workers' hands-on clinical skills, such as performing physical examinations or surgeries ($p=0.008$). This reduction in practical skills could be a potential drawback of relying on AI in healthcare settings. On the other hand, AI tools were also seen as contributing positively to professional development and enhancing mental capabilities ($p=0.002$), suggesting that they may support growth in other important aspects of healthcare roles. Yet, there are concerns about the long-term effects of AI use, particularly regarding its impact on healthcare workers' ability to make independent decisions and retain essential skills ($p=0.007$). This underscores the need for a balanced approach to integrating AI into healthcare to ensure that professionals maintain their core competencies while benefiting from technological advancements.

The results showed significant differences ($p\leq0.05$) in several areas related to the frequency of exposure to AI tools in healthcare. One key finding was that the use of AI tools reduced the need for traditional or manual medical practices in healthcare work, with a *p*-value of 0.001, indicating that AI is increasingly replacing more conventional methods. This shift, however, raised concerns, as some participants believed that AI tools could negatively impact the development of new skills, particularly those related to hands-on patient care (p=0.000), suggesting that over-reliance on AI might hinder the growth of practical clinical skills.

On a more positive note, AI was perceived to have a positive impact on overall professional development and mental capabilities (p=0.001), pointing to the benefits of AI in enhancing cognitive functions and professional growth. Despite these advantages, concerns were raised about the long-term effects of AI tools on healthcare staff's ability to make independent decisions and retain professional skills, as reflected by a p-value of 0.016, suggesting a potential risk of dependency on AI over time. Further concerns were expressed regarding the negative impact of AI tools on the professional practice of healthcare (p=0.000), indicating that some professionals fear AI could undermine the quality of care or professional standards. In addition, AI tools were seen to reduce the cognitive load associated with decision-making in clinical practice (p=0.000), making decision-making more efficient but possibly at the cost of diminishing critical thinking skills. Finally, there was a significant concern that AI tools could limit healthcare staff's ability to make independent decisions in clinical practice (p=0.000), suggesting that over-exposure to AI may reduce healthcare workers' autonomy in clinical settings.

The independent samples t-test results demonstrated significant differences in perceptions of AI tools in clinical practice. There was a noticeable variation in how much colleagues rely on AI tools for critical decisions (p=0.003), with one group reporting less reliance. Similarly, one group viewed AI tools as less efficient in processing medical data (p=0.010), reflecting a more negative perspective on their usefulness. Confidence in decision-making also varied significantly (p=0.002), with one group feeling less confident when using AI tools. The influence of AI on clinical decisions differed between groups, with one group perceiving a stronger impact ($p < 0.001$). A similar trend was observed in the perceived empowerment from AI tools in making better decisions (p=0.002), with one group feeling less empowered. There was also disagreement regarding whether AI tools helped retain critical clinical skills (p=0.002), with one group expressing more skepticism. AI's effect on reducing hands-on skills, such as physical examinations and surgeries, was perceived more negatively by one group (p=0.016). Finally, concerns about AI's long-term effects on decision-making and skill retention were more pronounced in one group ($p = 0.003$). These findings suggest varying experiences and perceptions of AI's role in clinical practice, indicating areas that may require further exploration to address concerns and clarify AI's impact on healthcare.

The test of homogeneity of variances (Levene's test) assessed whether different groups had similar variances in their responses. A significant result ($p \leq 0.05$) indicates unequal variances, challenging the assumption of homogeneity required for tests like analysis of variance. Several significant differences were found in perceptions of AI tools in clinical practice. For example, AI tools' ability to help healthcare staff process medical data efficiently showed significant variation (Levene statistic = 6.172, p=0.000), indicating differing levels of consistency in responses across groups. Similarly, the perception of AI tools limiting healthcare staff's independence in decision-making also had significant variability (Levene statistic = 10.053, p=0.000), suggesting inconsistency across groups. In addition, perceptions of empowerment through AI tools (Levene statistic = 3.046, p=0.022) and the retention of key clinical skills (Levene statistic = 2.498, p=0.050) revealed significant variance, showing inconsistent views on AI's role in these areas. The impact of AI tools on hands-on clinical skills (Levene statistic = 13.665, p=0.000) and the need for traditional medical practices (Levene statistic = 6.118, p=0.000) also demonstrated substantial differences across groups. Furthermore, concerns about AI's impact on skill development and professional practices (Levene statistics = 6.549 and 6.870, respectively; both p=0.000) highlighted variability in opinions. Finally, the perception that AI tools reduce cognitive load in decision-making (Levene statistic = 5.289, p=0.001) showed differing views among groups. These findings suggest significant differences in how groups perceive the impact of AI tools in clinical practice, highlighting varied experiences and opinions that warrant further investigation.

5. Discussion

This study explores the increasing integration of AI tools into clinical practices, affecting healthcare professionals' cognitive workload, decision-making, and professional development. The results are consistent with existing literature that emphasizes AI's potential to enhance healthcare by reducing cognitive load and increasing efficiency (Buntinx et al., 2020; Jha et al., 2020). However, concerns have emerged regarding AI's impact on hands-on clinical skills and independent decision-making, which are crucial for healthcare professionals' long-term autonomy.

A key finding is that 59.2% of healthcare professionals believe AI reduces cognitive load. This aligns with studies indicating that AI's ability to process large datasets enables clinicians to focus on higher-level tasks (Amir et al., 2022). Furthermore, 59.3% of respondents reported enhanced critical thinking and problem-solving, supporting the CLT, which asserts that reducing extraneous cognitive load through technology improves decision-making (Sweller, 2011). This positive perception reinforces the idea that AI complements human expertise in healthcare rather than replacing it.

However, concerns about AI undermining hands-on clinical skills, particularly in areas such as physical examinations and surgeries (49.2%), were raised. This is consistent with the concept of "skill degradation," where over-reliance on technology may hinder the development of manual and diagnostic skills (Mann & Stokes, 2019). Literature on automation advocates for a hybrid model, where AI assists rather than replaces human capabilities, ensuring sustainable healthcare development (Davenport & Kalakota, 2019).

The study also reveals significant variation in responses based on demographic factors such as years of experience and gender. More experienced healthcare professionals tend to view AI as a supportive tool, while those newer to the field are more dependent on AI for guidance (Brock et al., 2021). This highlights the need for tailored AI training and implementation strategies to address different levels of expertise and familiarity with technology. In addition, 44.4% of respondents expressed concerns about AI limiting independent decision-making, fearing excessive reliance on AI. This is supported by the dependency on automation theory, which suggests that increasing AI integration may reduce professionals' ability to make decisions without technological assistance (Parasuraman & Riley, 1997). While AI can enhance decision-making, it is essential that healthcare professionals retain the cognitive and practical skills needed to navigate complex clinical environments.

AI in healthcare has garnered significant attention, with studies exploring both its advantages and drawbacks. A prominent theme in the literature is the enhancement of cognitive functions, such as decision-making and data interpretation. AI tools assist healthcare professionals by efficiently analyzing large datasets, thus reducing the cognitive load (Topol, 2019). This study's findings align with these themes, as respondents reported improvements in cognitive functions and critical thinking. However, balancing technology with human skills remains a challenge. Researchers such as Brynjolfsson and McAfee (2014) emphasize AI's dual nature – improving efficiency while potentially undermining traditional skills. The study's finding that AI could reduce hands-on clinical skills echoes these concerns, especially in fields like surgery that requires manual dexterity. Hybrid models combining AI with human expertise are recommended to preserve essential clinical skills (Davenport & Kalakota, 2019).

The study highlights the integration of AI tools into clinical practices, and specific examples of these tools can clarify their practical relevance. AI tools in healthcare, such as clinical decision support systems, machine learning algorithms for diagnosing medical conditions, and advanced systems like robot-assisted surgery and AI-driven data analytic platforms, enable healthcare professionals to analyze large datasets, offer evidence-based recommendations, and assist in surgeries. Some examples include:

(i) Clinical decision support systems: These tools assist in diagnosing medical conditions by analyzing patient data and offering recommendations. For example, IBM Watson Health helps diagnose cancer by analyzing extensive medical literature and patient records.
(ii) AI-driven data analytics: Platforms like Google Health's AI algorithms for interpreting medical imaging, such as radiology scans, enhance diagnostic accuracy by identifying patterns that are invisible to the human eye. This study suggests that AI tools like these significantly improve healthcare workers' ability to interpret complex data, leading to more accurate diagnoses and treatment plans.
(iii) Robot-assisted surgery: Systems such as Da Vinci Surgical Systems provide enhanced precision in complex procedures. While AI supports the technical aspects of surgery, it does not replace the surgeon's expertise but rather augments it to improve patient outcomes.

Including these examples allows the study to better illustrate the practical applications of AI tools in healthcare settings, demonstrating their tangible impact on cognitive workload and clinical outcomes.

In terms of professional development, AI is viewed as enhancing healthcare workers' cognitive capabilities and empowering them to solve complex problems. The theory of professional identity development suggests that AI support can bolster healthcare workers' professional confidence (Simmons et al., 2021). This study found that most healthcare professionals believe AI positively influences their professional development, supporting this theory.

However, concerns about AI's long-term effects on decision-making autonomy are well-documented. The concept of "human-in-the-loop" decision-making stresses the importance of maintaining human oversight in AI-driven processes (Gunkel, 2018). The findings of this study reflect these concerns, as participants believe that AI could restrict their ability to make independent clinical decisions.

In summary, the integration of AI in healthcare presents both opportunities and challenges. While the study emphasizes AI's role in reducing cognitive load and enhancing professional development, concerns regarding the erosion of hands-on clinical skills and decision-making autonomy persist. These findings highlight the need for a balanced approach to AI adoption, ensuring it complements human expertise. Furthermore, future research should explore the long-term effects of AI on skill retention, decision-making autonomy, and the development of training programs to help healthcare professionals navigate technological advancements.

As AI tools continue to be integrated into healthcare, several ethical considerations must be addressed to ensure their responsible use, particularly regarding data privacy and equitable access. AI tools often require access to large datasets, including patient histories, test results, and imaging, raising significant privacy concerns. Maintaining the confidentiality of patient data is crucial, and AI tools such as IBM Watson Health and Google's diagnostic systems must comply with healthcare regulations such as HIPAA (Health Insurance Portability and Accountability Act) in the United States and GDPR (General Data Protection Regulation) in Europe to mitigate these risks. Healthcare institutions must prioritize data encryption, anonymization, and secure cloud storage, along with compliance with data protection laws. Transparency about how patient data is collected and used for AI purposes will foster trust between healthcare professionals, patients, and AI developers.

While AI has the potential to significantly improve healthcare delivery, its benefits may not be universally accessible. Factors such as geography, economic disparities, and healthcare infrastructure can limit access to AI-powered solutions. In low-resource areas, such as rural or underserved regions, AI tools may be unavailable due to high costs or lack of infrastructure, which could exacerbate health inequities. To ensure equitable access, global healthcare organizations and policymakers should advocate for affordable, open-source AI solutions that can be implemented across diverse healthcare environments. Training local healthcare workers to use these technologies effectively is also crucial. Collaboration between governments, NGOs, and AI developers can help create scalable solutions to ensure that AI benefits are accessible to all, regardless of income or location.

6. Conclusion

This study provides valuable insights into the growing role of AI tools in clinical practices, highlighting their positive impact on healthcare professionals' cognitive workload, decision-making, and professional development. The results indicate that AI significantly reduces cognitive load, allowing healthcare workers to handle complex tasks more efficiently and focus on higher-order decision-making and critical thinking. These benefits align with existing research suggesting that AI tools streamline processes, improve clinical outcomes, and enhance healthcare delivery by processing large datasets in real time.

However, concerns about AI's potential negative effects, particularly on hands-on clinical skills and independent decision-making, have emerged. As AI becomes more integrated into healthcare workflows, there is a risk of diminishing manual and diagnostic capabilities, which are crucial for long-term professional effectiveness and patient care. This concern is particularly evident in fields such as surgery and physical examinations, where direct patient interaction is essential.

The study underscores the importance of adopting AI in a manner that complements human expertise rather than replacing it. A balanced integration of AI should enhance healthcare professionals' roles while preserving core competencies for independent decision-making and direct patient care. Achieving this requires designing AI systems and training programs that promote a collaborative relationship between humans and machines, ensuring that healthcare workers can maximize the benefits of AI without sacrificing their autonomy or clinical competence.

Furthermore, the findings align with theoretical frameworks like CLT and professional identity development, demonstrating AI's potential to enhance cognitive functions and decision-making abilities. However, the study also raises concerns about skill degradation, echoing the theory of dependency on automation. The integration of AI into healthcare has profound societal implications. While AI can improve healthcare delivery by reducing cognitive load and supporting decision-making, it is essential to ensure that it does not compromise hands-on care. A balanced approach will help advance healthcare efficiency while preserving essential human skills, ultimately benefiting both patients and society.

For healthcare professionals, AI offers valuable support in decision-making, data analysis, and critical thinking, potentially enhancing professional development. However, there is a risk of over-reliance on AI, which could undermine autonomy and reduce hands-on clinical skills. Training programs should therefore focus on integrating AI while preserving core competencies, ensuring that professionals remain capable of independent decision-making and maintaining practical skills. From an organizational standpoint, healthcare institutions must implement AI in ways that complement human expertise, offering tailored training for staff based on their experience and fostering collaboration between AI and human practitioners. In addition, AI integration should not compromise decision-making autonomy or clinical skills.

This study's limitations include its reliance on self-reported data, which may be biased, and its focus on a specific set of AI tools and healthcare professionals, limiting generalizability. Future research should examine AI's long-term effects on skill retention and professional development in various healthcare settings. In addition, studies on AI's impact on patient outcomes and the development of comprehensive AI training programs are needed for effective integration into clinical practice.

The following are several recommendations for AI training programs:

(i) Hybrid training model: AI training should complement traditional medical education. Healthcare professionals should be trained to use AI tools effectively while maintaining clinical judgment, promoting decision-making that integrates AI recommendations without over-relying on technology.
(ii) Scenario-based training: Incorporate AI into scenario-based training programs that simulate real clinical situations. This will help professionals integrate AI with traditional practices and build confidence in making independent decisions when AI may be less reliable.
(iii) Regular skills refresher courses: Healthcare professionals should participate in refresher courses to maintain core competencies, such as physical examinations and diagnostics. This ensures that AI complements, rather than replaces, essential hands-on skills.
(iv) Focus on ethical use of AI: Training should cover the ethical use of AI, including its limitations, human oversight, and potential biases. Healthcare workers must learn to critically assess AI recommendations and understand its implications across patient populations.
(v) Differentiated training based on experience: Training should be tailored to the professional's experience level. For instance, senior doctors may focus on advanced AI applications, while newcomers may learn foundational concepts and the daily use of AI tools.
(vi) Continuous feedback and evaluation: AI training should be ongoing, with regular feedback on professionals' use of AI in clinical settings. Performance evaluations should track the impact of AI on practice and identify areas for improvement.

By implementing these recommendations, healthcare organizations can ensure AI supports professionals without undermining their clinical skills or autonomy, ultimately enhancing healthcare delivery and professional development.

Acknowledgment

The author would like to extend his gratitude to MBA student Mosa Alhusaini for facilitating the distribution of the survey.

References

Adegbesan, A., Akingbola, A., Aremu, O., Adewole, O., Amamdikwa, J.C., & Shagaya, U. (2024). From scalpels to algorithms: The risk of dependence on artificial intelligence in surgery. *Journal of Medicine, Surgery, and Public Health*, 3, 100140. https://doi.org/10.1016/j.glmedi.2024.100140

Ahmad, S.F., Han, H., Alam, M.M., Rehmat, M.K., Irshad, M., Arraño-Muñoz, M., et al. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanities and Social Sciences Communications*, 10, 311. https://doi.org/10.1057/s41599-023-01787-8

Al Hadithy, Z.A., Al Lawati, A., Al-Zadjali, R., & Al Sinawi, H. (2023). Knowledge, attitudes, and perceptions of artificial intelligence in healthcare among medical students at Sultan Qaboos University. *Cureus*, 15(9), e44887. https://doi.org/10.7759/cureus.44887

Al Riyami, N. (2024). *Advancing AI in Omani Medical Research: Progress, Challenges, and Ethics* [Preprint]. https://doi.org/10.2196/preprints.66991

Amir, M., Khan, M.R., & Raza, S. (2022). Maternal nutrition education as a tool to optimize the infant feeding practices: A narrative review. *International Medical and Health Sciences,* 2(1), 36–42. https://doi.org/36-42-02(1)2022MHS22-136

Ariffin, N.H.M., Yusoff, F.H., Mansor, N.N.A., Javed, S., & Nazir, S.N. (2024). Assessment of sustainable healthcare management policies and practices in Oman: A comprehensive analysis using an ai-based literature review. *Multidisciplinary Reviews*, 7(11), 2024257. https://doi.org/10.31893/multirev.2024257

Baburajan, V., De Abreu e Silva, J., & Pereira, F.C. (2022). *Open vs Closed-Ended Questions in Attitudinal Surveys: Comparing, Combining, and Interpreting using Natural Language Processing*. [arXiv Preprint]. https://doi.org/10.48550/arXiv.2205.01317

Bekbolatova, M., Mayer, J., Ong, C.W., & Toma, M. (2024). Transformative potential of AI in healthcare: Definitions, applications, and navigating the ethical landscape and public perspectives. *Healthcare (Basel)*, 12(2), 125. https://doi.org/10.3390/healthcare12020125

Brock, A., De, S., Smith, S.L., & Simonyan, K. (2021). High-Performance Large-Scale image Recognition without Normalization. In: *International Conference on Machine Learning*. PMLR, p. 1059–1071.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. WW Norton and company, United States.

Buntinx, F., Claes, P., Gulikers, M., Verbakel, J., Jan, D.L., Van der Elst, M., Van Elslande, J., Van Ranst, M., & Vermeersch, P. (2020). Added value of anti-SARS-CoV-2 antibody testing in a Flemish nursing home during an acute COVID-19 outbreak in April 2020. *Acta Clinica Belgica,* 77(2), 295–300. https://doi.org/10.1080/17843286.2020.1834285

Çela, E., Fonkam, M., & Potluri, R.M. (2024). Risks of AI-assisted learning on student critical thinking. *International Journal of Risk and Contingency Management,* 12(1), 1–19. https://doi.org/10.4018/IJRCM.350185

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2), 94–98.

El-Khoury, M., & Albarashdi, S. (2024). Navigating the privacy landscape of healthcare-driven ai in the middle east: Case studies from Oman, Qatar, and Saudi Arabia. *Social Sciences and Humanities Open*, 11, 101492. https://doi.org/10.1016/j.ssaho.2025.101492

Fox, S., & Rey, V.F. (2024). A cognitive load theory (CLT) analysis of machine learning explainability, transparency, interpretability, and shared interpretability. *Machine Learning and Knowledge Extraction,* 6(3), 1494–1509. https://doi.org/10.3390/make6030071

Fritsch, S.J., Blankenheim, A., Wahl, A., Hetfeld, P., Maassen, O., Deffge, S., et al. (2022). Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *Digital Health,* 8, 1–10. https://doi.org/10.1177/20552076221116772

Gunkel, D.J. (2018). The other question: Can and should robots have rights? *Ethics and Information Technology,* 20(2), 87–99. https://doi.org/10.1007/s10676-017-9442-4

Hazarika, I. (2020). Artificial intelligence: Opportunities and implications for the health workforce. *International Health,* 12(4), 241–245. https://doi.org/10.1093/inthealth/ihaa007

Jha, D., Riegler, M.A., Johansen, D., Halvorsen, P., & Johansen, H.D. (2020). DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. In: *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, p558–564. https://doi.org/10.1109/CBMS49503.2020.00111

Karalis, V.D. (2024). The integration of artificial intelligence into clinical practice. *Applied Biosciences,* 3, 14–44. https://doi.org/10.3390/applbiosci3010002

Khalifa, M., & Albadawy, M. (2024). AI in diagnostic imaging: Revolutionizing accuracy and efficiency. *Computer Methods and Programs in Biomedicine Update,* 5, 100146. https://doi.org/10.1016/j.cmpbup.2024.100146

Khan Rony, M.K., Alrazeeni, D.M., Akter, F., Nesa, L., Das, D.C., Uddin, M.J., et al. (2024). The role of artificial intelligence in enhancing nurses' work-life balance. *Journal of Medicine, Surgery, and Public Health*, 3, 100135. https://doi.org/10.1016/j.glmedi.2024.100135

Khosravi, M., Zare, Z., Mojtabaeian, S.M., & Izadi, R. (2024). Artificial intelligence and decision-making in healthcare: A thematic analysis of a systematic review of reviews. *Health Services Research and Managerial Epidemiology,* 11, 1–10. https://doi.org/10.1177/23333928241234863

Lovett, R.M., Benavente, J.Y., Opsasnick, L.A., Weiner-Light, S., Curtis, L.M., & Wolf, M.S. (2023). Associations between cognitive impairment severity and barriers to healthcare engagement among older adults. *Journal of Applied Gerontology*, 42(7), 1387–1396. https://doi.org/10.1177/07334648231166289

Mashabab, M.F., Al Sheniff, M.S., Alsharief, M.S., & Al Kulayb, A.H.A. (2024). The role of artificial intelligence in healthcare: A critical analysis of its implications for patient care. *Journal of Ecohumanism,* 3(7), 597–604. https://doi.org/10.62754/joe.v3i7.4228

Masudin, I., Restuputri, D.P., Amalia, F., & Oktiarso, T. (2024). The role of smart technology, managerial initiatives and human factors on sustainable manufacturing: A case study of Indonesian oil and gas workers. *Ergonomics*, 67(12), 1884–1908. https://doi.org/10.1080/00140139.2024.2360095

Mikalef, P., Islam, N., Parida, V., Singh, H., & Altwaijry, N. (2023). Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective. *Journal of Business Research*, 164, 113998. https://doi.org/10.1016/j.jbusres.2023.113998

Mrayyan, M.T., Abu Khait, A., Rababa, M., Algunmeeyn, A., Al-Rawashdeh, S., Al-Atiyyat, N., et al. (2024). Professional autonomy in nursing: A concept analysis. *Sage Open,* 14(4). https://doi.org/10.1177/21582440241302129

Natali, C., Marconi, L., Dias Duran, L.D., Miglioretti, M., & Cabitza, F. (2025). *AI-Induced Deskilling in Medicine: A Mixed Method Literature Review for Setting a New Research Agenda*. SSRN, Rochester. https://doi.org/10.2139/ssrn.5166364

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors,* 39(2), 230–253.

Pavuluri, S., Sangal, R., Sather, J., & Taylor, R.A. (2024). Balancing act: The complex role of artificial intelligence in addressing burnout and healthcare workforce dynamics. *BMJ Health and Care Informatics,* 31(1), e101120. https://doi.org/10.1136/bmjhci-2024-101120

Rashid, Z., & Rötting, M. (2021). Evaluation of manual skill degradation due to automation in apparel manufacturing. *Applied Sciences*, 11(23), 11098. https://doi.org/10.3390/app112311098

Salvatore, D., Numerato, D., & Fattore, G. (2018). Physicians' professional autonomy and their organizational identification with their hospital. *BMC Health Services Research*, 18, 775. https://doi.org/10.1186/s12913-018-3582-z

Simmons, C., Conley, M.I., Gee, D.G., Baskin-Sommers, A., Barch, D.M., Hoffman, E.A., Huber, R.S., Iacono, W.G., Nagel, B.J., Palmer, C.E., Sheth, C.S., Sowell, E.R., Thompson, W.K., & Casey, B.J. (2021). Responsible use of open-access developmental data: The adolescent brain cognitive development (ABCD) study. *Psychological Science,* 32(6), 866–870. https://doi.org/10.1177/09567976211003564

Smith, J.D., & Hasan, M. (2020). Quantitative approaches for the evaluation of implementation research studies. *Psychiatry Research*, 283, 112521. https://doi.org/10.1016/j.psychres.2019.112521

Sweller, J. (2011). Cognitive load theory. In: Ross, B.H., editor. *Psychology of Learning and Motivation*. Vol. 55. Academic Press, United States, p37–76. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Taie, M. (2014). Skill acquisition theory and its important concepts in SLA. *Theory and Practice in Language Studies,* 4(9), 1971–1976. https://doi.org/10.4304/tpls.4.9.1971-1976

Tao, D., Chen, Z., Qin, M., & Cheng, M. (2023). Modeling consumer acceptance and usage behaviors of m-health: An integrated model of self-determination theory, task-technology fit, and the technology acceptance model. *Healthcare,* 11(11), 1550. https://doi.org/10.3390/healthcare11111550

Varnosfaderani, M.S., & Forouzanfar, M. (2024). The role of AI in hospitals and clinics: Transforming healthcare in the 21st century. *Bioengineering,* 11(4), 337. https://doi.org/10.3390/bioengineering11040337

Williamson, S.M., & Prybutok, V. (2024). Balancing privacy and progress: A review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare. *Applied Sciences,* 14(2), 675. https://doi.org/10.3390/app14020675

Zhai, C., Wibowo, S., & Li, L.D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments,* 11, 28. https://doi.org/10.1186/s40561-024-00316-7

AUTHOR BIOGRAPHY

Raed Awashreh is a Professor of Management and Public Administration with over 25 years of experience in academia and industry, spanning non-governmental organizations, government, and the private sector. He has taught undergraduate and postgraduate students in the United Arab Emirates, Oman, and the West Bank, supervised theses, and served as a postgraduate examiner and journal reviewer. He has published extensively on topics such as human capital, governance, education, management, policy, leadership, and artificial intelligence applications. Beyond academia, he provides consulting services on strategy, governance, and organizational development. He holds a Doctorate from Flinders University, Australia, and a Master's from the Middlebury Institute, United States of America.

Ensemble learning for enhanced brain tumor diagnosis: A new approach for early detection

Nayla Faiq Othman[1]*, Shahab Wahhab Kareem[1,2]

[1]Department of Technical Information Systems Engineering, Erbil Technical Engineering College, Erbil Polytechnic University, Erbil, Iraq

[2]Department of Computer Science, Bayan University, Erbil, Iraq

*Corresponding author E-mail: nayla.othman@epu.edu.iq

(Received 06 January 2025; Final version received 11 March 2025; Accepted 23 April 2025)

Abstract

Brain tumors represent one of the most extreme and complex types of cancer, requiring unique analysis for powerful remedy and management. Accurate and early identification of brain tumors can greatly enhance patient outcomes and decrease mortality. Nowadays, deep learning aids the medical field a lot by diagnosing magnetic resonance imaging images in brain tumors. The potential of deep learning architectures to improve brain tumor diagnosis accuracy was explored in this work. This study evaluated three different convolutional neural network architectures: AlexNet, VGG16, and ResNet18 as an ensemble model. By leveraging the complementary strengths of these models and applying them to a dataset sourced from local hospitals and public repositories, this research aims to address the challenges in accurate and early brain tumor detection. Our ensemble technique achieved excessive accuracy, demonstrating its potential for reliable computer-aided diagnosis (CAD) in medical imaging. However, while the results indicate an improvement in class overall performance, the novelty of this approach is restrained because it builds upon existing methodologies as opposed to offering a completely new framework. The gathered dataset was used to train and test the models. To enhance the dataset's balance and the models' performance, data were collected from Rizgary Hospital (Erbil) and Hiwa Hospital (Slemani), addressing the underrepresentation of cases from the Kurdistan Region of Iraq (KRI). These image enhancement techniques were applied to two categories: normal and abnormal brain tumors. Several brain tumor datasets are available online for the development of CADs, but not KRI cases, which pose challenges in their classification through deep learning models. This study was implemented with Python programming language. Out of the three models, ResNet had the highest accuracy of 98.66%, VGG16 had an accuracy of 97.8%, and AlexNet had an accuracy rate of 97.666%. The ensemble, using both majority voting and weighting voting strategies, achieved an accuracy of 98.33%.

Keywords: Brain Tumor, Computer-aided Diagnosis, Magnetic Resonance Imaging, Transfer Learning

1. Introduction

This document is a template for Microsoft Word. Absolute confidence, human needs, technological paradigms, hospitals, and organizations have all significantly improved in the twenty-first century, with computing and records technology displacing other fields like an octopus (Alfonse & Salem, 2016; Nawaz et al., 2022). The human brain, situated within the skull, is an essential organ responsible for various functions, governed by a network of billions of neurons that coordinate electrical and chemical impulses, influencing our experiences and lives (Mathivanan et al., 2024). The brain is an extraordinary organ, serving as a cornerstone in the domains of cognition, emotion, and humanity. The mind consists of remarkable components, each with distinct functions, exemplifying complexity. The cerebral cortex, a convoluted outer layer, governs awareness, whereas the cerebellum is responsible for balance and coordination. The formation of an abnormal cellular proliferation, manifested as a mass or lump, is referred to as a tumor or neoplasm (Lauko et al., 2022). It is an aberrant proliferation of cells in the essential spinal canal or the brain. Two types of brain tumors exist, classified as benign and malignant. Meningiomas and low-grade gliomas are alternative designations for benign tumors, whereas malignant tumors are referred to as glioblastoma multiforme and high-grade gliomas. The most common type of brain neoplasm is referred to as a malignant tumor (Leena & Jayanthi, 2020). Brain tumors characterized by a homogeneous structure are classified as benign, as they do not include malignant cells. This can be entirely resolved with surgical removal or radiological monitoring, ensuring they do not recur. The life-threatening tumor is a malignant neoplasm characterized by a heterogeneous structure containing the majority of cancerous cells. The treatment for malignant tumors involves chemotherapy, radiation, or a combination of both. Consequently, timely analysis of brain tumors is a crucial element in the advancement of therapy (Harnod et al., 2014). Consequently, it is essential to ascertain the dimensions of the brain tumor and establish the associated grade. Recently, advancements in magnetic resonance imaging (MRI) have significantly improved the detection rate of brain malignancies. It has established a crucial foundation in the study of tumor diagnosis and image registration (El Kader et al., 2021). Tasked with fostering awareness of self and surroundings, including influencing muscular actions (Khairandish et al., 2022). Every responding notion, emotion, and plan is facilitated by the brain. MRI and computed tomography scans are diagnostic modalities utilized to reveal the internal structure of the brain. MRI is advantageous for delineating soft tissues and revealing the internal architecture of the framework. The MRI delineates the contrast between normal and abnormal tissues. This work utilizes MRI data to identify the damaged areas in the brain. MRI utilizes a powerful magnet and radio waves to generate a distinctive image of the patient's internal organs (Praveen & Agrawal, 2016). MRI poses no danger to the human body as it does not involve radiation. It provides statistics regarding anomalous tissues for diagnostic purposes. MRI is non-invasive, making it highly popular among individuals and commonly utilized for assessing tumor size, shape, and type (Ramtekkar et al., 2023).The MRI scans can accurately reflect the brain's structure and function, allowing for multi-angle and multi-modal imaging with minimal harm to the human body. Consequently, it is extensively utilized in the identification of neurological disorders (Fang & Wang, 2022). The amalgamation of deep learning and artificial intelligence has markedly enhanced medical image processing, resulting in substantial progress in the detection, diagnosis, and characterization of diverse medical disorders. This has allowed healthcare providers to make better-informed decisions, especially in the precise classification of cancer kinds, including lung and breast cancer (Siva Raja & Rani, 2020). This integration has led to earlier diagnoses, enhanced treatment decisions, and better patient outcomes. Artificial intelligence is essential in surgical planning, facilitating accurate segmentation of lesion margins and cerebral structures while balancing intervention with the preservation of quality of life (Arumugam et al., 2024). It forecasts problems, recurrence rates, and therapeutic responses, directing appropriate follow-up tactics and facilitating personalized patient management through customized screening protocols. Transfer learning is a machine learning methodology that has garnered considerable interest in the medical domain, emphasizing the utilization of pre-trained models on extensive datasets for particular tasks. Transfer learning is an essential instrument in medical image analysis, facilitating the development of high-performance models while minimizing training duration and computing expenses (Pacal, 2024). With the evolution of the field, transfer learning is anticipated to assume a more pivotal role in enhancing patient care. Numerous transfer learning models, such as VGG16, ResNet, and AlexNet, have demonstrated significant effectiveness in this domain. Transfer learning models, employing the depth and intricacy of neural networks, are utilized to discern complicated patterns in medical images (Mathivanan et al., 2024). This versatile approach extends beyond these well-known architectures, with numerous other models contributing to the growing range of tools for medical imaging analysis (Kumar & Ma, 2024). Transfer learning in medical imaging has markedly accelerated the development process and enhanced the performance and accuracy of pre-trained models, facilitating swifter and more precise diagnoses of malignant tumors, especially in their detection and categorization (Remzan et al., 2024). The efficiency improvements realized by transfer learning models have substantial implications for patient care, as early diagnosis and accurate categorization of cancer types are crucial for commencing prompt and focused treatment methods (Mandle et al., 2024). The interplay among deep learning, artificial intelligence, and transfer learning is set to revolutionize medical picture analysis. The integration of these technologies enhances the capacities of healthcare workers and has the potential to improve patient outcomes and transform medical diagnostics (Gül & Kaya, 2024). This study makes its key contribution through pre-trained convolutional neural network (CNN) ensemble methods that enhance brain tumor diagnostic accuracy. This study combines the AlexNet, ResNet18, and VGG16 networks through an ensemble system, which shows the advantages of multiple architecture implementations during prediction. Through the combination of majority and weighted voting systems, the classification precision increases. The first section of this paper is the introduction, and the second section is the related work, where we show the work of other papers, their methods, and results. Section 3 of the paper discusses the methods of our paper and how they are utilized. Section 4 is about our preparations and the experiments we have done. The fifth section is the result of our experiments and comparison with other datasets. Finally, Section 6 concludes the paper.

2. Related Work

The main goal of this study is to review and comprehend brain tumor classification and recognition strategies established globally between 2015 and 2024. The current study reviews the most widespread procedures for detecting brain cancer that have been made accessible universally, in addition to observing how effective computer-aided diagnosis systems are in this process. Current relevant review papers, along with their respective specifics and highlights, are discussed in this section.

ZainEldin et al. (2023) propose a method for diagnosing and classifying brain tumors using CNNs, and an adaptive dynamic sine-cosine fitness grey wolf optimizer is presented in this paper. The proposed model, Brain Connectivity Matrix-CNN, outperforms other models when evaluated on the Brain Tumor Segmentation Challenge (BRATS) 2021 dataset with an accuracy of 99.99%. It features both hyperparameter tuning of a CNN and segmentation capabilities utilizing a 3D volumetric data segmentation (3D UNet) architecture. In a study by Abdusalomov et al. (2023), a new method of brain tumor detection using deep learning was developed, the central idea of which is based on the enhancement of the You Only Look Once (YOLOv7) model integrated with additional components such as the Convolutional Block Attention Module, Spatial Pyramid Pooling Fast Plus(), and Bidirectional Feature Pyramid Network. This model is implemented and trained using a dataset of MRI images, which allowed for the achievement of satisfactory accuracy equal to 99.5%. It focuses on the correct detection of glioma, meningioma, and pituitary tumors and performs better than previously existing methods in this regard. Problems of small tumor size and instability of localization were resolved due to effective feature extraction. Ranjbarzadeh et al. (2021) suggest an innovative method for brain tumor segmentation based on CNNs using a Distance-Wise Attention Mechanism. The model achieves computational efficiency and lower overfitting by concentrating only on the more localized areas of interest and using pre-processing for unnecessary information. Experiments on the BRATS 2018 dataset provide competitive results with good performance in tumor localization and segmentation. Mahmud et al. (2023) aim to improve the efficiency of deep learning in detecting brain tumors from MRI scans. A CNN framework is presented and evaluated against other architectures, including ResNet-50, VGG16, and Inception V3. It used a sample of 3264 MRI scans and emerged with the CNN model, achieving the highest performance at 93.3% accuracy and an area under the curve (AUC) at 98.43%. Based on the study, it focuses on the early stage of tumor detection to further reduce the mortality rate of the patients using the pre-processing and data augmentation methods. Noreen et al. (2020) recommend using deep learning methodology, attempting an automated diagnosis of brain tumors focused on differentiating among the three types: gliomas, meningiomas, and pituitary tumors, utilizing MRI. It uses two pre-trained models, Inception-v3 and DenseNet201, to extract features from different layers and combine them in order to improve the classification. The feature-level fusion technique is known to provide multiscale and dense embeddings, which are better than other existing techniques. The Inception-v3 model achieved an accuracy of 99.34%, and the DenseNet 201 model achieved 99.51% accuracy on the testing dataset. The results support the second hypothesis and show that brain cancer classification through feature fusion is effective and can be considered a viable approach for medical imaging tasks. Mostafa et al. (2023) tackled the problem of brain tumor segmentation through deep learning techniques with MRI images. Khan et al. (2022) present a solution that offers complete automation based on a CNN model, which is a modern deep-learning technique making use of datasets containing multimodal MRI images, including the BRATS. Major results include the segmentation of brain tumors into classes: Necrotic and edema enhancing, with the achievement of validation accuracy of 98% even in different settings. The methodology includes: preprocessing, data augmentation, model training including U-Net sampling techniques, and model optimization with cross-entropy loss and categorical method. The research focuses on progress made in the area of automated brain tumor diagnosis and efforts made to enhance efficiency and the level of precision to be used in assisting medical personnel. The paper describes a novel hierarchical deep-learning model for brain tumor identification. The developed hierarchical deep learning with 2D and 3D features for brain tumor segmentation (HDL2B-TUMOR-CLASSIFIER) system consists of CNNs that divide the studied brain tumors into four types: Glioma, meningioma, pituitary, and no tumor. The model achieved a 92.13% accuracy, which is significantly higher than many of the methods known: the model's miss rate was 7.87%. A dataset of 3,264 images was used in three stages: Preparation, training, and validation. It focuses on assisting clinical diagnoses of brain tumors by increasing the speed of detection and the accuracy of classification. Ghaffari et al. (2020) show that the classification and segmentation of brain tumors demonstrate substantial development when deep learning and machine learning models are applied to the process. The researchers utilized multimodal MRI together with machine learning and CNN on the BRATS dataset to tackle the problem of costly and time-consuming physical brain tumor segmentation. The research analyzed multimodal MRI to develop a benchmark that showed more than 90% success through previous work comparisons. In a study by Gao et al. (2022), a multi-scale CNN was used on the BRATS dataset for dealing with heavy computational needs that affect 3D convolution networks and the independent nature of single-view 2D slices. This model implemented multi-scale approaches to effectively show how neighborhood size affected segmentations in three-dimensional CNNs while achieving superior tumor segmentation performance at enhancing tumor (75%), whole tumor (90%), and tumor core (84%). According to a study by Albalawi et al. (2024), the application of CNN technology to process the Kaggle database successfully solved problems with manually detecting brain tumors from MRI images because of diverse tumor dimensional characteristics and variations in shape and brightness. The study created a high-performing classification system that differentiated glioma from meningioma and pituitary tumors with a 98.04% success rate. The hybrid deep learning framework by Rasheed et al. (2023) combined AlexNet with ResNet-18 using a support vector machine (SVM) to analyze MRI data collected at Nanfang Hospital and Tianjin Medical University General Hospital. Researchers worked on tumor brain classification through the combination of CNN algorithms with SVM methods because they wanted improved precision in detecting tumor and non-tumor patterns. The combined model reached a performance level of 95.1% accuracy coupled with 95.25% sensitivity and 98.50% specificity. Jader et al. (2024) tackled painstaking issues associated with the segmentation and classification of brain tumors that are prone to human error. Its tasks include the classification of MRI images into four categories: pituitary, glioma, meningioma, and nontumorous. Subsequently, the research carried out employed VGG-16, ResNet-50, and AlexNet models, which were based on transfer learning, and consolidated them into an ensemble model in order to increase classification accuracy. Achieving greater classification accuracy than the older methods, the ensemble model performed better than other methods such as Naïve Bayes, decision trees, random forests, and deep neural networks. The results returned were 99.16% accuracy, 98.47% sensitivity, 98.57% specificity, 98.74% precision, 98.49% recall, and 98.18% F1 score. The primary innovation of this paper is brain tumor classification ensemble transfer learning, which is a new diagnosis approach. The results prove this method is effective. The deep learning method for classifying various stages of Alzheimer's disease (AD) presented in Ramzan et al. (2020) was conducted on a sample that consists of resting state functional MRI images collected by the AD Neuroimaging Initiative, which is comprised of 138 subjects. The model classifies six stages of AD, which include: Cognitively normal, significant memory concern, early mild cognitive impairment, mild cognitive impairment, late mild cognitive impairment, and AD. The goal of this study's research is to enhance the diagnostic process for the classification of AD at its early stages, which poses a great challenge due to similarities in symptom presentation and a lack of multiclass classification approaches. Doing so allows for the contribution of this research to be more clinically relevant. Residual learning, transfer learning, and deep learning methods of this work have been shown to significantly improve classification performance. Results show that the average accuracy of the ResNet-18 model that was fine-tuned is 97.88%. This achieves the best results to date and exceeds accuracy in classifier development for all stages of AD. In Roopa et al. (2022), a CNN algorithm is suggested for detecting brain tumors in MRI images. The study employed a dataset of 3,264 MRI scans, classified into four categories: Glioma, meningioma, pituitary tumor, and no tumor. The challenge in focus is the timely and precise identification of brain tumors for effective treatment management. The study analyzes the performance of the designed CNN against the ResNet-50, VGG16, and Inception V3 benchmark models, measuring the accuracy, recall, AUC, and loss. The main contribution of the paper is uncovering the effectiveness and benefit of the proposed CNN model, which outshone other models by achieving an accuracy of 93.3%, an AUC of 98.43%, a recall of 91.19%, and a loss of 0.25. The authors' conclusions indicate that the CNN algorithm put forward is a dependable answer for detecting brain tumors from MRI images and is more accurate and robust than popular transfer learning algorithms. Putzu et al. (2020) present a classifier model that is based on CNNs and is built on top of AlexNet. Their research attempts to solve the content-based image retrieval problem using CNNs for feature extraction and relevance feedback (RF). The dataset encompasses Caltech-101, Caltech-256, Flowers-102, and SUN-397, which have different classes of images for the retrieval tests. The most important part of this paper is the introduction of two CNN architectures for RF, one of which has an original AlexNet depth but a last layer adapted to feedback, and the other one has an additional layer for better user feedback incorporation. The work also investigates some query refinement approaches such as relevance scoring and mean feature computation to improve retrieval accuracy. The experimental results confirm that the proposed CNN based RF methods increase the retrieval accuracy, where RF using classification outperformed feature extraction in the later iterations, which proves the effectiveness of tuned CNNs on interactive image search systems. The study by Al-Hadidi et al. (2020) features an advanced technique that employs a multi-model CNN using Xception, DenseNet-201, and EfficientNet-B3 as classifiers to identify brain tumors. The study seeks to address the issue associated with low classification accuracy in brain MRI images, which stems from variations in tumor size, shape, and position. The data set contains THOMAS (Dataset 1) and NICKPARVAR (Dataset 2), which each contain MRI photographs of glioma, meningioma, pituitary tumor, and other no tumor categories. The algorithm developed in this work is innovative in the sense that it bases model selection on the validation accuracy and false positive rate to combine multiple CNN models and thus results in better classification efficiency. The results show that multi-model CNN outperformed single CNN models, achieving 97.74% accuracy on Dataset 1 and 99.69% on Dataset 2, which is an improvement between 1.29% and 4.19% for Dataset 1 and 0.22–0.61% for Dataset 2. Based on these results, it can be concluded that traditional single-model approaches to brain tumor detection are less effective than the multi-model approach. Abdullah et al. (2024) suggest incorporating deep CNN (DCNN) techniques, which are based on the VGG-16 model for tumor identification in MRI scans. This dataset has 253 MRI images of the brain—155 with identified tumors and 98 without. These images were obtained from Kaggle. The research is aimed at providing a solution for automated and accurate brain tumor detection, which circumvents manual methods that are labor-intensive and highly subjective. The primary alteration the authors made to the architecture of VGG-16 was a substitution of the last max-pooling with Global Average Pooling, which mitigated the effects of overfitting and improved generalization. The accuracy achieved by the proposed DCNN model is 96%, which is higher than the accuracy achieved using conventional methods. The model performed remarkably well, as it achieved a precision of 0.93, a sensitivity of 1.00, an F1-score of 0.97, Cohen's kappa of 0.91, and an AUC of 0.95, an indicator of the effectiveness of the tool for clinical experts to improve brain tumor detection and expedite treatment measures. Asif et al. (2022) focus on the classification of brain tumors using pre-trained DCNNs like VGG-19, VGG-16, ResNet50, and Inception V3 on MRI images. The dataset consists of 305 brain MRI images, including tumorous and non-tumorous cases, collected from publicly available sources. The study addresses the problem of fully automated and accurate classification of brain tumors, which is significant for prompt diagnosis and therapeutic intervention. The main effort stems from determining how effective transfer learning is with the pre-trained DNNs with the small sample set, and the poster shows that high accuracy can be achieved. The results indicate that VGG-19 achieved the highest accuracy of 99.48%, followed by VGG-16 (99%), ResNet50 (97.92%), and Inception V3 (81.25%). These results support the claim that the automated classification of brain tumors with models assisted by a transfer learning framework is practical and does not require elaborate feature extraction procedures.

3. Methodology

3.1. Dataset

Three deep learning models, namely AlexNet, ResNet18, and VGG16, were utilized for the implementation of this proposed approach. The trained deep-learning models operated on data obtained from both local hospitals and public repositories. The main research objective of this study does not revolve around transfer learning innovations but rather deploying ensemble voting methods to enhance classification accuracy. The ensemble framework carried out predictive analysis by using majority voting and weighted voting methods to boost classification outcomes. The proposed approach uses existing ensemble learning techniques for its effectiveness while mainly depending on established principles. The main advancement occurs from the practical application of these models in medical imaging contexts instead of creating new computational structures. Future explorations need to use deep features or advanced fusion techniques to improve methodological innovation.

For developing the model, a dataset was collected, which had 200 cases of patient MRI images using Digital Imaging and Communication in Medical. Of those 200 cases, 100 of them were normal, and 100 of them were abnormal; the dataset normal cases had 38 male patients and 62 female patients, and in the abnormal 67 male and 33 female, which means males are infected with brain cancer, the age of the patients was in the range of (19–98) years old. It required 3 months of data collection from Rizgary Hospital and Hiwa Hospital from the cases of 2024; 2 months were spent for data cleaning, which gave the outcome of 3000 images 1500 of them were normal, and 1500 were abnormal with a resolution of 512 × 512, and 15 images from each case, and Table 1 shows the detail of dataset split, which was divided into 80% for training and 20% for testing. The dataset was organized into two main folders: Normal and abnormal. Within each folder, images were further divided into test and train subfolders. Specifically, 300 images were allocated for testing and 1,200 images for training in each category, as shown in Table 1.

Overall, training and testing the data for each model took between 2 and 5 h, different for each model, using a Legion 13th Gen Intel (R) Core (TM) i7-13650HX 2.60 GHz having an Nvidia RTX 4060 GPU with 16 GB of RAM DDR5 4,800 MHz and 1.5 TB SSD NVME hard disk 3,500 mb/s.

3.2. Methods

The proposed model, illustrated in Fig. 1, employs three well-known transfer learning approaches—ResNet, VGG16, and AlexNet, an ensemble model—to create three classes for analyzing and estimating: the recommended frame. The data undergoes three transfer learning techniques, and following analysis, it's divided into an 80% training set and a 20% testing set. Using Pytorch, this code implements and assesses three deep learning models, AlexNet, ResNet, and VGG16, based on an image dataset. To begin with, the code imports and processes the image dataset by creating a training and testing split. The models are adjusted for their required number of output classes, and weights have been initialized from pre-trained models. At the phase of training, loss and accuracy scores are computed for each batch and overall raw and weighted accuracy statistics are recorded across a number of epochs. After training has been completed, it tests the models on the test data and scores them in terms of accuracy, confusion matrix, as well as a classification report. Also, the code performs an ensemble of the models by using majority and weighted voting and computes the ensemble accuracy as well. Multi-class receiver operating characteristic (ROC) and AUC metrics to analyze the model's multi-class classification performance are also calculated and presented. Confusion matrices and comprehensive classification reports for training and test sets are presented to evaluate the performance of the models according to the various classes. Fig. 2 shows the normal and abnormal images of brain MRI.

Table 1. Data split detail

Phase	Abnormal (50%)	Normal (50%)	Total
Train	1,200	1,200	2,400
Test	300	300	600
Total	1,500	1,500	3,000

3.3. Pre-processing

Pre-processing refers to the input images that will go for further analysis to improve the effective analysis. It comprises eliminating artifacts for better focus, concentrating on the area of the brain by excluding non-brain tissues, and cutting up the image with segmentation to define its meaningful parts (Khairandish et al., 2022). In the context of image processing, pre-processing regularly includes resizing pixels to a hard and fast size, normalizing pixel values to a trendy range, applying data augmentation techniques like flipping or rotating to growth variety, and converting photos into tensor formats appropriate for computation. These steps collectively ensure the statistics are optimized for effective model learning and evaluation (Rasheed et al., 2023).

3.3.1. ResNet

ResNet, which stands for residual network, is a deep-learning architecture proposed by Ramzan et al. (2020) as a way of solving problems posed during classification. The primary advancement of this system is the deployment of the so-called residual or skip connections, which make it possible to directly add portions of a layer back into a layer (Jader et al., 2024; (Mahmud et al., 2023). The network is able to learn residual mappings, which makes the whole optimization process less complicated and also reduces the problems of vanishing gradients that are experienced in deep networks. ResNet architecture is built using modular residual blocks each of which contains convolutional layers that perform batch normalization and rectified linear unit (ReLU) activation in conjunction with skip connections (Ramzan et al., 2020). ResNet reliability and versatility have thus made it a more commonly used model to transfer learning between several tasks. With the ability to train very deep networks, ResNet has served as a point of reference in deep learning, proving that if done well within trained deeper models, they yield better results (Ramtekkar et al., 2023). The model consists of several blocks, as shown in Fig. 3.

3.3.2. VGG16

The structure is distinguished by a uniform and deep architecture, which contains 16 weight layers of three fully connected and thirteen convolutional layers. The unique feature of VGG16 is the use of small 3 × 3 convolutional kernels in each layer of the network. A smaller kernel is not a drawback, as many convolutional layers may be accustomed to compute

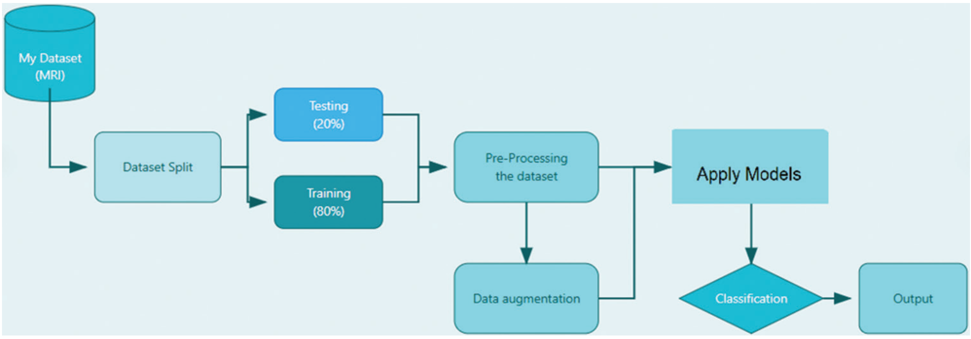


Fig. 1. Proposed model architecture

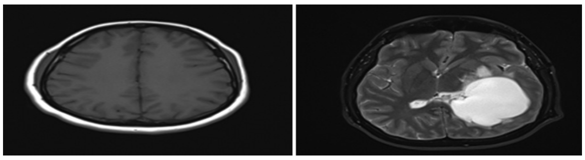


Fig. 2. Normal (left) and abnormal (right) images of brain magnetic resonance imaging

complex and hierarchical features (Roopa et al., 2022). Definitely a typical approach for the mentioned architectures supporting wider kernels as opposed to the deeper network (Mahmud et al., 2023); (Sulistyowati et al., 2023). The structural design of VGG16 is quite simple, adopted sequential architecture without skip or residual connections, and this makes it very easy to build and comprehend. The architecture's depth and its design enable it to perform effectively in retrieving the details of visual features; hence, it is widely used for image classification, transfer learning, and feature extraction (Bansal et al., 2023). Fig. 4 depicts the architecture of VGG16.

3.3.3. AlexNet

AlexNet features eight layers in total, namely, five convolutional and three fully connected layers. The model proposed multiple measures to train deep neural networks more efficiently (Mathivanan et al., 2024). As for the architecture, the authors employed ReLU activations, which considerably sped up the training process, dealing with the vanishing gradient problem that is extremely common for the use of sigmoid or tanh activations. Moreover, AlexNet included dropout into the network to prevent the issue of overfitting and also used data augmentation to extend the limited training set and its variability. The biggest advancement, however, was the training on GPUs, which allowed for harnessing their parallelism in order to tackle the numerous computationally heavy operations needed in the model (Jader et al., 2024). It was also the case that AlexNet implemented local response normalization, which sought to increase the feature learning of the model through neuronal competition (Putzu et al., 2020). The success of the architecture proved the ability of deep learning to tackle complicated vision problems, which in turn led to the construction. AlexNet allowed for the development of the modern architecture of CNNs, establishing it as a landmark in the evolution of deep-learning algorithms (Nawaz et al., 2022). The model consists of several layers, as shown in Fig. 5.

4. Preparation and Assessment Experiments

In this experiment, a large dataset of 3,000 images was collected from 200 cases of patients, 1,500 of them were benign, and 1,500 were Malignant. All the data was collected within 3 months in Erbil Rzgray Hospital and Sulaymaniyah Hewa Hospital to ensure the effectiveness of the training and testing phases. We also collected two more datasets to access a robust computing environment. The datasets that we collected from Kaggle were Brain Tumor Image Dataset with Grayscale Normalization and Zoom and Brain MRI Images for Brain Tumor Detection, which were used to compare with the data we collected. Importantly, the same datasets were utilized for all advanced models, encompassing both the training set and the testing set. The success of our models can be attributed to the collaborative contributions of Sklearn, TensorFlow, and PyTorch. For optimal performance in all high-end models, a block size of 32 was determined to be the most effective. Table 2 illustrates the hyperparameter details of transfer learning models.

The evaluated model, ResNet18, exhibited superior performance, achieving the lowest testing loss of 0.0235 at epoch 30 and testing accuracy of 99.33.00%. Among the evaluated models, AlexNet exhibited the lowest testing loss of 0.1026 at epoch 32 but experienced the most fluctuation in testing accuracy. It ultimately achieved testing accuracy of 98.17%. VGG16 demonstrated promising results, achieving training accuracy of 98.83% and loss of 0.0426 at epoch 39, respectively, Figs. 6-8 describe the result of the three models from each epoch.

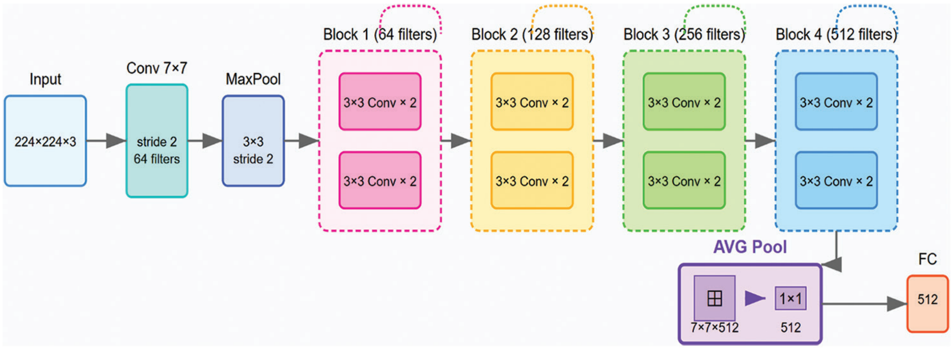


Fig. 3. Architecture of ResNet 18

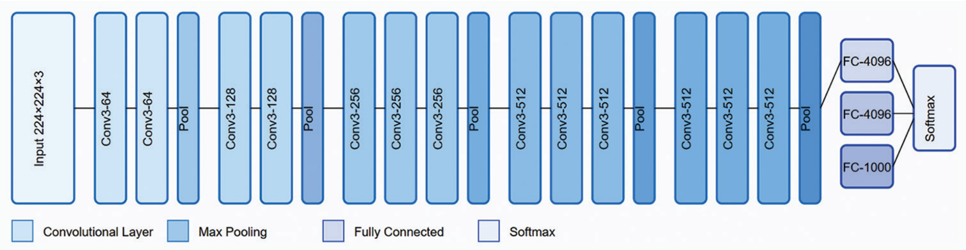


Fig. 4. Architecture of VGG16

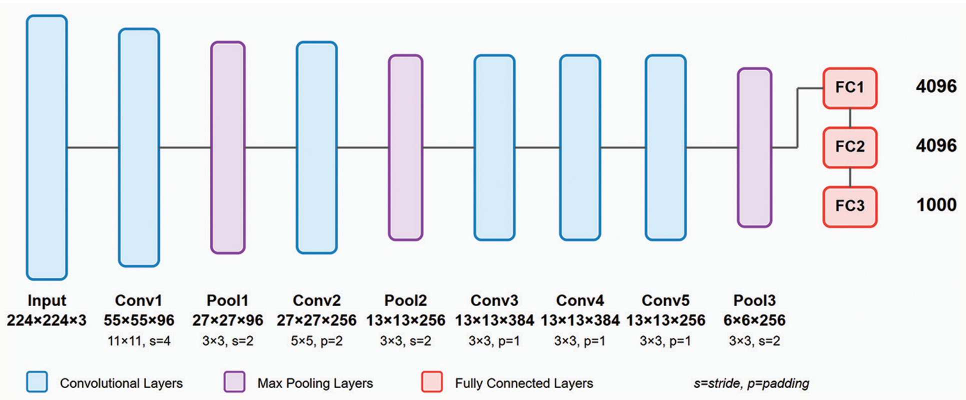


Fig. 5. Architecture of AlexNet

Table 2. Hyperparameters of transfer learning models for image classification

Quantifying performance and evaluation	Assessing measurement outcomes
Size of the batch	32
Optimizer	Adam
No. of epochs	50
Rate of learning	0.0001
Evaluation criterion	Cross-entropy loss function
Training	Confusion matrices

An ensemble of three models combines their predictions to enhance accuracy and robustness. This can be executed through strategies such as voting (majority or weighted), where the final output is primarily based on the consensus or confidence of each model. For brain tumor classification, an ensemble of ResNet18, VGG16, and AlexNet should use majority or weighted voting to supply extra reliable predictions. The pseudocode below shows the process of combining all three models and running them together. It also shows the result of the training and testing data accuracy for each of the models. Furthermore, we got the ensemble predictions of all the models together; the majority voting was (98.33%), and the weighting voting was (98.33%). Lastly, we also found the ROC curves, an AUC of 1, and plot accuracies for the models over the epochs as shown in Figs. 9-11, which show the architecture of the ensemble model.

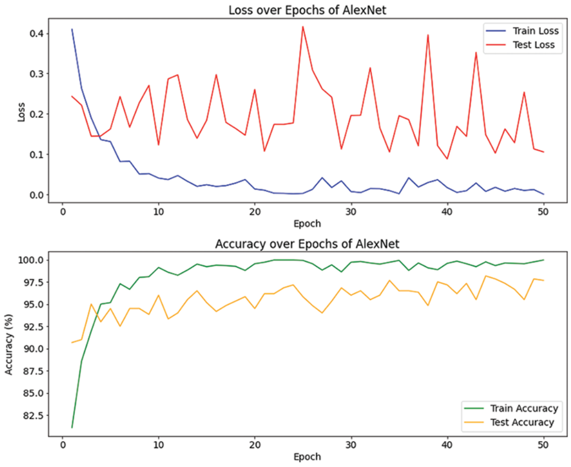


Fig. 6. Training and testing loss and accuracy of AlexNet

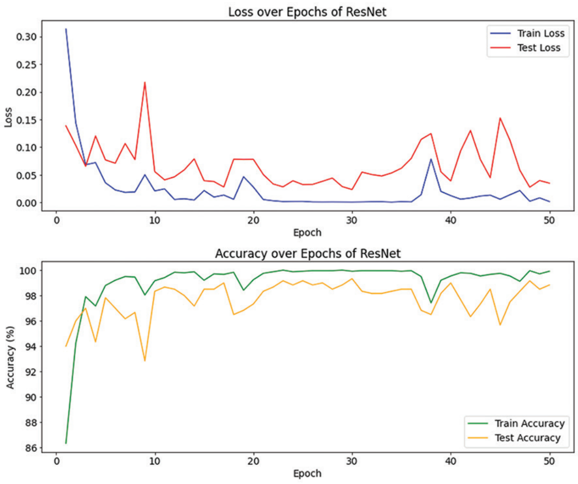


Fig.7. Training and testing loss and accuracy of ResNet

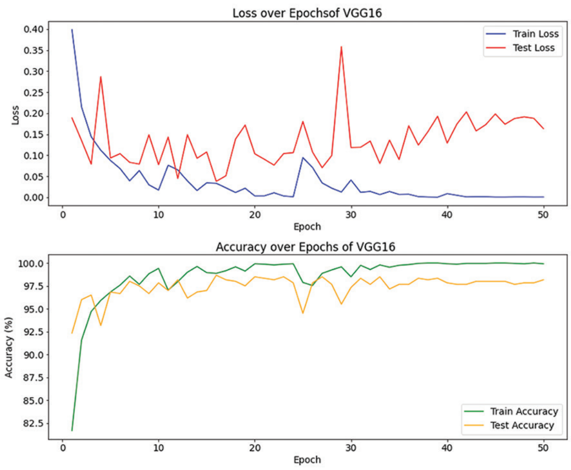


Fig.8. Training and testing loss and accuracy of VGG16

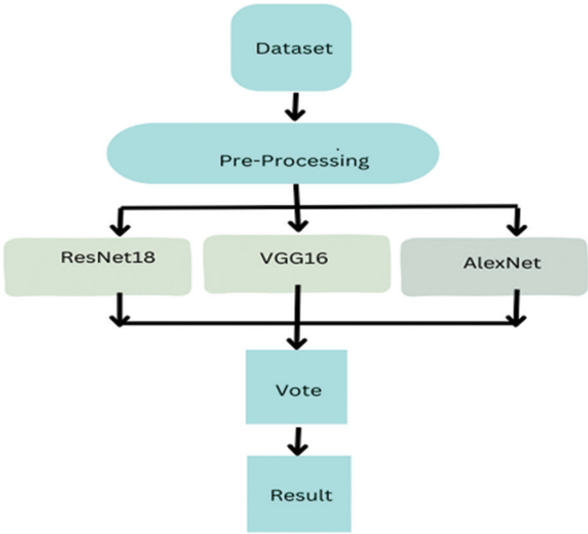


Fig. 9. Ensemble of AlexNet, ResNetv18, and VGG16

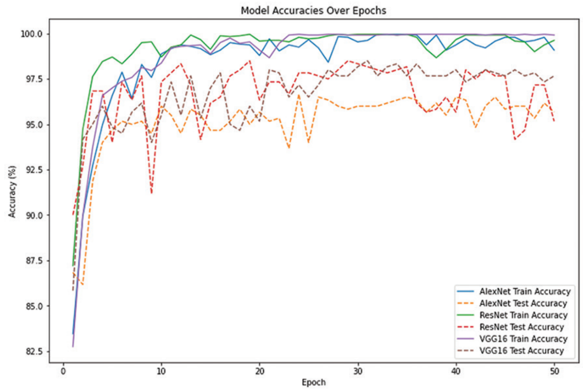


Fig. 10. Ensemble accuracy of AlexNet, ResNetv18, and VGG16

5. Experimental Results and Comparison

5.1. Results and Discussion

The primary aim of the learning technique presented is to develop several models in this study. The initial ResNet represents a specific learning methodology. The second model is VGG16, and the third is AlexNet. The accuracy, specificity, and sensitivity computed during the testing in this work serve as the evaluation metrics for the three models in this system.

Table 3 proposes an evaluation metric for the ResNet18, VGG16, and AlexNet models' performance in training and testing out datasets, differentiating between "normal" and "abnormal" instances of the dataset. These metrics contain precision, F1-rating, and support for both datasets.

- Accuracy: Accuracy refers to the ratio of the true patterns to the summation of entire patterns. It can be expressed as

1 Setup and Preprocessing
- Define dataset path and image transformations.
- Load dataset, split into training (80%) and testing (20%).
- Create Data Loaders for training and testing.

2 Define Models
- Create AlexNetModel, ResNetModel, VGG16Model classes.
- Modify last layer of each model to match the number of classes.

3 Initialize
- Load models and move to CPU/GPU.
- Set training parameters (epochs, learning rate, loss function).

4 Train Model
- For each batch in training data:
- Forward pass, calculate loss, backpropagate, update weights.
- Track training loss and accuracy.

5 Evaluate Model
- For each batch in testing data:
- Forward pass, calculate loss, track predictions.
- Compute test loss and accuracy.

6 Train and Evaluate All Models
- For each model (AlexNet, ResNet, VGG16):
- Train and evaluate for each epoch.
- Store training/testing metrics.

7 Ensemble Predictions
- Predict on test set using all models.
- Combine predictions using majority voting and weighted voting.
- Calculate ensemble accuracy.

8 Performance Visualization
- Plot accuracies for all models over epochs.
- Compute and plot ROC curves (binary or multi-class).

9 Metrics
- Calculate and display confusion matrix, accuracy, and classification report for training and testing datasets.

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positives} + \text{False Positives} + \text{True Negative} + \text{False Negative}} \times 100$$

- Precision: The percentage of accurately projected positive observations to the total projected positives

Table 3. Training and testing all models

Model	**Type**	**Precision**	**Recall**	**F1-score**
ResNet18				
Testing Abnormal	Normal	0.99	0.98	0.99
	0.98	0.99	0.99	
Training Abnormal	Normal	1	1	1
	1	1	1	
VGG16				
Testing Abnormal	Normal	0.97	0.99	0.98
	0.99	0.97	0.98	
Training Abnormal	Normal	1	1	1
	1	1	1	
AlexNet				
Testing Abnormal	Normal	0.97	0.99	0.98
	0.99	0.97	0.98	
Training	Normal	1	1	1
	Abnormal	1	1	1

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Positives}}$$

Greater precision shows fewer false positives.

- Recall (or sensitivity): The percentage of accurately projected positive observations to all actual positives

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Negative}}$$

Greater recall shows fewer false negatives.

- F1-Score: The harmonic mean of precision and recall

$$F1_{\text{score}} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

A high F1-score shows that the model balances precision and recall in a good way.

Fig. 12 illustrates the ROC curves for all three models. ResNet is exceptionally proficient in differentiating between the two classes, exhibiting minimal false positives and a high rate of true positives across various thresholds. An AUC of 1 demonstrates high model performance. VGG16 also had a proficient ROC curve with a result of 1, and lastly, AlexNet had an ROC curve result of 1.

5.2. Comparison

Table 4 provides a detailed performance evaluation for the testing of three transfer learning

DOI: 10.6977/IJoSI.202506_9(3).0007
N. F. Othman, S. W. Kareem//Int. J. Systematic Innovation, 9(3), 109-123 (2025)

Table 4. Performance comparison of the author's dataset with two other datasets

a-our dataset (3000imges)				**b-Brain Tumor Image Dataset with Grayscale Normalization and Zoom (3096 images)**				**c-Brain MRI images for brain tumor recognition (252 images)**			
	Precision	**Recall**	**F1-score**		**Precision**	**Recall**	**F1-score**		**Precision**	**Recall**	**F1-score**
Normal	0.98	0.99	0.98	Normal	0.97	0.99	0.98	Normal	0.86	0.90	0.88
abnormal	0.99	0.97	0.98	Abnormal	0.99	0.98	0.99	Abnormal	0.93	0.90	0.92
AlexNetAccuracy: 97.666%				AlexNet Accuracy: 98.13%				AlexNet Accuracy: 94.11%			
ResNet Accuracy: 98.66%				ResNet Accuracy: 99.25%				ResNet Accuracy: 96.07%			
VGG16 Accuracy: 97.8%				VGG16 Accuracy: 98.88%				VGG16 Accuracy: 94.117%			
Ensemble (majority vote) accuracy: 98.33%				Ensemble (Majority Vote) Accuracy: 98.5%				Ensemble (Majority Vote) Accuracy: 92.15%			
Ensemble (weighted vote) Accuracy: 98.33%				Ensemble (Weighted Vote) Accuracy: 98.5%				Ensemble (Weighted Vote) Accuracy: 92.15%			
AlexNet: Epoch 32 is the best, with a Test Loss of 0.10261 and a Test Accuracy of 98.17%.				AlexNet: Epoch 30 is the best, with a Test Loss of 0.1883 and a Test Accuracy of 98.88%.				AlexNet: Epoch 4 is the best, with a Test Loss of 0.2836 and a Test Accuracy of 96.08%			
ResNet: Epoch 30 is the best, with a Test Loss of 0.0235 and a Test Accuracy of 99.33%.				ResNet: Epoch 35 is the best, with a Test Loss of 0.0345 and a Test Accuracy of 99.63%.				ResNet: Epoch 16 is the best, with a Test Loss of 0.1488 and a Test Accuracy of 96.08%			
VGG16: Epoch 39 is the best, with a Test Loss of 0.0426 and a Test Accuracy of 98.83%.				VGG16: Epoch 27 is the best, with a Test Loss of 0.0493 and a Test Accuracy of 99.63%.				VGG16: Epoch 10 is the best, with a Test Loss of 0.1074 and a Test Accuracy of 96.8%.			

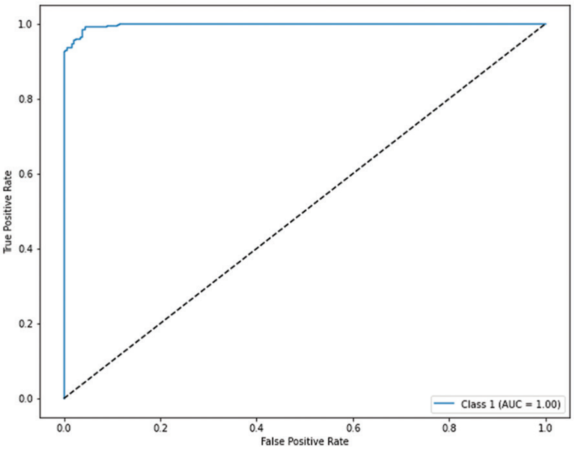


Fig. 11. Receiver operating characteristic curve Ensemble of AlexNet, ResNetv18, and VGG16

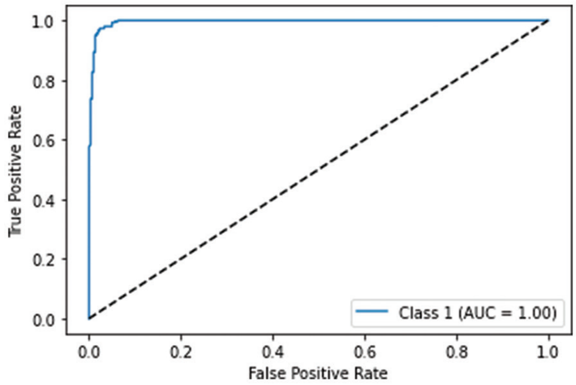


Fig. 12. Receiver operating characteristic curve

models, AlexNet, ResNet, and VGG16, on three datasets of different sizes and complexity. Below is a performance comparison focusing on the key targets. I used two other datasets to compare with my dataset. The results show that my dataset has a better accuracy rate than the other two datasets.

- Dataset Size Impact:
  - My dataset (3000 images) enhances the overall.
  - Performance due to higher model generalization, as visible in the 98.33% accuracy on your dataset the usage of ensemble strategies.
  - Smaller datasets display barely decreased accuracy due to constrained schooling variety.
- Model Performance:
  - ResNet continuously achieves the highest accuracy across datasets, highlighting its capability to handle complex datasets successfully
  - VGG16 and AlexNet carry out comparably, with slight differences in smaller datasets.

- Ensemble Models:
  - Ensemble (Majority Vote) and Ensemble (Weighted Vote) methods continually outperform character models, emphasizing their robustness in combining a couple of predictions.
- Optimal Epochs:
  - Models obtain their fine overall performance at varying epochs depending on the dataset complexity and size.

5.2.1. Brain tumor image dataset with grayscale normalization and zoom (3096 images)

The dataset was brought from Kaggle; it consists of 3096 brain images. That contains normal and abnormal. As shown in Table 5, the method used in my model had the highest accuracy and the method used in (Al-Hadidi et al., 2020). Had the lowest accuracy. Furthermore, (Wang et al., 2024) Had the highest precision.

5.2.2. Brain MRI images for brain tumor detection (252 images)

The dataset was also brought from Kaggle; it contains 252 images of normal and abnormal brain images. As shown in Table 6, the methods used in (Asif et al., 2022) had the highest accuracy among all the methods, with InceptionResNetV2 having the highest accuracy, while (Bakr Siddiaue et al., 2020) had the lowest accuracy among all. Moreover, InceptionResNetV2 had 100% in precision.

Table 5. Performance comparison of brain tumor image dataset with grayscale normalization

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Multi-model of Xception, DenseNet-201, and EfficientNet-B3. (Santoso *et al*., 2024)	97.74	97.7	98.06	97.92
U-Net, CNN, VGG19 (Dhiman & Satpute, 2019)	93.7	-	93.1	-
CNN (Al-Hadidi *et al*., 2020)	75	-	-	-
AlexNet, VGG, ResNet (Wang *et al*., 2024)	96.94	99.32	-	-
ResNet50 and VGG16 (Abdullah *et al*., 2024)	92.6	-	-	-
Ensemble model (ResNet18, VGG16, and AlexNet)	98.5	98	98.5	98.5

Table 6. Performance comparison with the dataset of brain MRI images for brain tumor detection

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Deep convolutional neural network (DCNN) (Bakr Siddiaue *et al*., 2020)	96	93	100	97
VGG-19, VGG-16, ResNet50, Inception V3) (Krishnapriya & Karuna, 2023)	99.48	100	98.76	99.17
	99	100	98.18	99.08
	97.92	77.77	87.27	82.24
	81.25	53.84	63.25	58.16
Xception, NasNet Large, DenseNet121, InceptionResNetV2 (Asif *et al*., 2022)	99.67	99.68	99.68	99.68
	99.34	99.36	99.36	99.36
	99.00	98.72	99.36	99.04
	99.67	100.00	99.36	99.68
DCNN (Ramtekkar *et al*., 2023)	98.9	-	-	-
Detection (Phase 01: DBFS-EC Framework), Classification (Phase 02: HFF-BTC Framework) (Khan *et al*., 2022)	99.56	99.91	98.99	99.45
	99.20	99.13	99.06	99.09
Ensemble model (ResNet18, VGG16, and AlexNet)	92.15	89.5	0.90	90
Ensemble model (VGG-16, ResNet-50, and AlexNet) (Jader *et al*., 2024)	99.16	98.74	98.49	98.18

6. Conclusion

This research has proven the efficiency of employing transfer learning architectures, ResNet18, VGG16, and AlexNet, for the identification of brain tumors through MRI images. Of the three models, ResNet18 outperformed the others thanks to its outstanding accuracy of 98.66%. This enabled it to proficiently navigate complex data patterns as well as solve common issues experienced by convolutional networks, one of them being the vanishing gradient problem. The performance of VGG16 was equally promising, as the model recorded an accuracy of 97.8%, aided by its deep convolutional configuration that enabled the model to learn advanced features. Yet, AlexNet, a less complicated architecture, also performed reasonably well with an accuracy of 97.66%, demonstrating its applicability in less demanding environments. The technique of ensemble learning further brought to the surface the fact that improving the accuracy of individual models enhances the overall model accuracy to 98.33%, thereby supporting the case for combined models in improving diagnosis. This work not only supports the credibility of deep learning approaches to medical imaging but also the transformative nature of transfer learning in medicine, where the challenges posed by low datasets have been streamlined. In the future, studies might investigate the addition of further transfer learning frameworks, hyperparameter tuning, and multi-modal imaging data to enhance their diagnostic capabilities. This study's findings are valuable in improving the diagnosis and management of a brain tumor, which provides a basis for better medical solutions that are more accurate and efficient. Future work could discover innovative alternatives, fusion strategies, architectural modifications, or hybrid models to further enhance diagnostic accuracy and model robustness.

References

Abdullah, F., Jamil, A., Alazawi, E.M., & Hameed, A.A. (2024). Exploring deep learning-based approaches for brain tumor diagnosis from MRI images. In: *2024 IEEE 3rd International Conference on Computing and Machine Intelligence, ICMI 2024 - Proceedings*. https://doi.org/10.1109/ICMI60790.2024.10585851

Abdusalomov, A.B., Mukhiddinov, M., & Whangbo, T.K. (2023). Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers (Basel)*, 15(16), 4172. https://doi.org/10.3390/cancers15164172

Al-Hadidi, M.R., AlSaaidah, B., & Al-Gawagzeh, M.Y. (2020). Glioblastomas brain tumour segmentation based on convolutional neural networks. *International Journal of Electrical and Computer Engineering*, 10(5), 4738–4744. https://doi.org/10.11591/ijece.v10i5.pp4738-4744

Albalawi, E., Thakur, A., Dorai, D.R., Khan, S.B., Mahesh, T.R., Almusharraf, A., et al. (2024). Enhancing brain tumor classification in MRI scans with a multi-layer customized convolutional neural network approach. *Frontiers in Computational Neuroscience*, 18, 1-19. https://doi.org/10.3389/fncom.2024.1418546

Alfonse, A., & Salem, A.B. (2016). An automatic classification of brain tumors through MRI using support vector machine. *Egyptian Computer Science Journal*, 40(3), 1110–2586

Arumugam, M., Thiyagarajan, A., Adhi, L., & Alagar, S. (2024). Crossover smell agent optimized multilayer perceptron for precise brain tumor classification on MRI images. *Expert Systems with Applications,* 238, 121453. https://doi.org/10.1016/j.eswa.2023.121453

Asif, S., Yi, W., Ain, Q.U., Hou, J., Yi, T., & Si, J. (2022). Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from MR images. *IEEE Access*, 10, 34716–34730. https://doi.org/10.1109/ACCESS.2022.3153306

Bakr Siddiaue, M.A., Sakib, S., Rahman Khan,M.M., Tanzeem, A.K., Chowdhury, M., & Yasmin, N. (2020). Deep Convolutional Neural Networks Model-Based Brain Tumor Detection in Brain MRI Images. In: *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, p909–914. https://doi.org/10.1109/I-SMAC49090.2020.9243461

Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. (2023). Transfer learning for image classification using VGG19: Caltech-101 image data set. *Journal of Ambient Intelligence and Humanized Computing*, 14(4), 3609-3620. https://doi.org/10.1007/s12652-021-03488-z

Dhiman, A., & Satpute, P.B.S. (2019). Brain tumor segmentation in MRI images. *International Journal of Research in Advent Technology*, 7(9), 10-14. https://doi.org/10.32622/ijrat.78201916

El Kader, I. A., Xu, G., Shuai, Z., & Saminu, S. (2021) Brain tumor detection and classification by hybrid CNN-DWA model using MR images. *Current Medical Imaging,* 17(10), 1248-1255. https://doi.org/10.2174/157340561766621022411 3315

Fang, L., & Wang, X. (2022). Brain tumor segmentation based on the dual-path network of multi-modal MRI images. *Pattern Recognition,* 124, 108434. https://doi.org/10.1016/j.patcog.2021.108434

Gao, P., Shan, W., Guo, Y., Wang, Y., Sun, R., Cai, J., *et al.* (2022). Development and validation of a deep learning model for brain tumor diagnosis and classification using magnetic resonance imaging. *JAMA Network Open*, 5(8), e2225608. https://doi.org/10.1001/jamanetworkopen.2022.25608

Ghaffari, M., Sowmya, A., & Oliver, R. (2020). Automated brain tumor segmentation using multimodal brain scans: A survey based on models submitted to the BraTS 2012-2018 Challenges. *IEEE Reviews in Biomedical Engineering*, 13, 156-168. https://doi.org/10.1109/RBME.2019.2946868

Gül. M., & Kaya, Y. (2024). Comparing of brain tumor diagnosis with developed local binary patterns methods. *Neural Computing and Applications*, 36(13), 7545-7558. https://doi.org/10.1007/s00521-024-09476-6

Harnod, T., Lin, C.L., Sung, F.C., & Kao, C.H. (2014). An association between benzodiazepine use and occurrence of benign brain tumors. *Journal of the Neurological Sciences*, 336(1-2), 8-12. https://doi.org/10.1016/j.jns.2013.11.009

Jader, R.F., Kareem, S.W., & Awla, H.Q. (2024). Ensemble deep learning technique for detecting MRI brain tumor. *Applied Computational Intelligence and Soft Computing*, 2024, 6615468. https://doi.org/10.1155/2024/6615468

Khairandish, M.O., Sharma, M., Jain, V., Chatterjee, J.M., & Jhanjhi, N.Z. (2022). A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. *Irbm*, 43(4), 290–299. https://doi.org/10.1016/j.irbm.2021.06.003

Khan, A., Ghafoor, U., & Bhutta, M.R. (2022). A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI. *Sensors*, 22(7), 2726.

Krishnapriya, S., & Karuna, Y. (2023). Pre-trained deep learning models for brain MRI image classification. *Frontiers in Human Neuroscience*, 17, 1150120. https://doi.org/10.3389/fnhum.2023.1150120

Kumar, A., & Ma, Y. (2024). *Localization and Classification of Brain Tumor Using Multi-Layer Perceptron*. Vol. 1132. Singapore: Springer Nature. https://doi.org/10.1007/978-981-99-8853-2_6

Mahmud, M.I., Mamun, M., & Abdelgawad, A. (2023). A deep analysis of brain tumor detection from MR images using deep learning networks. *Algorithms*, 16(4), 1–19. https://doi.org/10.3390/a16040176

Mathivanan, S.K., Sonaimuthu, S., Murugesan, S., Rajadurai, H., Shivahare, B.D., & Shah, M.A. (2024). Employing deep learning and transfer learning for accurate brain tumor detection. *Scientific Reports*, 14(1), 1–15. https://doi.org/10.1038/s41598-024-57970-7

Mandle, A. K., Sahu, S. P., & Gupta, G. P. (2024). WSSOA: Whale social spider optimization algorithm for brain tumor classification using deep learning technique. *International Journal of Information Technology,* 16(1), 1–17. https://doi.org/10.1007/s41870-024-01782-5

Mostafa, A.M., Zakariah, M., & Aldakheel, E.A. (2023). Brain tumor segmentation using deep learning on MRI images. *Diagnostics,* 13(9), 1-22. https://doi.org/10.3390/diagnostics13091562

Nawaz, S.A., Khan, D.M., & Qadri, S. (2022). Brain tumor classification based on hybrid optimized multi-features analysis using magnetic resonance imaging dataset. *Applied Artificial Intelligence*, 36(1), 2031824. https://doi.org/10.1080/08839514.2022.2031824

Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., & Shoaib, M. (2020). A deep learning model based on concatenation approach for the diagnosis of brain tumor. *IEEE Access,* 8, 55135-55144. https://doi.org/10.1109/access.2020.2978629

Lauko, A., Lo, A., Ahluwalia, M.S., & Lathia, J.D. (2022). Cancer cell heterogeneity and plasticity in glioblastoma and brain tumors. *Seminars in Cancer Biology*, 82, 162–175. https://doi.org/10.1016/j.semcancer.2021.02.014

Leena, B., & Jayanthi, A. (2020). Brain tumor segmentation and classification via adaptive CLFAHE with hybrid classification. *International Journal of Imaging Systems and Technology*, 30(4), 874–898. https://doi.org/10.1002/ima.22420

Pacal, I., (2024). A novel Swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *International Journal of Machine Learning and Cybernetics*., 15(9), 3579-3597. https://doi.org/10.1007/s13042-024-02110-w

Praveen, G. B., & Agrawal, A. (2016). *Hybrid approach for brain tumor detection and classification in magnetic resonance images*. In: 2015 Communication, Control and Intelligent Systems, CCIS. IEEE, United States, p162-166. https:///doi.org/10.1109/ccintels.2015.7437900

Putzu, L., Piras, L., & Giacinto, G. (2020). Convolutional neural networks for relevance feedback in content based image retrieval:

A Content based image retrieval system that exploits convolutional neural networks both for feature extraction and for relevance feedback. *Multimedia Tools and Applications*, 79(37–38), 26995–27021. https://doi.org/10.1007/s11042-020-09292-9

Ramtekkar, P.K., Pandey, A., & Pawar, M.K. (2023). Accurate detection of brain tumor using optimized feature selection based on deep learning techniques. *Multimedia Tools and Applications*, 82(29), 1-31. https://doi.org/10.1007/s11042-023-15239-7

Ramzan, F., Khan, M.U.G., Rehmat, A., Iqbal, S., Saba, T., Rehman, A., *et al*. (2020). A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. *Journal of Medical Systems*, 44(2), 37. https://doi.org/10.1007/s10916-019-1475-2

Ranjbarzadeh, R., Kasgari, A.B., Ghoushchi, S.J., Anari, S., Naseri, M., & Bendechache, M. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Scientific Reports*, 11(1), 1-17. https://doi.org/10.1038/s41598-021-90428-8

Rasheed, Z., Ma, Y.K., Ullah, I., Al Shloul, T., Tufail, A.B., Ghadi, Y.Y., *et al*. (2023). Automated classification of brain tumors from magnetic resonance imaging using deep learning. *Brain Science*, 13(4), 602. https://doi.org/10.3390/brainsci13040602

Remzan, N., Tahiry, K., & Farchi, A. (2024). Advancing brain tumor classification accuracy through deep learning: harnessing radimagenet pre-trained convolutional neural networks, ensemble learning, and machine learning classifiers on MRI brain images. *Multimedia Tools and Applications*, 83, 0123456789. https://doi.org/10.1007/s11042-024-18780-1

Roopa, Y.M., Kumar, G.N., Harsha, Y.V.S., & Aditya, P.S. (2022). Detection of Brain Tumor Types Using Deep Learning. In: *Proceedings of the 2^{nd} International Conference on Artificial Intelligence and Smart Energy, ICAIS.* Vol. 6, p459–465. https://doi.org/10.1109/ICAIS53314.2022.9742731

Santoso, I.B., Supriyono, & Utama, S.N. (2024). Multi-model of convolutional neural networks for brain tumor classification in magnetic resonance imaging images. *International Journal of Intelligent Engineering and Systems*, 17(5), 741–758. https://doi.org/10.22266/ijies2024.1031.56

Siva Raja, P. M., & Rani, A.V. (2020). Brain tumor classification using a hybrid deep autoencoder with Bayesian fuzzy clustering-based segmentation approach. *Biocybernetics and Biomedical Engineering*, 40(1), 440–453. https://doi.org/10.1016/j.bbe.2020.01.006

Sulistyowati, T., Purwanto, P., Alzami, F., & Pramunendar, R.A. (2023). VGG16 deep learning architecture using imbalance data methods for the detection of apple leaf diseases. *Moneter: Jurnal Keuangan Dan Perbankan*, 11(1), 41–53. https://doi.org/10.32832/moneter.v11i1.57

Wang, J., He, L., & Zhou, X. (2024). Optimizing Inception-V3 for Brain Tumor Classification Using Hybrid Precision Training and Cosine Annealing Learning Rate. In: *2024 7^{th} International Conference on Advanced Algorithms and Control Engineering, ICAACE*, p528–532. https://doi.org/10.1109/ICAACE61206.2024.10548577

ZainEldin, H., Gamel, S.A., El-Kenawy, E.M., Alharbi, A.H., Khafaga, D.S., Ibrahim, A., Talaat, F.M. (2023). Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization. *Bioengineering,* 10(1), 1-19. https://doi.org/10.3390/bioengineering10010018

AUTHOR BIOGRAPHY

Shahab Wahhab Kareem. He received his BSc in Control and Computer Engineering from the University of Technology, Baghdad, in 2001, MSc in Software Engineering from Salahadeen University in 2009, and Ph.D. in Yasar University Izmir, Turkey in 2020. My research interests include Machine learning and BIG DATA. He is a lecturer at the Information System Eng. (ISE) Department.

Nayla Faiq Othman. She received her BSc in Information Systems Engineering from Erbil Polytechnic University in 2018 and an MSc student in Information Systems Engineering from Erbil Polytechnic University in 2024. Her research interests include deep learning and machine learning.

DOI: 10.6977/IJoSI.202506_9(3).0008
T. R. I. Putra, M. I. Fajri/Int. J. Systematic Innovation, 9(3), 124-141 (2025)

Sentiment of the largest state, first mover, and largest private banks digital performance in Indonesia: Strategic perspective

Teuku Roli Ilhamsyah Putra*, Muhammad Iqbal Fajri

Department of Management, Universitas Syiah Kuala, Banda Aceh, Aceh, Indonesia

*Corresponding author E-mail: teuku.roli@usk.ac.id

(Received January 13 2025; Final version received 13 March 2025; Accepted 07 April 2025)

Abstract

Digitalization plays an essential role in improving company performance, including banking. Understanding consumer sensitivity in digital banking applications is essential for strategic decisions. This research aims to analyze the user's sensitivity of the largest state, first mover, and largest private bank digital application in Indonesia, using the Naïve Bayes technique through Python. The data were taken from the Google Play Store, a software provider application for computer/laptop and mobile users, with a time range of 3 months from April 2023 until July 2023. The applications as the subject were XAA because it is owned by the largest state-owned bank, namely Bank XA, XBB because it belongs to the first mover for digital banks, and XCC, which is an application from the largest privately owned bank, namely Bank XC. The results reveal that most digital bank application users in Indonesia perceive that existing digital bank applications in Indonesia have yet to be able to meet their expectations. This is explained by the higher average negative value of their feedback answers than the existing positive value. Furthermore, this conclusion was revealed from the finding that XAA and XCC, which still had a positive score, had a higher negative score. Meanwhile, the XBB application, which is a first mover, was found to have a positive value higher than a negative one. We clearly compare the three applications divided into positive and negative categories and discuss the existing negative comments using a Digital Business Capabilities perspective.

Keywords: Digital Banking, Digital Business Capabilities, Sentiment Analysis, Strategic Perspective

1. Introduction

Along with the world's digital transformation, which is developing rapidly, organizations hope to improve their performance by using digital work patterns, especially in companies such as banks, which must operate with fast management and service processes. Digital banking is often seen as a transformation from conventional to online banking. In terms of service performance, changes in people's consumption patterns toward digital have greatly encouraged banks to accelerate the transformation process toward digital banking. Digital transactions worldwide from 2017–2021 grew by 118%, from USD 3.09 trillion in 2017 to USD 6.75 trillion in 2021 (Statista, 2021). Digital transformation has been shown to improve the performance of banks mainly through cost reduction and broader customer reach (Begenau et al., 2018; Chiu & Koeppl, 2019; Fuster et al., 2019; Zhu, 2019). Lee et al. (2021) examined a case within the financial sector as it undergoes paradigm shifts driven by technological advancements, focusing specifically on Deutsche Bank's competitive actions in the rapidly digitizing European banking landscape. The study concluded that Deutsche Bank's innovative initiatives are essential for achieving successful performance in a hyper-competitive environment. Xie & Wang (2023) established an index system to evaluate the digital transformation of banks across three key dimensions: strategic transformation, business transformation, and management transformation. By analyzing the data from Chinese commercial banks, they assessed the progress made in digital transformation within these institutions. Their empirical findings suggest that digital transformation not only improves bank performance but also alleviates the negative impacts of new technological entrants and supports the gradual discontinuation of offline channels.

DOI: 10.6977/IJoSI.202506_9(3).0008
T. R. I. Putra, M. I. Fajri/Int. J. Systematic Innovation, 9(3), 124-141 (2025)

Fig. 1 shows the trend of Internet and mobile banking transactions in Thailand has grown rapidly in the last 7 years, beating neighboring countries such as Malaysia, Indonesia, and Vietnam. Internet banking is a banking service available through websites or Internet sites. Meanwhile, mobile banking is a banking service available through a smartphone application. Although the way to access them differs, both services require an internet connection and can be used for digital banking transactions, such as bill payments, money transfers, balance checks, etc. In 2015, Thailand's Internet/mobile banking transactions were only 7,786 per 1,000 residents. However, in 2021, the volume increased to 238,245 transactions per 1,000 residents. Cumulatively, the volume of Internet and mobile banking transactions per 1,000 residents in Thailand grew by around 2,959% during 2015–2021. The trend of rapid growth is also seen in Malaysia and Vietnam. During the 2015–2021 period, the volume of internet/mobile banking transactions per 1,000 residents in Malaysia grew by 539%, and in Vietnam, it grew by 1,754%. Meanwhile, the growth rate in Indonesia was only 106% in the same period. As a result, the volume of Internet/mobile banking transactions per 1,000 residents in Indonesia in 2021 will be lower than Malaysia and Thailand and will almost be overtaken by Vietnam, as shown in Fig. 1 (Databoks, 2022).

In Indonesia, the growth of digital transactions has experienced a remarkable acceleration, achieving an impressive increase of 1,556 percent from 2017 to 2020. In 2021, electronic money transactions reached IDR 786.35 trillion. This figure reflects a rise of IDR 281.39 trillion (55.73%) compared to the previous year's total of IDR 504.96 trillion (Detik, 2022). Several banks are very familiar in various circles of society. This digital bank is so loved by millennials in Indonesia today because it is considered very practical and efficient. Users can make transactions and activate accounts online. They can have an account with a device, such as a smartphone, computer, or laptop, connected to the Internet to access digital banking (Fortune, 2022). Recent developments indicate that digital transformation affects commercial banks in two primary ways. The first involves outside-bank digital transformation, which refers to the competitive pressures stemming from non-banking institutions that offer similar digital services. This phenomenon not only intensifies competition but also facilitates technological spillover from fintech companies into banking institutions. A growing body of the literature explores the impact of this outside digital transformation on banks (Guo & Shen, 2016). Second, inside-bank digital transformation pertains to banks' adoption of advanced modern technologies. These technologies encompass artificial intelligence, blockchain, big data, cloud computing, and more (Khattak et al., 2023).

Fig. 2 shows Indonesia's domestic digital banking transactions value reached IDR 4,264.8 trillion, almost IDR 4.3 quadrillion. This includes various transactions like Internet and mobile banking, as defined by the Indonesian Financial Services Authority. According to the regulation issued by the institution, digital banking services are defined as banking services delivered through electronic media, optimized by leveraging customer data. In April 2023, the value of digital banking transactions in Indonesia declined by 11.8% from March 2023 and was 20.1% lower compared to April 2022. However, when we compare it to 5 years prior, the value of digital banking transactions in April 2023 represents a remarkable growth of 158% relative to April 2018. Despite experiencing monthly fluctuations, the long-term trend indicates a strengthening usage of digital banking services in Indonesia, as depicted in Fig. 2 (Databoks, 2023). Banks that invest significantly in technology are viewed as being more digitally transformed. A study by Do et al. (2022) indicates that digital transformation enhances the performance of banks. Theiri & Alareeni (2023) propose that digital transformation can serve as an innovative strategy during crises, such as the COVID-19 pandemic. Given that technology is seen as a cost-effective investment, a positive relationship

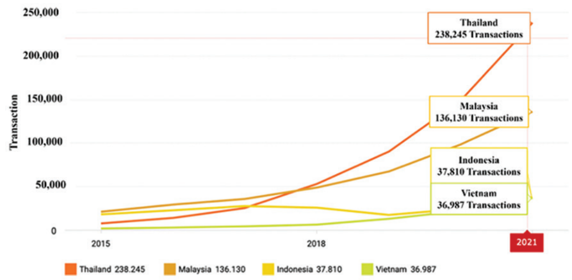


Fig. 1. Internet and mobile banking transaction volume per 1,000 residents in ASEAN countries (2015–2021)
Source: Databoks (2022)

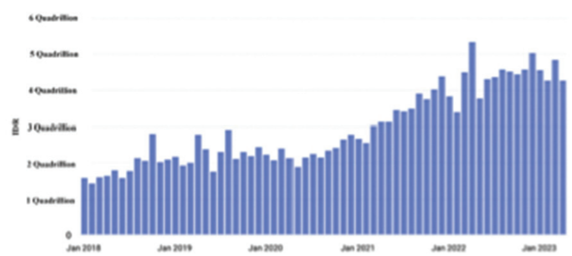


Fig. 2. Indonesia digital banking transactions value per month (January 2018–April 2023)
Source: Databoks (2023)

is anticipated, with digital transformation leading to improved bank stability. Utilizing microdata from Brazil, Silva et al. (2023) discovered that branches of more digitalized banks were less affected by local borrowers' circumstances and were able to broaden their clientele. These branches provided credit to borrowers in remote areas that were less impacted by COVID-19, thereby positioning themselves more favorably compared to their less digitalized counterparts.

Currently, in the aftermath of the COVID-19 pandemic, digital banks in Indonesia are experiencing significant growth, promising enhanced service quality that transcends physical and temporal limitations. The number of banks is on the rise, each offering a variety of digitally focused services. It is important to distinguish between digital banks and conventional banks that merely provide digital services such as mobile and internet banking. Digital banks uniquely enable customers to conduct all banking activities from opening accounts and making transfers to depositing and closing accounts exclusively through smartphones or electronic devices, eliminating the need for physical visits to a bank branch. In addition, a key distinction is that digital banks typically operate without physical branches (aside from a head office) or may have a limited number of physical locations (OJK, 2021a). Conventional banks that offer digital services often struggle to provide all of their offerings online. In addition, these banks are typically associated with numerous physical branch locations. However, there are at least two significant potential benefits that can emerge from the digital transformation of banks. First, it can expand access to banking services. Second, it can enhance the competitiveness of Indonesian banking. Digital banking will facilitate easier public access and improve banking efficiency, ultimately fostering greater economic activity (OJK, 2021b).

The driving factors of digital development can be categorized into three main aspects: digital opportunities, behavior, and transactions. Digital opportunities encompass the influence of demographics, economic conditions, financial potential, internet penetration rates, and the expanding consumer base. Digital behavior pertains to the ownership of devices and the utilization of mobile applications. Digital transactions refer to online trading (e-commerce), digital banking, and electronic money transfers. While the banking industry can capitalize on numerous opportunities presented by digital transformation, it also faces several challenges that require careful consideration. These challenges include the protection of personal data and the risk of data breaches, the possibility that technology investments may not align with business strategies, the misuse of artificial intelligence, cybersecurity threats, outsourcing risks, the necessity to foster readiness within digitally oriented institutional frameworks, ensuring financial inclusion for individuals with disabilities, enhancing digital financial literacy, rapid deployment of information technology (IT) infrastructure, and the need for an appropriate regulatory framework to support (OJK, 2021b). Furthermore, Khattak et al. (2023) discovered that banks experience a positive impact from digital transformation on stability, especially when they maintain a lower level of diversification. As banks diversify further, the advantages of digital transformation start to diminish, potentially turning negative at higher levels of diversification. Their research suggests that increased diversification heightens bank risk in the context of digital transformation, whereas banks with lower levels of diversification benefit from digital transformation, contributing to greater stability.

2. Literature Review

2.1. Digital Banking

Numerous studies indicate a need for further research into strategic decisions concerning the use of digital banking products, particularly in comparison to the existing studies on the service quality of these digital offerings. Hoehle et al. (2012) explored digital banking utilization and noted limitations in previous research due to the phenomena studied and the methodologies employed. Waite & Harrison (2015) advocate for alternative research perspectives to examine digital banking, while Piyathasanan et al. (2015) point out the scarcity of experiential guidelines in this domain. It is essential for banks to leverage digital banking to enhance customer benefits (Patsiotis et al., 2012). The scope of digital banking encompasses electronic banking services accessed via digital devices, including t-banking, e-banking, mobile banking, contactless cards (such as tap-and-go), ATMs, and point-of-sale systems, while excluding platforms like PayPal that serve as intermediaries (e.g., eBay) for banking transactions. These studies imply that banks can enhance customer experience and overall performance by utilizing digital banking channels. Therefore, the impact of digital banking requires clarification (Mbama et al., 2018). Some ways are by detecting consumer satisfaction based on what they feel as users of digital applications.

2.2. Performance versus Perceived Performance

Digital banking has empowered banks to enhance customer satisfaction by providing instant services through various distribution channels (Tam & Oliveira, 2017). As customer expectations continue to rise, firms are compelled to adopt a more customer-centric approach and invest in delivering quality services, improving performance (Pekovic & Rolland, 2016), and strengthening their brand image (Fritz et al., 2017). The study by Julien & Tsoni (2013) focuses on the perceptions of banking staff and customers, aiming to identify gaps in perception and enhance the co-production of service quality. In recent years, mobile banking (m-banking) has emerged as a significant strategic shift for banks (Tam & Oliveira, 2017). Perceived performance reflects customers' views on service attributes, benefits, and overall outcomes (Rahi et al., 2023; Venkatesh et al., 2003). So far, studies have only measured the quality of digital banking product services by detecting user satisfaction based on their feelings. As stated by Eren (2021), satisfaction serves as the basis for user behavior and can be evaluated by examining the confirmation or disconfirmation of user expectations alongside the performance of the service. Customer expectation refers to the perceptions customers have about what they anticipate receiving in exchange for utilizing a service (Franque et al., 2021; Rahi et al., 2023; Venkatesh et al., 2003). However, this method must, of course, be seen from two sides. After dissecting the customer's perspective, it is necessary to study the strategic perspective to see possible future performance.

Additionally, research into user responses to technology services like artificial intelligence (AI) indicates that users experience either confirmation or disconfirmation of their expectations when using AI-enabled banking. When user expectations align with their needs and enhance customer satisfaction, the experience is positive. However, if expectations are not met, it can lead to negative attitudes toward AI-based banking (Azizi et al., 2021; Bhattacherjee, 2001; Bastari et al., 2020). Positive confirmation enhances perceived performance, which in turn leads to greater user satisfaction and a stronger acceptance of artificial intelligence (Al-Okaily, 2024; Brill et al., 2019). Previous research has employed expectation confirmation theory as a framework to explore how users' expectations of technology services are confirmed or disconfirmed. This theory has been instrumental in examining the relationship between users' satisfaction and the perceived usefulness of these services, shedding light on how initial anticipations influence overall user experience and acceptance of technology (Bhattacherjee, 2001; Brill et al., 2019; Eren, 2021; Kim, 2018; Kosiba et al., 2020).

2.3. Digital Business Capability (DBC)

DBC is an essential aspect for organizations that aim to thrive in the competitive landscape shaped by rapid technological advancements (Stocker et al., 2024). This capability empowers companies to quickly and effectively respond to shifting market dynamics and the diverse demands of today's consumers. By harnessing DBC, organizations can streamline their operations, optimizing workflows to enhance efficiency and productivity while simultaneously fostering a culture of innovation. This ultimately leads to the development and delivery of cutting-edge products and services that resonate with their target audience (Hall et al., 2020). Moreover, the process of digital transformation not only paves the way for reduced operational costs and increased revenue streams but also plays a pivotal role in elevating the overall customer experience. It cultivates an environment of collaboration within teams, enabling them to work together more seamlessly to meet shared goals (Balakrishnan & Das, 2020; Meena et al., 2024). Wielgos et al. (2021) found that DBC is more valuable for business-to-consumer than for business-to-business firms and revealed that DBC is more valuable under high than low levels of structural flux. Exciting insights are presented by Khattak et al. (2023), highlighting that investments in technology and increased diversification can render banks riskier and more vulnerable. Li et al. (2018) examined small and medium enterprises that lack adequate capabilities and resources, concluding that managerial capabilities are the primary catalyst for digital transformation. Vial (2019) underscored the significance of IT capabilities, arguing that they enhance firm performance, with digital transformation initiatives serving as a mediating factor.

DBC is a multifaceted construct that comprises three interrelated components: (i) Digital strategy (DS), (ii) digital integration, and (iii) digital control, as articulated by various scholars (Ranjan, 2024; Sorescu & Schreier, 2021; Wielgos et al., 2021; Xia et al., 2024). The first component, DS, plays a pivotal role in defining and articulating distinct objectives and goals aimed at harnessing the power of digital technologies to propel business success. This strategic approach not only involves a thorough comprehension of the ever-changing market dynamics but also requires a deep awareness of customer needs, ensuring harmony with the organization's broader business strategies (Bresciani et al., 2021; Henderson & Venkatraman, 1999; Gobble, 2018; Guinan et al., 2019; Warner & Wäger, 2019; Yeow et al., 2018; Zaki, 2019). Moreover, this process necessitates the effective implementation of various digital technologies across different layers of an organization, accompanied by a commitment to ongoing monitoring and adaptation of these strategies to sustain a competitive edge within the marketplace (Warner & Wäger, 2019). The significance of a well-defined DS cannot be overstated; it is essential not merely for maintaining competitiveness but also for fostering long-term growth and innovation (Tardieu et al., 2020). Importantly, DS is not a singular, isolated endeavor but rather a continuous process that demands regular evaluation and refined adjustments to respond to evolving market conditions and technological advancements (Chirumalla, 2021; Grover & Malhotra, 1997; Guinan et al., 2019; Szelągowski & Berniak-Woźny, 2022).

Digital integration emphasizes the comprehensive implementation of various strategies that permeate every facet of an organization, encompassing areas such as marketing, operations, and beyond (Berman, 2012; Chaffey & Smith, 2022; Gunasekaran & Ngai, 2004; Holopainen et al., 2022; Leeflang et al., 2014; Yeow et al., 2018). In today's dynamic and constantly evolving business landscape, it is imperative for organizations to cultivate agility and responsiveness. By seamlessly weaving digital technologies into all operational segments, businesses can achieve greater efficiency, streamline workflows, and significantly enhance the overall customer experience (Zaki, 2019). This holistic approach not only positions companies favorably within their competitive landscape but also empowers them to swiftly adapt to emerging market trends and consumer demands (Knudsen et al., 2021).

Digital control encompasses the processes of closely monitoring and adjusting various strategic initiatives to ensure their effectiveness in driving desired outcomes (Liu et al., 2009; Nylén & Holmström, 2015; Proia et al., 2022). In the face of today's rapidly changing business landscape, companies must cultivate mastery over three essential capabilities to flourish in the digital era. By consistently analyzing the data and assessing performance metrics, organizations are empowered to make well-informed decisions and adapt their strategies as circumstances evolve (Bajwa et al., 2017; Kirtley & O'Mahony, 2023; Kokina et al., 2017). This proactive methodology toward digital control allows businesses to maintain agility and swiftly respond to fluctuating market conditions, ensuring they remain competitive and relevant (Berman & Marshall, 2014; Giacosa et al., 2022; Gunasekaran & Yusuf, 2002; Yeow et al., 2018).

As technology evolves at an astonishing rate, companies must prioritize digital integration and control to stay competitive and fulfill the changing demands of their customers (Berman & Marshall, 2014; Giacosa et al., 2022; Gunasekaran & Yusuf, 2002; Yeow et al., 2018). However, having access to data and metrics alone does not guarantee success. Even organizations that make substantial investments in data analysis and performance tracking can struggle to effectively utilize these insights, resulting in poor decision-making and ultimately impeding their ability to thrive in the digital landscape (Dahlbom et al., 2020). Success is not solely dependent on having the right technology; it equally requires cultivating the expertise and strategic vision needed to leverage these tools effectively (Henderson & Venkatraman, 1999; Luftman et al., 1993; Quinn et al., 1999. Without the capability to effectively harness insights from data analysis, organizations risk falling behind in an increasingly competitive market (Ranjan & Foropon, 2021; Sharma et al., 2014; Wang et al., 2018).

2.4. Research Questions (RQ)

The discussions above reveal the importance of looking at the user's perspective as the basis for a strategy to improve banking digitalization performance. The author tries to comprehensively analyze this perspective in this case. The RQ are:

1. RQ1: What are the sentiment analysis results for the three banking applications analyzed in Indonesia?
2. RQ2: How can banks improve their digital business capabilities to satisfy customers?

3. Method, Data, and Analysis

This research used the subject of Digital Banks in Indonesia in 2023. The selected digital banks are three digital banks from different classifications, and the basis is the application, which is currently still developing, namely:

(i) XAA is a digital banking application from the largest state-owned bank in Indonesia
(ii) XBB is a digital banking application from Bank XB and is a first-mover application in digital banking
(iii) XCC is a digital banking application from the largest privately owned bank in Indonesia.

The respondents were Indonesian people who commented on the Google Play Store. Data is pulled from Google Play Store using a Python application with the Naïve Bayes technique. Naïve Bayes is a classification algorithm based on Bayes' theorem, assuming conditional independence between features, which is why it is considered "naïve" (Allam et al., 2025). Despite rarely meeting this assumption in real data, it has proven effective in tasks like document categorization, spam filtering, and sentiment analysis (Mendhakar & Tilmatine, 2023). This research employs a Multinomial Naïve Bayes, which models word frequency per class and works well with word counts (Ridho et al., 2022; Zhang, 2004). The text is represented using the Bag-of-Words model, where documents are collections of words with their frequency. Features can include unigrams or n-grams, though overly long n-grams may reduce accuracy (Nandwani & Verma, 2021). To enhance feature quality, weighting schemes like term frequency-inverse document frequency are used to emphasize unique words within documents, aiding Naïve Bayes in identifying important keywords (Allam et al., 2025). In this research, data are collected from the Google Play Store using the google_play_scraper library in Python. User ratings classify comments for digital banking applications: ratings of 1 and 2 indicate negative sentiments, while ratings of 4 and 5 represent positive sentiments; ratings of 3 are considered neutral and ignored. Preprocessing is conducted to enhance data quality, involving steps such as: Handling missing values by removing incomplete entries, tokenization for breaking text into manageable word units, removing stop-words that carry little significance in sentiment analysis, and lemmatization to normalize words to their base forms. A Multinomial Naïve Bayes model is then employed for sentiment classification, trained on the processed data, and evaluated using metrics like f1-score, precision, and recall. The model was built using a dataset split of 80:20 for training and testing. The implementation leverages the "MultinomialNB" classifier from "sklearn.naive_bayes," with default hyperparameters maintained, including an alpha of 1.0 for Laplace smoothing. The model predicts sentiment based on the review text, assessing user perception of the digital banking application.

Furthermore, this study's first assessment (RQ1) was carried out using sentiment analysis, which looks at public opinion about digital banks. The data are in the text form, which is an opinion expressed in Indonesian. The sample taken is data for 3 months, from April 2023 until July 2023; the data found are a total of 3000 data, which for XAA is 1,000 Data, XBB 1,000 Data, and XCC 1,000 Data. The analysis will divide user perspectives into positive and negative categories for the three applications studied. The result is aggregate data for comparative analysis. After answering RQ1, the second assessment of this study (RQ2) is conducted through a DBC perspective (Ranjan, 2024; Sorescu & Schreier, 2021; Wielgos et al., 2021; Yi et al., 2024). Discussion on the DBC perspective is deemed necessary to fulfill the perspective within existing application capabilities. A slight difference from the DBC survey, which uses indicators for surveying, DBC in this study was studied based on negative comments from application users. Positive reviews have their aspects, which are not discussed in this study. We deem the selection of negative aspects necessary because the comments made by users are direct and occur spontaneously of their own accord, so they are deemed necessary to study. In contrast to particular survey results where the scale has been determined, and there is a need to accommodate the surveyor's request to fill out a questionnaire or answer interview questions, spontaneous results in the Google Play Store application are seen as the user's desire to provide comments naturally (their desires and needs for themselves) so that the response to negative comments is also seen as very honest and has a critical weight to respond to.

The existing negative comments are related to DBC descriptively, namely the researcher's considerations through indicators, namely DS, Integration (DI), and Control (DC). For the DS, the perspective used is the extent to which the DS creates new opportunities to add value for our customers (DS1), add value for the firm and its partners (DS2), and continuously deliver innovations in digital products and services (DS3). For digital integration, the perspective used is the extent to which the firm is becoming more connected online with customers, suppliers, and partners (DI1). Digital business transformation is becoming integral and interconnected across all areas of the company (DI2), digital skills are transforming business processes within firms (DI3), business processes across the value chain are becoming more digitally interconnected (DI4), and, the firm is becoming more interconnected through digital platforms (DI5). Meanwhile, for Digital Control, the perspective used is the extent to which the firm has clear specifications for implementing digital business transformation (DC1), the firm consistently tracks its digital business transformation progress (DC2), and the firm regularly analyzes performance metrics to evaluate its digital business transformation (DC3). The comment data generated from Google Play Store are summarized and analyzed descriptively so that comments that tend to be the same will be combined into one (representative) so that it appears there are only a few comments. These results become material for this study's findings based on the DBC.

4. Result and Discussion

4.1. Data Description

A data description is needed to see the statistical value. Fig. 3 shows the result for XAA.

The XAA data obtained are 1,000, with a minimum value of 1.0, namely, a polarity value of strong negative and a max value of 5.0 with a strong positive polarity. However, if we look at the mean owned by XAA, which is in the range of 2.06, it can be interpreted that the comments owned by XAA are in negative territory.

Next, Fig. 4 shows the data description for XBB.

XBB has a total of 1,000 data, with a minimum value of 1.0, which indicates the polarity or intensity of the most negative response is a strong negative, and the max value is 5.0, which indicates a strong positive polarity of sentiment, with a mean of 3.62 or rounding off 4 concludes that the comments that are owned by genius are in positive territory.

DOI: 10.6977/IJoSI.202506_9(3).0008
T. R. I. Putra, M. I. Fajri/Int. J. Systematic Innovation, 9(3), 124-141 (2025)

Furthermore, Fig. 5 shows the data description for XCC.

XCC has 1000 data, with min is 1.0, so the polarity is negative, and max is 5.0, so the polarity is strongly positive. XCC's mean value is 2.6, which is in the negative to neutral quadrant.

	score
count	1000.000000
mean	2.064000
std	1.420536
min	1.000000
25%	1.000000
50%	1.000000
75%	3.000000
max	5.000000

Fig. 3. XAA data description

	score
count	1000.000000
mean	3.627000
std	1.795298
min	1.000000
25%	1.000000
50%	5.000000
75%	5.000000
max	5.000000

Fig. 4. XBB data description

	score
count	1000.00000
mean	2.66700
std	1.68373
min	1.00000
25%	1.00000
50%	2.00000
75%	5.00000
max	5.00000

Fig. 5. XCC data description

4.2. Model Validity

The Naïve Bayes model applied in this study shows fairly high performance, with the following accuracy rates: XAA: 91%, XBB: 92%, and XCC: 88%. To ensure a more comprehensive evaluation, the model was analyzed using the f1-score metric, which is the harmonic mean between precision and recall. The evaluation results show that the model performs well in distinguishing between positive and negative sentiments:

- XAA: Negative (0.95) and positive (0.69)
- XBB: Positive (0.94) and negative (0.88)
- XCC: Negative (0.89) and positive (0.85).

Data are considered valid if it achieves an f1-Score >0.7 (Allam et al., 2025; Nandwani & Verma, 2021). In the data collected for this study, most classes exhibit an f1-score above this threshold. However, the positive class XAA has an F1 Score of <0.01 (0.69). Despite this, it is still regarded as having substantial validity (Allam et al., 2025). Therefore, the model can be classified as robust and capable of effectively explaining the findings.

4.3. Multinomial Naïve Bayes

The Naïve Bayes algorithm is a machine learning algorithm for classification based on the Bayesian theorem. This algorithm studies the probability of an object with specific characteristics belonging to a particular class or group. This method is often used to solve opinion-mining problems. In this sentiment analysis research, the approach used is Multinomial Naïve Bayes, a model that calculates the frequency of occurrence of words from a document (Zhang, 2004). Hence, the results obtained per application are as follows.

The results for the XAA application state the same thing as the results in Table 1: the majority of independent reviews are in the negative category. The value that becomes the benchmark is the f1 score because the f1 score is the harmonic average value of precision and recall. The picture above reveals that a negative f1 score of 0.95 is greater than a positive f1 score (0.69) with an accuracy rate of 91%.

The results obtained from the multinomial Naïve Bayes are the same as those in Table 2, which states that genius is in the positive category. If we look at the f1 score obtained, the positive result for genius (0.94) is greater than the negative result for genius (0.88), with an accuracy of 92%. The value that becomes the benchmark is the f1 score.

Table 1. XAA result

Item Analyzed	Prec.	Recall	f1	Support
Negative	0.90	1.00	0.95	143
Positive	1.00	0.53	0.69	34
Accuracy			0.91	177
Macro avg	0.95	0.76	0.82	177
Weighted avg	0.92	0.91	0.90	177

Table 2. XBB result

Item Analyzed	Prec.	Recall	F1	Support
Negative	0.80	0.98	0.88	57
Positive	0.99	0.90	0.94	135
Accuracy			0.92	192
Macro avg	0.90	0.94	0.91	192
Weighted avg	0.93	0.92	0.92	192

The data in Table 3 states that the f1 score obtained is a negative value (0.89), slightly higher than a positive value (0.85), with an accuracy of 88%. The multinomial results have the same results as the XCC data description results, where these results state that there is very little difference between positive and negative comments. So, the comparison of negative feedback between XAA, XBB, and XCC is more clearly presented in Fig. 6.

Meanwhile, the comparison of the positive feedback is shown in Fig. 7.

RQ1: In general, it shows that of the three banks, the XAA application is in a strongly negative category, while XBB has a positive category. Moreover, if XCC is included even though it is positive, it reveals that XCC is still in the neutral to positive category. This confirms the results obtained in the previous data description, with XBB mean of 3.62 to 4, which has reviews that tend to be positive; XCC has a mean of 2.66 toward three (toward neutral). It explains that XCC's positive and negative reviews are slightly balanced, and the Mean XAA value of 2.06 states that the review owned by XAA is in the negative category. However, the average comparison of positive (0.83) and negative (0.91) reviews concludes that currently, the majority of users in Indonesia still view digital bank applications as needed to meet user expectations. Satisfaction is influenced by several factors, including expectation confirmation, perceived performance, trendiness, visual attractiveness, problem-solving capabilities, customization, and communication quality. These elements reveal a significant variance of $R^2 = 51.1\%$ in digital banking user satisfaction (Alnaser et al., 2023). The integration of artificial intelligence features such as trendiness, visual appeal, and problem-solving has enhanced the attractiveness, appeal, and innovativeness of banking operations. For instance

Table 3. XCC result

Item Analyzed	Prec.	Recall	F1	Support
Negative	0.90	0.88	0.89	104
Positive	0.84	0.86	0.85	72
Accuracy			0.88	176
Macro avg	0.87	0.87	0.87	176
Weighted avg	0.88	0.88	0.88	176

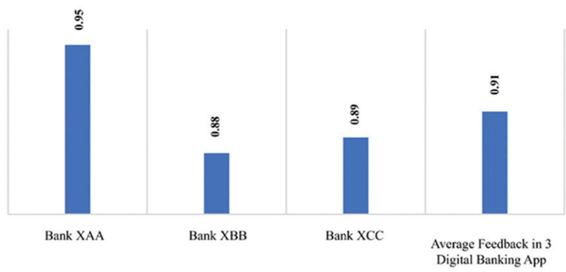


Fig. 6. Negative feedback comparison

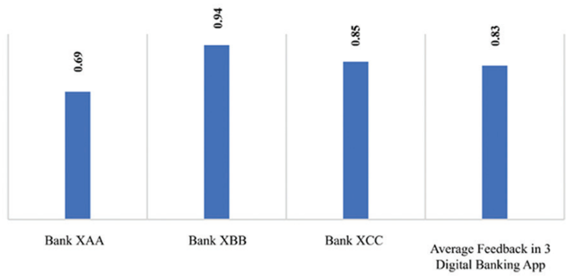


Fig. 7. Positive feedback comparison

(Chung et al., 2020) noted that customers tend to favor trendy services over traditional options. Additionally, there has been a significant shift in the business landscape, as the role of salespersons has diminished; customers now increasingly rely on online systems to enhance their lifestyles (Godey et al., 2016). Previous research has demonstrated that artificial intelligence-driven digital banking aligns well with customers' needs and contemporary lifestyles, ultimately leading to increased user satisfaction (Chung et al., 2020; Godey et al., 2016; Zolkepli & Kamarulzaman., 2015). Aside from trendiness, digital banking services should be attractive and appealing (Bhandari et al., 2019; Gupta et al., 2023; Ho et al., 2023).

4.4. Negative Comments

(1) XAA Perceived Performance

Analyzed data summarized several negative reviews. Of the 100 negative reviews, there were reviews as representations of the XAA application. The first problem that occurred with the XAA application was:

- Every time a transaction is made, the application is asked to update the clock automatically; this is a problem that disrupts banking activities; the majority of reviews give 1 star (worst) for the problem of updating automatic hours.
- When the transfer activity is needed, there are always problems with the network, and the most frequent complaint is the notification of "transfer failed, connection error, please check your connection and try again."
  - These two things are the main problems that application providers must address regarding digital solutions. From the DBC perspective, these two problems we consider lie in the dimensions of DS, Integration, and Control. The application provider has yet to be able to meet two indicators in the DS dimension. Namely, the provider's DS has not been able to create value for customers (DS1) and has yet to be fully able to continuously deliver innovations in digital products and services (DS3). For the Digital Integration dimension, what directly impacts these comments is the indicator that the firm is becoming more connected online with customers, suppliers, and partners (DI1). Meanwhile, this problem is related to the overall indicators in the Digital Control dimension; namely, the firm has clear specifications for implementing digital business transformation (DC1); the firm consistently tracks its digital business transformation progress (DC2), and the firm regularly analyzes performance metrics to evaluate its digital business transformation (DC3).

(2) XBB Perceived Performance

The result reveals that there were still some negative comments about the performance of the XBB application. The negative comments received on the XBB application tend to be in the direction of:

- Difficulty in registration: Almost all complained in video calls about data verification.
- Unlinking the device cannot be done via the application but via customer service via telephone.
- There is no cost transparency or detailed cost explanations by the genius team in the field, so there are many hidden fees that users do not know about
  - These three things are the main problems that application providers must address regarding digital solutions. From the DBC perspective, these problems lie in the dimensions of DS, integration, and control. The application provider has yet to be able to meet two indicators in the DS dimension; the provider's DS has not been able to create value for customers (DS1) and has not been able to continuously deliver innovations in digital products and services (DS3). Meanwhile, in the digital Integration dimension, the provider has yet to make the business processes across the value chain more digitally interconnected (DI4). So for the Control Dimension, this problem is related to the overall indicators; namely, the firm has clear specifications for implementing digital business transformation (DC1), the firm consistently tracks its digital business transformation progress (DC2), and the firm regularly analyzes performance metrics to evaluate its digital business transformation (DC3). These facts explain that in general existing business processes have not been able to optimally meet user needs, namely that the digital transformation has not answered the need for service flexibility and cost transparency.

(3) XCC Perceived Performance

The result proves that there are negative reviews that users often discuss. This review only focuses on two things, namely:

- Difficulty in "signing in" the application
- Difficulty in verifying personal data.
  - These two things are the main problems that application providers must address regarding digital solutions. From the DBC perspective, we consider that this problem lies in the three existing dimensions: DS, Integration, and Control. Application providers have yet to be able to meet two indicators in the DS dimension. Namely, the provider's DS has not been able to create value for customers (DS1) and has not been able to continuously deliver innovations in digital products and services (DS3). Meanwhile, in the digital Integration dimension, the provider has yet to make the business processes across the value chain more digitally interconnected (DI4). So for the Control Dimension, this problem is related to the overall indicators; namely, the firm has clear specifications for implementing digital business transformation (DC1), the firm consistently tracks its digital business transformation progress (DC2), and the firm regularly

analyzes performance metrics to evaluate its digital business transformation (DC3). This means that the company still needs to implement these control indicators fully.

4.5. Discussion

The results show that (RQ1) of the three banks, the XAA application is strongly negative, while XBB has a positive category. Moreover, if XCC is included even though it is positive, it reveals that XCC is still in the neutral to positive category. XBB reviews tend to be positive; XCC's positive and negative reviews are slightly balanced, and the XAA reviews tend to be negative. However, the average comparison of positive (0.83) and negative (0.91) reviews reveals that the majority of users in Indonesia still view digital bank applications as not being able to meet user expectations. Theoretically, measurement via DBC is relevant to the development of banking digitalization in Indonesia. The negative reviews show similarities between the three applications regarding DBC indicators, namely DS1, DS3, DI4, DC1, DC2, and DC3. This finding explains that negative comments on Google Playstore lead to the same needs for these indicators, while other indicators on DBC receive a portion outside of negative comments. Thus, the results obtained that still require encouragement by banking digitalization practitioners to transform their digital banking applications are:

1. Create the value of "convenience and fast service" (DS1), and continue to deliver digital products and service innovations (DS3). These factors can be enhanced by leveraging the insights gained from data analysis, and organizations can customize their digital services to better meet the needs of their target audience, ultimately leading to increased customer satisfaction and loyalty. In today's digital age, the ability to effectively harness data is essential for organizations to not only survive but thrive in a competitive market (Amankwah-Amoah & Adomako, 2019; Grover et al., 2018; Oppong et al., 2005; Sultana et al., 2022). Companies that effectively implement digital strategies are able to adapt to changing market conditions and customer preferences, ultimately leading to increased revenue and market share (Katsikeas et al., 2020; Leeflang et al., 2014; Zaki, 2019). By investing in digital technologies and staying ahead of the curve, businesses can secure their position as industry leaders and continue to drive growth in the long term.
2. Become more connected online with customers, suppliers, and partners (DI1). The disrupted distribution (usage) value chain for customers must be repaired. It is in order to adapt and stay relevant in an ever-changing landscape. This corrective action can be achieved by continuously analyzing and interpreting data to understand customer behavior and preferences, allowing organizations to make informed decisions and deliver personalized experiences (Chakravarty, 2014; Cordon et al., 2016; Roden et al., 2017). By addressing these disruptions and staying ahead of the curve, organizations can ensure long-term success and sustainable growth in the digital marketplace. Ultimately, successful digital integration can lead to increased revenue, growth, and long-term sustainability in today's competitive landscape (Knudsen et al., 2021). It can lead to increased revenue, growth, and long-term sustainability in today's competitive landscape, giving companies a significant advantage over their competitors (Bereznoy, 2019).
3. Concrete specifications for implementing digital business transformation (DC1), regularly track its digital business transformation progress (DC2), and regularly analyze performance metrics to evaluate its digital business transformation (DC3). These factors require the right staff to control the transformations. This will help ensure that organizations can adapt quickly to change and remain competitive (Holbeche, 2007; Prastacos et al., 2002). Additionally, implementing regular training programs for employees can help them stay up-to-date on the latest technologies and trends in the industry (Kamble et al., 2018; Tavitiyaman et al., 2022; Vashishth et al., 2024). By fostering a culture of innovation and adaptability, organizations can create a dynamic and agile workforce that is able to respond effectively to any disruptions in the market (Ajgaonkar et al., 2022; Dyer & Shafer, 2003; Ulrich & Yeung, 2019. With the right tools, resources, and team in place, organizations can position themselves as leaders in their respective industries and thrive in the digital age. It is also crucial for organizations to invest in ongoing training and development programs to keep their workforce up-to-date with the latest advancements in technology and industry best practices (Dachner et al., 2021; Gope et al., 2018; Sonnentag et al., 2004). Additionally, fostering a culture of collaboration and open communication can help employees feel empowered to share innovative ideas and solutions (Ahsan, 2024; Barczak et al., 2010; Çakar & Ertürk, 2010; Islam et al., 2024; Lemon & Sahota, 2004; Jaskyte, 2004); Lin, 2007; Waseel et al., 2023). With a strong foundation of adaptability, innovation, and a skilled workforce,

Table 4. Digital improvement solution matrix

Commented Problem / Indicator Related	XAA • Every time a transaction is made, the application is asked to update the clock automatically; this is a problem that disrupts banking activities; the majority of reviews give 1-star (worst) for the problem of updating automatic hours. • The “transfer failed, connection error, please check your connection and try again”.	XBB • Difficulty in registration: Almost all complained in video calls about data verification. • Unlinking the device cannot be done via the application but via customer service via telephone. • There is no cost transparency or detailed cost explanations, many hidden fees that users do not know about	XCC • Difficulty in “signing in” the application • Difficulty in verifying personal data.
The firm digital strategy creates value for customers (DS1)	Creating the value of “convenience and fast service” with Total Quality Management (TQM) related to the efficiency of updating clocks, and preventing transfer failures due to the network. This value can open up other new opportunities to provide greater value to customers.	Creating the value of “convenience, fast service, and accurate information dissemination” with TQM related to the registration process, efficient device unlinking process for customers, and cost transparency. This value can open up other new opportunities to provide greater value to customers.	Creating the value of “convenience and fast service,” with TQM related to the ease of the “sign in” process, and the efficiency of the personal data verification process. This value can open up other opportunities to provide greater customer value.
The firm continuously delivers innovations in digital products and services (DS3)	Continue to provide digital services by expanding/enlarging the benefits and appeal and providing appropriate usage methods for customers.	Continue to provide digital services by expanding/ enlarging the benefits and appeal and providing appropriate usage methods for customers.	Continue to provide digital services by expanding/enlarging the benefits and appeal and providing appropriate usage methods for customers.
The firm business processes across the value chain are becoming more digitally interconnected (DI4)	The disrupted distribution (usage) value chain for customers must be repaired, namely the system clock update system and transfer failures due to network disruptions, so that it does not disrupt components others in the value chain	The disrupted distribution (usage) value chain for customers must be repaired, namely registration barriers, inefficient processes for unlinking the device, and decreased customer trust due to hidden costs.	The disrupted distribution (usage) value chain for customers must be repaired, namely obstacles to the “sign-in” process and inefficient data verification processes.
The firm has clear specifications for implementing digital business transformation (DC1)	Specifications that are not yet concrete must be refined, namely the Standard Operation Procedure (SOP) for preventing errors during the service process, and the SOP for maintaining the network when digital banking transactions occur.	Specifications that are not yet concrete must be refined regarding efficient registration SOPs, appropriate SOPs for unlinking the device, and appropriate SOPs regarding cost disclosure.	Specifications that are not yet concrete must be refined regarding the SOP for the “sign-in” service to eliminate obstacles, and the SOP for the personal data verification process is appropriate and makes it easier for consumers.
The firm consistently tracks its digital business transformation progress (DC2)	The transformation monitoring system must be more stringent. Place the right staff to control it.	The transformation monitoring system must be more stringent. Place the right staff to control it.	The transformation monitoring system must be more stringent. Place the right staff to control it.
The firm regularly analyzes performance metrics to evaluate its digital business transformation (DC3)	Transformation performance metrics must continue to be analyzed accurately. Place the right staff to maintain the accuracy.	Transformation performance metrics must continue to be analyzed accurately. Place the right staff to maintain the accuracy.	Transformation performance metrics must continue to be analyzed accurately. Place the right staff to maintain the accuracy.

organizations can stay ahead of the curve and remain competitive in the fast-paced digital age.

4.6. Managerial Implication

Furthermore, to answer RQ2, we formulated a matrix that analyzes the findings of negative comments based on the affected DBC indicators and recommends practical solutions for parties involved in application transformation in digital banking in Indonesia, which is shown in Table 4.

The digital improvement solution matrix above explains how the next treatment for XAA, XBB, and XCC applications should be. However, of course, a transformation will be realized well if the existing management is also able to do it. Li et al. (2018) stated that managerial capabilities are the main driving force for digital transformation. Indeed, digital banking should have enabled banks to delight customers with instant services through distribution channels (Tam & Oliveira, 2017). Thus, this study has confirmed user expectations and strengthened them with digital improvements based on DBC theory and other theoretical supports as predictors of the success of XAA, XBB, and XCC.

5. Conclusion and Suggestion

The results reveal that most digital bank application users in Indonesia perceive that existing digital bank applications in Indonesia have yet to be able to meet their expectations. This is explained by the higher average negative value of their feedback answers than the existing positive value. Furthermore, this conclusion was revealed from the finding that XAA and XCC, which still had a negative score, had a higher positive score. Meanwhile, the XBB application, which is a first mover, was found to have a positive value higher than a negative one. This perceived performance view is a benchmark for the performance of digital banks in Indonesia to continue to develop their effectiveness and efficiency in managing service applications in direct contact with their users. The uniqueness of this research is that it uses three selected digital applications in Indonesia with categories from the Largest State Bank, First Mover, and Largest Private Bank. The limitation of this study is that the data analyzed are in the form of comments sourced from the Play Store, with a span of 3 months, with data sampling of 1000 data per application commented on. Academically, these findings enrich performance theory, especially regarding digital banks, that the positive and negative numbers illustrate the necessary strengthening of the current digital bank management model. For further research, this finding can be used to strengthen research on more effective application management models through the digital banking ecosystem. It is necessary to research the DBC connectivity on various variables, either affecting DBC or being affected by DBC. DBC can also continue to be tested for its suitability for measuring the effectiveness of an application, such as the banking system application in this study, by continuing to observe this theory and adapting to developments in the elements involved in digital bank operations. These findings can also be utilized by practitioners, especially the leaders of digital banks in Indonesia, to express user needs so that they are served according to their expectations. Several suggestions for existing application providers have also been presented in the managerial implications section.

References

Ahsan, M.J. (2024). Cultivating a culture of learning: The role of leadership in fostering lifelong development. *The Learning Organization*, 32(2), 282–306. https://doi.org/10.1108/TLO-03-2024-0099

Ajgaonkar, S., Neelam, N.G., & Wiemann, J. (2022). Drivers of workforce agility: A dynamic capability perspective. *International Journal of Organizational Analysis*, 30(4), 951–982. https://doi.org/10.1108/IJOA-11-2020-2507

Allam, H., Makubvure, L., Gyamfi, B., Graham, K.N., & Akinwolere, K. (2025). Text classification: How machine learning is revolutionizing text categorization. *Information*, 16(2), 130. https://doi.org/10.3390/info16020130

Alnaser, F.M., Rahi, S., Alghizzawi, M., & Ngah, A.H. (2023). Does artificial intelligence (AI) boost digital banking user satisfaction? Integration of expectation confirmation model and antecedents of artificial intelligence enabled digital banking. *Heliyon*, 9(8), e18930. https://doi.org/10.1016/j.heliyon.2023.e18930

Al-Okaily, M. (2024). Assessing the effectiveness of accounting information systems in the era of COVID-19 pandemic. *VINE Journal of Information and Knowledge Management Systems*, 54(1), 157–175. https://doi.org/10.1108/VJIKMS-08-2021-0148

Amankwah-Amoah, J., & Adomako, S. (2019). Big data analytics and business failures in data-Rich environments: An organizing framework. *Computers in Industry*, 105, 204–212. https://doi.org/10.1016/j.compind.2018.12.015

Azizi, M.R., Atlasi, R., Ziapour, A., Abbas, J., & Naemi, R. (2021). Innovative human resource management strategies during the COVID-19 pandemic: A systematic narrative review approach. *Heliyon*, 7(6), e07233.

https://doi.org/10.1016/j.heliyon.2021.e07233

Bajwa, S.S., Wang, X., Nguyen Duc, A., & Abrahamsson, P. (2017). Failures to be celebrated: an analysis of major pivots of software startups. *Empirical Software Engineering*, 22, 2373–2408. https://doi.org/10.1007/s10664-016-9458-0

Balakrishnan, R., & Das, S. (2020). How do firms reorganize to implement digital transformation? *Strategic Change*, 29(5), 531–541. https://doi.org/10.1002/jsc.2362

Barczak, G., Lassk, F., & Mulki, J. (2010). Antecedents of team creativity: An examination of team emotional intelligence, team trust and collaborative culture. *Creativity and Innovation Management*, 19, 332–345. https://doi.org/10.1111/j.1467-8691.2010.00574.x

Bastari, A., Eliyana, A., Syabarrudin, A., Arief, Z., & Emur, A.P. (2020). Digitalization in banking sector: The role of intrinsic motivation. *Heliyon*, 6(12), e05801. https://doi.org/10.1016/j.heliyon.2020.e05801

Begenau, J., Farboodi, M., & Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, 97, 71–87. https://doi.org/10.1016/j.jmoneco.2018.05.013

Bereznoy, A. (2019). Changing competitive landscape through business model innovation: The new imperative for corporate market strategy. *Journal of the Knowledge Economy*, 10(4), 1362–1383. https://doi.org/10.1007/s13132-015-0324-x

Berman, S., & Marshall, A. (2014). The next digital transformation: from an individual-centered to an everyone-to-everyone economy. *Strategy and Leadership*, 42(5), 9–17. https://doi.org/10.1108/SL-07-2014-0048

Berman, S.J. (2012). Digital transformation: Opportunities to create new business models. *Strategy and Leadership*, 40(2), 16–24. https://doi.org/10.1108/10878571211209314

Bhandari, U., Chang, K., & Neben, T. (2019). Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. *Information and Management*, 56(1), 85–93. https://doi.org/10.1016/j.im.2018.07.003

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25(3), 351–370. https://doi.org/10.2307/3250921

Bresciani, S., Ferraris, A., Romano, M., & Santoro, G. (2021). Building a digital transformation strategy. In: *Digital Transformation Management for Agile Organizations: A Compass to Sail the Digital World*. Emerald Publishing Limited, Leeds, p5–27. https://doi.org/10.1108/978-1-80043-171-320211002

Brill, T.M., Munoz, L., & Miller, R.J. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, 35(15–16), 1401–1436. https://doi.org/10.1080/0267257X.2019.1687571

Çakar, N.D., & Ertürk, A. (2010). Comparing innovation capability of small and medium-sized enterprises: Examining the effects of organizational culture and empowerment. *Journal of Small Business Management*, 48(3), 325–359. https://doi.org/10.1111/j.1540-627X.2010.00297.x

Chaffey, D., & Smith, P.R. (2022). *Digital Marketing Excellence: Planning, Optimizing and Integrating Online Marketing*. 6th ed. Routledge, England. https://doi.org/10.4324/9781003009498

Chakravarty, A.K. (2014). *Supply Chain Transformation*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41911-9

Chirumalla, K. (2021). Building digitally-enabled process innovation in the process industries: A dynamic capabilities approach. *Technovation*, 105, 102256. https://doi.org/10.1016/j.technovation.2021.102256

Chiu, J., & Koeppl, T. (2019). Blockchain-based settlement for asset trading. *The Review of Financial Studies*, 32(5), 1716–1753. https://doi.org/10.1093/rfs/hhy122

Chung, M., Ko, E., Joung, H., & Kim, S.J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, 117, 587–595. https://doi.org/10.1016/j.jbusres.2018.10.004

Cordon, C., Garcia-Milà, P., Vilarino, T.F., & Caballero, P. (2016). *Strategy is Digital*. 1st ed. Springer, Cham. https://doi.org/10.1007/978-3-319-31132-6

Dachner, A.M., Ellingson, J.E., Noe, R.A., & Saxton, B.M. (2021). The future of employee development. *Human Resource Management Review*, 31(2), 100732. https://doi.org/10.1016/j.hrmr.2019.100732

Dahlbom, P., Siikanen, N., Sajasalo, P., & Jarvenpää, M. (2020). Big data and HR analytics in the digital era. *Baltic Journal of Management*, 15(1), 120–138. https://doi.org/10.1108/BJM-11-2018-0393

Databoks. (2022). *Growth of Internet Banking Transactions in Indonesia Not Yet as Strong as Neighboring Countries*. Available from: https://databoks.katadata.co.id/keuangan/

statistik/7108401f77b080f/pertumbuhan-transaksi-internet-banking-ri-belum-sekuat-negara-tetangga [Last accessed on 2024 Feb 15].

Databoks. (2023). *Digital Banking Transactions in Indonesia Grow 158% in Last 5 Years*. Available from: https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/41234dcb958d96e/transaksi-digital-banking-di-indonesia-tumbuh-158-dalam-5-tahun-terakhir [Last accessed on 2024 Feb 15].

Detik. (2022). *Digital Transformation in Sharia Banking: BSI with the People*. Available from: https://news.detik.com/kolom/d-6480110/transformasi-digital-dalam-perbankan-syariah-bsi-bersama-rakyat [Last accessed on 2024 Feb 20].

Do, T.D., Pham, H.A.T., Thalassinos, E.I., & Le, H.A. (2022). The impact of digital transformation on performance: Evidence from vietnamese commercial banks. *Journal of Risk and Financial Management*, 15(1), 21. https://doi.org/10.3390/jrfm15010021

Dyer, L., & Shafer, R. (2003). *Dynamic Organizations: Achieving Marketplace and Organizational Agility with People*. 1st ed. Psychology Press. Available from: https://www.taylorfrancis.com/chapters/edit/10.4324/9781410607508-3/dynamic-organizations-lee-dyer-richard-shafer [Last accessed on 2024 Feb 13].

Eren, B.A. (2021). Determinants of customer satisfaction in chatbot use: Evidence from a banking application in Turkey. *International Journal of Bank Marketing*, 39(2), 294–311. https://doi.org/10.1108/IJBM-02-2020-0056

Fortune. (2022). *List of 7 Best Digital Banks in Indonesia 2022, Popular with Young People*. Available from: https://www.fortuneidn.com/finance/surti/bank-digital-di-indonesia?page=all [Last accessed on 2024 Feb 20].

Franque, F.B., Oliveira, T., & Tam, C. (2021). Understanding the factors of mobile payment continuance intention: Empirical test in an African context. *Heliyon,* 7(8), e07807. https://doi.org/10.1016/j.heliyon.2021.e07807

Fritz, K., Schoenmueller, V., & Bruhn, M. (2017). Authenticity in branding - exploring antecedents and consequences of brand authenticity. *European Journal of Marketing*, 51(2), 324–348. https://doi.org/10.1108/EJM-10-2014-0633

Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The role of technology in mortgage lending. *The Review of Financial Studies,* 32(5), 1854–1899. https://doi.org/10.1093/rfs/hhz018

Giacosa, E. Culasso, F., & Crocco, E. (2022) Customer agility in the modern automotive sector: How lead management shapes agile digital companies. *Technological Forecasting and Social Change*, 175, 121362. https://doi.org/10.1016/j.techfore.2021.121362

Gobble, M.M. (2018). Digital strategy and digital transformation. *Research-Technology Management*, 61(5), 66–71. https://doi.org/10.1080/08956308.2018.1495969

Godey, B., Manthiou, A., Pederzoli, D., Rokka, J., Aiello, G., Donvito, R., & Singh, R. (2016). Social media marketing efforts of luxury brands: Influence on brand equity and consumer behavior. *Journal of Business Research*, 69(12), 5833–5841. https://doi.org/10.1016/j.jbusres.2016.04.181

Gope, S., Elia, G., & Passiante, G. (2018), The effect of HRM practices on knowledge management capacity: A comparative study in Indian IT industry, *Journal of Knowledge Management*, 22(3), 649–677. https://doi.org/10.1108/JKM-10-2017-0453

Grover, V., & Malhotra, M.K. (1997). Business process reengineering: A tutorial on the concept, evolution, method, technology and application, *Journal of Operations Management*, 15(3), 193–213. https://doi.org/10.1016/S0272-6963(96)00104-0

Grover, V., Chiang, R.H.L., Liang, T.P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, 35(2), 388–423. https://doi.org/10.1080/07421222.2018.1451951

Guinan, P.J., Parise, S., & Langowitz, N. (2019). Creating an innovative digital project team: Levers to enable digital transformation. *Business Horizons*, 62(6), 717–727. https://doi.org/10.1016/j.bushor.2019.07.005

Gunasekaran, A., & Ngai, E.W. (2004). Information systems in supply chain integration and management. *European Journal of Operational Research*, 159(2), 269–295. https://doi.org/10.1016/j.ejor.2003.08.016

Gunasekaran, A., & Yusuf, Y.Y. (2002). Agile manufacturing: A taxonomy of strategic and technological imperatives. *International Journal of Production Research*, 40(6), 1357–1385. https://doi.org/10.1080/00207540110118370

Guo, P., & Shen, Y. (2016). The impact of Internet finance on commercial banks' risk taking: Evidence from China. *China Finance and Economic Review*, 4, 16. https://doi.org/10.1186/s40589-016-0039-6

Gupta, B.B., Gaurav, A., Panigrahi, P.K., & Arya, V. (2023). Analysis of artificial intelligence-based technologies and approaches on sustainable entrepreneurship. *Technological Forecasting and Social Change*, 186, 122152. https://doi.org/10.1016/j.techfore.2022.122152

Hall, D.M., Whyte, J.K., & Lessing, J. (2020). Mirror-breaking strategies to enable digital manufacturing in silicon valley construction firms: A comparative case study. *Construction Management and Economics*, 38(4), 322–339. https://doi.org/10.1080/01446193.2019.1656814

Henderson, J.C., & Venkatraman, H. (1999). Strategic alignment: Leveraging information technology for transforming organizations. *IBM Systems Journal*, 38(2.3), 472–484. https://doi.org/10.1147/SJ.1999.5387096

Ho, M.T., Le, N.T.B., Mantello, P., Ho, M.T., & Ghotbi, N. (2023). Understanding the acceptance of emotional artificial intelligence in Japanese healthcare system: A cross-sectional survey of clinic visitors' attitude. *Technology in Society*, 72, 102166. https://doi.org/10.1016/j.techsoc.2022.102166

Hoehle, H., Scornavacca, E., & Huff, S. (2012). Three decades of research on consumer adoption and utilization of electronic banking channels: A literature analysis. *Decision Support Systems*, 54(1), 122–132. https://doi.org/10.1016/j.dss.2012.04.010

Holbeche. (2007). *Understanding Change*. Routledge, England. Available from: https://www.taylorfrancis.com/books/mono/10.4324/9780080481500/understanding-change-linda-holbeche [Last accessed on 2024 Feb 12].

Holopainen, M., Ukko, J., & Saunila, M. (2022). Managing the strategic readiness of industrial companies for digital operations. *Digital Business*, 2(2), 100039. https://doi.org/10.1016/j.digbus.2022.100039

Islam, T., Ahmed, I., & Alkharabsheh, O.H.M. (2024), Fostering a knowledge sharing culture: the role of fun, empowerment, and inclusive leadership. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-12-2023-0471

Jaskyte, K. (2004). Transformational leadership, organizational culture, and innovativeness in nonprofit organizations. *Nonprofit Management and Leadership*, 15, 153–168. https://doi.org/10.1002/nml.59

Julien, A., & Tsoni, C. (2013). Front-line employee versus customer perceptions of quality attributes. *Quality Management Journal*, 20(2), 48–59. https://doi.org/10.1080/10686967.2013.11918097

Kamble, S.S., Gunasekaran, A., & Gawankar, S.A. (2018). Sustainable industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. *Process Safety and Environmental Protection*, 117, 408–425. https://doi.org/10.1016/j.psep.2018.05.009

Katsikeas, C., Leonidou, L., & Zeriti, A. (2020), Revisiting international marketing strategy in a digital era: Opportunities, challenges, and research directions. *International Marketing Review*, 37(3), 405–424. https://doi.org/10.1108/IMR-02-2019-0080

Khattak, M.A., Ali, M., Azmi, W., & Rizvi, S.A.R. (2023). Digital transformation, diversification and stability: What do we know about banks? *Economic Analysis and Policy*, 78, 122–132. https://doi.org/10.1016/j.eap.2023.03.004

Kim, B. (2018). Understanding the role of conscious and automatic mechanisms in social networking services: A longitudinal study. *International Journal of Human Computer Interaction*, 34(9), 805–818. https://doi.org/10.1080/10447318.2017.1392079

Kirtley, J., & O'Mahony, S. (2023). What is a pivot? Explaining when and how entrepreneurial firms decide to make strategic change and pivot. *Strategic Management Journal,* 44, 197–230. https://doi.org/10.1002/smj.3131

Knudsen, E.S., Lien, L.B., Timmermans, B., Belik, I., & Pandey, S. (2021). Stability in turbulent times? The effect of digitalization on the sustainability of competitive advantage. *Journal of Business Research*, 128, 360–369. https://doi.org/10.1016/j.jbusres.2021.02.008

Kokina, J., Pachamanova, D., & Corbett, A. (2017). The role of data visualization and analytics in performance management: Guiding entrepreneurial growth decisions. *Journal of Accounting Education*, 38, 50–62. https://doi.org/10.1016/j.jaccedu.2016.12.005

Kosiba, J.P., Acheampong, A., Adeola, O., & Hinson, R.E. (2020). The moderating role of demographic variables on customer expectations in airport retail patronage intentions of travellers. *Journal of Retailing and Consumer Services*, 54, 102033. https://doi.org/10.1016/j.jretconser.2020.102033

Lee, S., Kwon, Y., Quoc, N.N., Danon, C., Mehler, M., Elm, K., Bauret, R., & Choi, S. (2021). Red queen effect in german bank industry: Implication of banking digitalization for open innovation dynamics. *Journal of Open Innovation Technology Market and Complexity*, 7(1), 90. https://doi.org/10.3390/joitmc7010090

Leeflang, P.S.H., Verhoef, P.C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European Management Journal*, 32(1), 1–12.

https://doi.org/10.1016/j.emj.2013.12.001

Lemon, M., & Sahota, P.S. (2004). Organizational culture as a knowledge repository for increased innovative capacity. *Technovation*, 24(6), 483–498. https://doi.org/10.1016/S0166-4972(02)00102-5

Li, L., Su, F., Zhang, W., & Mao, J. (2018). Digital transformation by SME entrepreneurs: A capability perspective. *Information Systems Journal*, 28(6), 1129–1157. https://doi.org/10.1111/isj.12153

Lin, H.F. (2007). Knowledge sharing and firm innovation capability: An empirical study. *International Journal of Manpower*, 28(3/4), 315–332. https://doi.org/10.1108/01437720710755272

Liu, Y.F., Meyer, E., & Liu, X. (2009). Recent developments in digital control strategies for DC/DC switching power converters. *IEEE Transactions on Power Electronics*, 24(11), 2567–2577. https://doi.org/10.1109/TPEL.2009.2030809

Luftman, J.N., Lewis, P.R., & Oldach, S.H. (1993). Transforming the enterprise: The alignment of business and information technology strategies. *IBM Systems Journal*, 32(1), 198–221. https://doi.org/10.1147/sj.321.0198

Mbama, C.I., Ezepue, P., Alboul, L., & Beer, M. (2018). Digital banking, customer experience and financial performance: UK bank managers' perceptions. *Journal of Research in Interactive Marketing*, 12(4), 432–451. https://doi.org/10.1108/JRIM-01-2018-0026

Meena, A., Dhir, S., & Sushil, S. (2024). Coopetition, strategy, and business performance in the era of digital transformation using a multi-method approach: Some research implications for strategy and operations management. *International Journal of Production Economics*, 270, 109068. https://doi.org/10.1016/j.ijpe.2023.109068

Mendhakar, A., & Tilmatine, M. (2023). Automatic authorship classification for German lyrics using naïve bayes. *Journal for Language Technology and Computational Linguistics*, 36(1), 171–182. https://doi.org/10.21248/jlcl.36.2023.242

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 81. https://doi.org/10.1007/s13278-021-00776-6

Nylén, D., & Holmström, J. (2015). Digital innovation strategy: A framework for diagnosing and improving digital product and service innovation. *Business Horizons*, 58(1), 57–67. https://doi.org/10.1016/j.bushor.2014.09.001

OJK. (2021a). *Banking Digital Transformation Blueprint*. OJK, Indonesia, p1–54. Available from: https://www.ojk.go.id/id/berita-dan-kegiatan/info-terkini/documents/pages/cetak-biru-transformasi-digital-perbankan/cetak%20biru%20transformasi%20digital%20perbankan%20(short%20version).pdf [Last accessed on 2024 Feb 10].

OJK. (2021b). *POJK No.12/POJK.03/2021 about Commercial Banks*. www.ojk.go.id, p1–113. Available on line at: https://sikepo.ojk.go.id/SIKEPO/DatabasePeraturan/PeraturanUtuh/84c36c57-c4bb-4815-9b13-c229 [Last accessed on 2024 Feb 10.

Oppong, S.A., Yen, D.C., Merhout, J.W. (2005). A new strategy for harnessing knowledge management in e-commerce. *Technology in Society*, 27(3), 413–435. https://doi.org/10.1016/j.techsoc.2005.04.009

Patsiotis, A.G., Hughes, T., & Webber, D.J. (2012). Adopters and non-adopters of internet banking: A segmentation study. *International Journal of Bank Marketing*, 30(1), 20–42. https://doi.org/10.1108/02652321211195686

Pekovic, S., & Rolland, S. (2016). Customer orientation and firm's business performance: A moderated mediation model of environmental customer innovation and contextual factors. *European Journal of Marketing*, 50(12), 2162–2191. https://doi.org/10.1108/EJM-08-2015-0584

Piyathasanan, B., Mathies, C., Wetzels, M., Patterson, P.G., & Ruyter, K.D. (2015). A hierarchical model of virtual experience and its influences on the perceived value and loyalty of customers. *International Journal of Electronic Commerce*, 19(2), 126–158. https://doi.org/10.1080/10864415.2015.979484

Prastacos, G., Söderquist, K., Spanos, Y., & Wassenhove, L.V. (2002). An integrated framework for managing change in the new competitive landscape. *European Management Journal*, 20(1), 55–71. https://doi.org/10.1016/S0263-2373(01)00114-1

Proia, S., Carli, R., Cavone, G., & Dotoli, M. (2022). Control techniques for safe, ergonomic, and efficient human-robot collaboration in the digital industry: A survey. *IEEE Transactions on Automation Science and Engineering*, 19(3), 1798–1819. https://doi.org/10.1109/TASE.2021.3131011

Quinn, J.B., & Anderson, P. (1999). *Leveraging Intellect*. 1st ed. Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780080509778-19/leveraging-intellect-james-brian-quinn-philip-anderson-sydney-finkelstein [Last accessed on 2024 Feb 15].

Rahi, S., Alghizzawi, M., & Ngah, A.H. (2023). Factors influence user's intention to continue use of e-banking during COVID-19 pandemic: The nexus between self-determination and expectation confirmation model. *EuroMed Journal of Business*, 18(3), 380–396. https://doi.org/10.1108/EMJB-12-2021-0194

Ranjan, J., & Foropon, C. (2021). Big data analytics in building the competitive intelligence of organizations. *International Journal of Information Management*, 56, 102231. https://doi.org/10.1016/j.ijinfomgt.2020.102231

Ranjan, P. (2024). Unraveling the mystery of the link between digital orientation and innovation performance: The interplay of digital business capability and environmental dynamism. *Technovation*, 131, 102966. https://doi.org/10.1016/j.technovation.2024.102966

Ridho, F.M., Wibawa, A.P., Kurniawan, F., Badrudin, B., & Ghosh, A. (2022). Performance analysis of naive bayes in text classification of islamophobia issues. *Science in Information Technology Letters*, 3, 22–30. https://doi.org/10.31763/sitech.v3i1.1211

Roden, S., Nucciarelli, A., Li, F., & Graham, G. (2017). Big data and the transformation of operations models: a framework and a new research agenda. *Production Planning and Control*, 28(11–12), 929–944. https://doi.org/10.1080/09537287.2017.1336792

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, 23(4), 433–441. https://doi.org/10.1057/ejis.2014.17

Silva, T.C., Souza, S.R.S.D., Guerra, S.M., & Tabak, B.M. (2023). COVID-19 and bank branch lending: The moderating effect of digitalization. *Journal of Banking and Finance*, 152, 106869. https://doi.org/10.1016/j.jbankfin.2023.106869

Sonnentag, S., Niessen, C., & Ohly, S. (2004). Learning at work: Training and development. In: *International Review of Industrial and Organizational Psychology*. Wiley, United States. https://doi.org/10.1002/0470013311.ch8

Sorescu, A., & Schreier, M. (2021). Innovation in the digital economy: A broader view of its scope, antecedents, and consequences. *Journal of the Academy of Marketing Science*, 49(4), 627–631. https://doi.org/10.1007/s11747-021-00793-z

Statista. (2021). *Digital Payments - Worldwide*. Available from: https://www.statista.com/outlook/dmo/fintech/digital-payments/worldwide [Last accessed on 2024 Feb 25].

Stocker, A., Kaiser, C., Lechner, G., & Fellmann, M. (2024). A conceptual framework for mobility data science. *IEEE Access*, 12, 117126–117142. https://doi.org/10.1109/ACCESS.2024.3445166

Sultana, S., Akter, S., & Kyriazis, E. (2022). Theorising data-driven innovation capabilities to survive and thrive in the digital economy. *Journal of Strategic Marketing*, 32(7), 864–890. https://doi.org/10.1080/0965254X.2021.2013934

Szelągowski, M., & Berniak-Woźny, J. (2022). How to improve the assessment of BPM maturity in the era of digital transformation. *Information Systems and E-Business Management*, 20(1), 171–198. https://doi.org/10.1007/s10257-021-00549-w

Tam, C., & Oliveira, T. (2017). Literature review of mobile banking and individual performance. *International Journal of Bank Marketing*, 35(7), 1044–1067. https://doi.org/10.1108/IJBM-09-2015-0143

Tardieu, H., Daly, D., Esteban-Lauzán, J., Hall, J., & Miller, G. (2020). Innovation-recipes for staying ahead of the competition in a digital world. In: *Deliberately Digital*. *Future of Business and Finance*. Springer, Cham. https://doi.org/10.1007/978-3-030-37955-1-14

Tavitiyaman, P., So, C.Y.A., Chan, O.L.K., & Wong, C.K.C. (2022). How task technology fits with employee engagement, organizational support, and business outcomes: Hotel executives' perspective. *Journal of China Tourism Research*, 18(6), 1212–1238. https://doi.org/10.1080/19388160.2022.2027834

Theiri, S., & Alareeni, B. (2023). Perception of the digital transformation as a strategic advantage through the Covid 19 crisis? case of Tunisian banks. *Journal of Sustainable Finance and Investment*, 13(1), 477–498. https://doi.org/10.1080/20430795.2021.1964809

Ulrich, D., & Yeung, A. (2019). Agility: The new response to dynamic change. *Strategic HR Review*, 18(4), 161–167. https://doi.org/10.1108/SHR-04-2019-0032

Vashishth, T.K., Sharma, V., Sharma, K.K., Kumar, B., Chaudhary, S., & Panwar, R. (2024). Industry 4.0 trends and strategies: A modern approach with focus on knowledge management. In: *Knowledge Management and Industry Revolution*. Springer Nature, Germany. https://doi.org/10.1002/9781394242641.ch5

Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. https://doi.org/10.2307/30036540

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, 28(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Waite, K., & Harrison, T. (2015). Online banking adoption: We should know better 20 years on. *Journal of Financial Services Marketing*, 20(4), 258–272. https://doi.org/10.1057/fsm.2015.19

Wang, Y., Kung, L., & Byrd, T.A., (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126, 3–13. https://doi.org/10.1016/j.techfore.2015.12.019

Warner, K.S.R., & Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*, 52(3), 326–349. https://doi.org/10.1016/j.lrp.2018.12.001

Waseel, A.H., Zhang, J., Shehzad, M.U., Saddiqa, A., Liu, J., & Hussain, S. (2023), Does empowering leadership help firms to establish collaborative culture and organizational commitment to stimulate frugal innovation? *Kybernetes*, 54(2), 1289–1318. https://doi.org/10.1108/K-05-2023-0786

Wielgos, D.M., Homburg, C., & Kuehnl, C. (2021). Digital business capability: Its impact on firm and customer performance. *Journal of the Academy of Marketing Science*, 49(4), 762–789. https://doi.org/10.1007/s11747-021-00771-5

Xia, Q., Xie, Y., Hu, S., & Song, J. (2024), Exploring how entrepreneurial orientation improve firm resilience in digital era: Findings from sequential mediation and FsQCA. *European Journal of Innovation Management*, 27(1), 96–122. https://doi.org/10.1108/EJIM-12-2021-0593

Xie, X., & Wang, S. (2023). Digital transformation of commercial banks in China: Measurement, progress and impact. *China Economic Quarterly International*, 3(1), 35–45. https://doi.org/10.1016/j.ceqi.2023.03.002

Yeow, A., Soh, C., & Hansen, R. (2018). Aligning with new digital strategy: A dynamic capabilities approach. *The Journal of Strategic Information Systems*, 27(1), 43–58. https://doi.org/10.1016/j.jsis.2017.09.001

Zaki, M. (2019). Digital transformation: harnessing digital technologies for the next generation of services. *Journal of Services Marketing*, 33(4), 429–435. https://doi.org/10.1108/JSM-01-2019-0034

Zhang, H. (2004). *The Optimality of Naive Bayes. Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference*. FLAIRS 2004, 2, 562–567.

Zhu, C. (2019). Big data as a governance mechanism. *The Review of Financial Studies*, 32(5), 2021–2061. https://doi.org/10.1093/rfs/hhy081

Zolkepli, I.A., & Kamarulzaman, Y. (2015). Social media adoption: The role of media needs and innovation characteristics. *Computers in Human Behavior*, 43, 189–209. https://doi.org/10.1016/j.chb.2014.10.050

AUTHOR BIOGRAPHY

Teuku Roli Ilhamsyah Putra is an Associate Professor in the Management Department at Universitas Syiah Kuala, Indonesia. He has extensive experience in both academia and practical applications. His academic background includes teaching and research, while his practical experience consists of working as a management consultant. His primary research areas include management science, strategic management, human resource management, and marketing management. His active email is: teuku.roli@usk.ac.id

Muhammad Iqbal Fajri is a lecturer in the Management Department at Universitas Syiah Kuala, Indonesia. He has extensive experience in practical management, particularly as a Data Analyst. Over the years, he has worked with several companies, and he is now fully dedicated to his role as a lecturer and researcher. His research interests encompass various areas of management, including digital business, marketing management, digital marketing, and tourism marketing. His active email is: iqblfjri@feb.usk.ac.id.

INSTRUCTIONS TO AUTHORS

Submission of papers

The International Journal of Systematic Innovation is a refereed journal publishing original papers four times a year in all areas of SI. Papers for publication should be submitted online to the IJoSI website (http://www.ijosi.org) In order to preserve the anonymity of authorship, authors shall prepare two files (in MS Word format or PDF) for each submission. The first file is the electronic copy of the paper without author's (authors') name(s) and affiliation(s). The second file contains the author's (authors') name(s), affiliation(s), and email address(es) on a single page. Since the Journal is blind refereed, authors should not include any reference to themselves, their affiliations or their sponsorships in the body of the paper or on Figs and computer outputs. Credits and acknowledgement can be given in the final accepted version of the paper.

Editorial policy

Submission of a paper implies that it has neither been published previously nor submitted for publication elsewhere. After the paper has been accepted, the corresponding author will be responsible for page formatting, page proof and signing off for printing on behalf of other co-authors. The corresponding author will receive one hardcopy issue in which the paper is published free of charge.

Manuscript preparation

The following points should be observed when preparing a manuscript besides being consistent in style, spelling, and the use of abbreviations. Authors are encouraged to download manuscript template from the IJoSI website, http://www.ijosi.org.

1. *Language.* Paper should be written in English except in some special issues where Chinese may be accepTable Each paper should contain an abstract not exceeding 200 words. In addition, three to five keywords should be provided.
2. *Manuscripts.* Paper should be typed, single-column, double-spaced, on standard white paper margins: top = 25mm, bottom = 30mm, side = 20mm. (The format of the final paper prints will have the similar format except that double-column and single space will be used.)
3. *Title and Author.* The title should be concise, informative, and it should appear on top of the first page of the paper in capital letters. Author information should not appear on the title page; it should be provided on a separate information sheet that contains the title, the author’s (authors') name(s), affiliation(s), e-mail address(es).
4. *Headings.* Section headings as well as headings for subsections should start front the left-hand margin.
5. *Mathematical Expressions.* All mathematical expressions should be typed using Equation Editor of MS Word. Numbers in parenthesis shall be provided for equations or other mathematical expressions that are referred to in the paper and be aligned to the right margin of the page.
6. *Tables and Figs.* Once a paper is accepted, the corresponding author should promptly supply original copies of all drawings and/or tables. They must be clear for printing. All should come with proper numbering, titles, and descriptive captions. Fig (or table) numbering and its subsequent caption must be below the Fig (or table) itself and as typed as the text.
7. *References.* Display only those references cited in the text. References should be listed and sequenced alphabetically by the surname of the first author at the end of the paper. For example:

Altshuller, G. (1998). *40 Principles: TRIZ Keys to Technical Innovation*, Technical Innovation Center.
Sheu, D. & Lee, H. (2011). A Proposed Process for Systematic Innovation, International Journal of Production Research, Vol. 49, No. 3, 2011, 847-868.

The International Journal of Systematic Innovation
Journal Order Form

Organization Or Individual Name	
Postal address for delivery	
Person to contact	Name: e-mail: Position: School/Company:
Order Information	**I would like to order ___ copy(ies) of the** *International Journal of Systematic Innovation***:** **Period Start: 1st/ 2nd half ____, Year:____(Starting 2010)** **Period End : 1st/ 2nd half ____, Year:** **Price:** **Institutions: US $150 (yearly) / NT 4,500 (In Taiwan only)** **Individuals: US $50 (yearly) / NT 1500 (In Taiwan only)** (Local postage included. International postage extra) **E-mail to**: IJoSI@systematic-innovation.org or fax: +886-3-572-3210 Air mail desired □ (If checked, we will quote the additional cost for your consent)
Total amount due	**US$**

Payment Methods:
1. **Credit Card (Fill up the following information and e-mail/ facsimile this form to The Journal office indicated below)**
2. **Bank transfer**
3. **Account:** The Society of Systematic Innovation
4. **Bank Name:** Mega International Commercial BANK
5. **Account No:** 020-53-144-930
6. **SWIFT Code:** ICBCTWTP020
7. **Bank code：** 017-0206
8. **Bank Address:** No. 1, Xin'an Rd., East Dist., Hsinchu City 300, Taiwan (R.O.C.)

VISA / Master/ JCB/ AMERICAN Cardholder Authorization for Journal Order

Card Holder Information

Card Holder Name	(as it appears on card)		
Full Name (Last, First Middle)			
Expiration Date	/ (month / year)	Card Type	□ VISA □ MASTER □ JCB
Card Number	□□□□-□□□□-□□□□-□□□□	Security Code	□□□ 
Amount Authorized		Special Messages	
Full Address (Incl. Street, City, State, Country and Postal code)			

Please Sign your name here _________________ (same as the signature on your card)

The Society of Systematic Innovation

6 F, #352, Sec. 2, Guanfu Rd,
Hsinchu, Taiwan, 30071, R.O.C.