ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202409_8(3)

# International Journal of Systematic Innovation

VOL. 08 NO. 03
September, 2024

Published by the Society of Systematic Innovation

## Opportunity Identification & Problem Solving

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202409_8(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
- web site: http://www.IJoSI.org
- Tel: +886-3572-3200

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202409_8(3)

# International Journal of Systematic Innovation

## Contents

September 2024 Volume 8 Issue 3

### Full Papers

DOI: 10.6977/IJoSI.202409_8(3).0001
S. Ch, J. Dong, etc./Int. J. Systematic Innovation, 8(3),1-11 (2024)

# Project-based teaching of product innovation design based on KJ/FAST/TRIZ

Siyuan Cheng[1], Jie Dong[2], Xuerong Yang[3]
School of Electro-mechanical Engineering, Guangdong University of Technology
* Corresponding author E-mail: imdesign@gdut.edu.cn

(Received 23 July 2023; Final version received 06 July 2024; Accepted 16 July 2024)

## Abstract

In the new era of "mass innovation," higher requirements are being placed on the cultivation of innovative talents in colleges and universities. In response to this, the project teaching process of product innovation design based on KJ/FAST/TRIZ is proposed. Firstly, the original user needs are obtained through investigation, and the KJ method is used to classify these needs in order to determine the target of product design. Subsequently, guided by these objectives, the FAST function tree is utilized to establish the functional model. Then the Function-Oriented Search in TRIZ theory is employed to address key sub-functions, and the conflicts within the functional model are resolved using the Contradiction Matrix. Ultimately, a product innovation design scheme that meets the needs of users is obtained. Furthermore, a case study of this project-based teaching process was provided using the innovative design of a Household Intelligent Storage Cabinet as an example. This teaching method takes user demand as the orientation takes project implementation as the carrier and utilizes a comprehensive application of various tools to guide students' completion of product innovation design projects in steps, thereby to enhance the cultivation of students' innovative practical abilities.

*Keywords: Innovative design, Project-based teaching, TRIZ, User needs*

## 1. Introduction

In the recent international trend of "returning to engineering," project-based teaching, as a student-centered pedagogical approach, has garnered significant attention and promotion in various engineering institutions. This is due to its proven ability to effectively enhance students' engineering practical skills, teamwork capabilities, problem analysis, and problem-solving abilities (Zhang et al., 2023). Project-based teaching revolves around student-centered learning, guided by instructors. It relies on a specific real or virtual project, or competition topics within a professional field. This approach integrates and interconnects the foundational knowledge and fundamental principles that need to be taught in the course with their corresponding practical components. It is presented through the process and methodology of hands-on project work within the teaching context. Students proactively acquire relevant knowledge based on the project, as guided by this instructional method (Zhao et al., 2019).

In June 2016, China became an official contracting member of the "Washington Accord", marking a new historical phase in the country's engineering education reform (Geng et al., 2018). According to the requirements of the "Washington Accord," programs accredited through engineering education should be able to cultivate students with the capability to solve complex engineering problems. The accreditation standards for engineering education in China also explicitly state the need to address complex engineering issues, while defining the "1+X" characteristics associated with these complex problems (Lin, 2016). Therefore, the cultivation of students' ability to solve complex engineering problems has become a significant challenge within engineering education. Project-based teaching is well-aligned with the "1+X" characteristics of complex engineering problems. Throughout the process of project-based learning, students engage in solving open-ended real-world engineering problems, applying theoretical knowledge to practical scenarios, and constantly attempting to find solutions to open-ended problems. This approach is a vital means of fostering innovative thinking and the skills to tackle complex engineering challenges.

In response to the era of "mass innovation" and in close alignment with the new demands placed on engineering and technical professionals by innovation-driven development, our university's mechanical department has introduced a course titled "Innovation Methods." This course teaches the Theory of Inventive Problem Solving (TRIZ), also known as the theory of inventing solutions to problems. TRIZ provides engineers with a systematic approach to innovation, which can lead to the reduction of product development time cycles and an increase in design efficiency (Cano-Moreno et al., 2022). However, TRIZ theory primarily focuses on solving engineering problems and may struggle to capture and assess the diverse requirements of users during the early stages of design, which doesn't cover the entire product design process. As a result, many scholars have proposed combining TRIZ theory with other design methods. This integration can offer better clarity regarding user requirements, the identification of relationships between requirements and product functionalities, and subsequently, the utilization of TRIZ theory to solve and optimize product function models, which forms a complete innovation design process. Murat Mayda et al. proposed integrating TRIZ theory into Pahl and Beitz's conceptual design process to assist problem solving and solution improvement (Mayda &Börklü,2014). Yuanwu Shi et al. combined FAST and TRIZ methods to introduce an innovative approach to designing elderly smart care robots (Shi & Pan,2018). Fiorineschi et al. suggested combining TRIZ with Problem Solution Network (PSN) to enhance conceptual design efficiency (Fiorineschi et al.,2018). Yang et al. proposed merging QFD with TRIZ to achieve Radical innovations (Yang et al.,2021). Fuqin Liu et al. presented a product conceptual design process based on KJ, KANO, and FAST models, exemplified by innovative design of a forest fire truck (Liu & Li, 2022). Anqi Wu et al. integrated AHP, QFD, and TRIZ theories to propose solutions for aspects like cabin, seats, battery life, and guidance in a shared mobility vehicle for scenic areas (Wu et al., 2022).

Existing research has demonstrated the value of integrating TRIZ theory into the product design process. However, in the context of project-based teaching, the characteristics of instructional implementation need to be considered. Process design should align with students' theoretical foundation, time constraints, and other practical aspects. This article addresses the cultivation of complex engineering problem-solving abilities outlined in engineering certification standards, as well as the advanced requirements of First-class courses (Yuan et al., 2022). Combining the development of the "Innovative Methods" First-class course, a project-based teaching approach that integrates KJ method, FAST method, and TRIZ theory is proposed. This approach is guided by the organic fusion of knowledge, skills, and qualities, providing a reference for fostering complex engineering problem-solving abilities and developing First-class courses in relevant disciplines.

## 2. Project-based Teaching Process Based on KJ/FAST/TRIZ

In courses such as "Innovative Methods" and "Project Design," a project-based teaching approach reform has been implemented. Students are required to engage in the concept design of products through three main steps: requirement analysis, establishment of functional models, and functional problem-solving, as illustrated in Figure 1. The initial phase is the requirement analysis stage. After posing the problem, market research is conducted to understand the market background. User's original requirement descriptions are obtained through surveys, interviews, and other methods. Subsequently, the obtained requirements are classified using the KJ method. Following this, the establishment of functional models phase takes place. The FAST functional tree is employed to transform the product's requirements into functions, identifying primary and secondary functions. Lastly, the functional problem-solving phase occurs. Utilizing TRIZ tools such as function-oriented search and the contradiction matrix, general solutions to common problems are derived. These solutions are then linked to specific issues, resulting in specific solutions for the design process. This process culminates in the formulation of a creative design scheme.

In this project-based teaching approach, teachers assign open, real-world, relatively independent projects to student groups for design and implementation. This encompasses tasks such as data collection, conceptualization of solutions, and project design. Regarding the instructional content, teachers, while delivering instruction on innovative methods, appropriately integrate relevant knowledge and methodologies like requirement analysis and functional decomposition. Student-centeredness during the project implementation process is emphasized and practical applications tailored to the project are carried out. By employing a problem-driven approach, students' curiosity and desire for knowledge are aroused, transitioning their traditional learning methods towards exploratory and challenging learning. Additionally, students are encouraged to transform their

design proposals into patentable outcomes or to participate in scientific and technological competitions.

In a longitudinal aspect, this project establishes a connection between professional education and societal development. Through the integration and application of the curriculum's knowledge structure within the project, an interactive pathway is forged between knowledge and the external world. This facilitates the practical implementation of professional teaching, catering to the needs of the service industry and societal development, while also embodying the characteristics of complex engineering problems. Horizontally, adhering to cognitive principles, the course content is progressively structured from basic to advanced, and from simple to complex. This incremental approach aligns with students' habits of adapting to actual engineering applications, gradually nurturing and enhancing their ability to solve complex engineering problems.

## 3. Innovative Design Case of a Household Intelligent Storage Cabinet

### 3.1 Project Design Background

In traditional household living, storage is typically reliant on conventional furniture such as shelves, cabinets, and coffee tables. However, these traditional storage methods fail to meet the requirements of special groups within the household context, such as individuals with conditions like Alzheimer's disease or lower limb disabilities. For instance, individuals with Alzheimer's disease often suffer from memory loss, leading them to frequently forget the locations where they have stored items, as well as everyday tasks like taking medication. On the other hand, individuals with lower limb disabilities face challenges due to physical constraints. They encounter difficulties in retrieving items placed at a distance or storing items that are positioned either too high or too low. For users lacking organized storage awareness, there's a possibility of important items getting lost due to haphazard placement, leading to a cluttered living environment and a decrease in quality of life. Therefore, to address the storage requirements of these special groups and enhance their daily lives, the design of an intelligent storage cabinet tailored to their requirements is essential. This case study aims to exemplify the process of project-based teaching by having students design a household intelligent storage cabinet that caters to the requirements of these special groups.

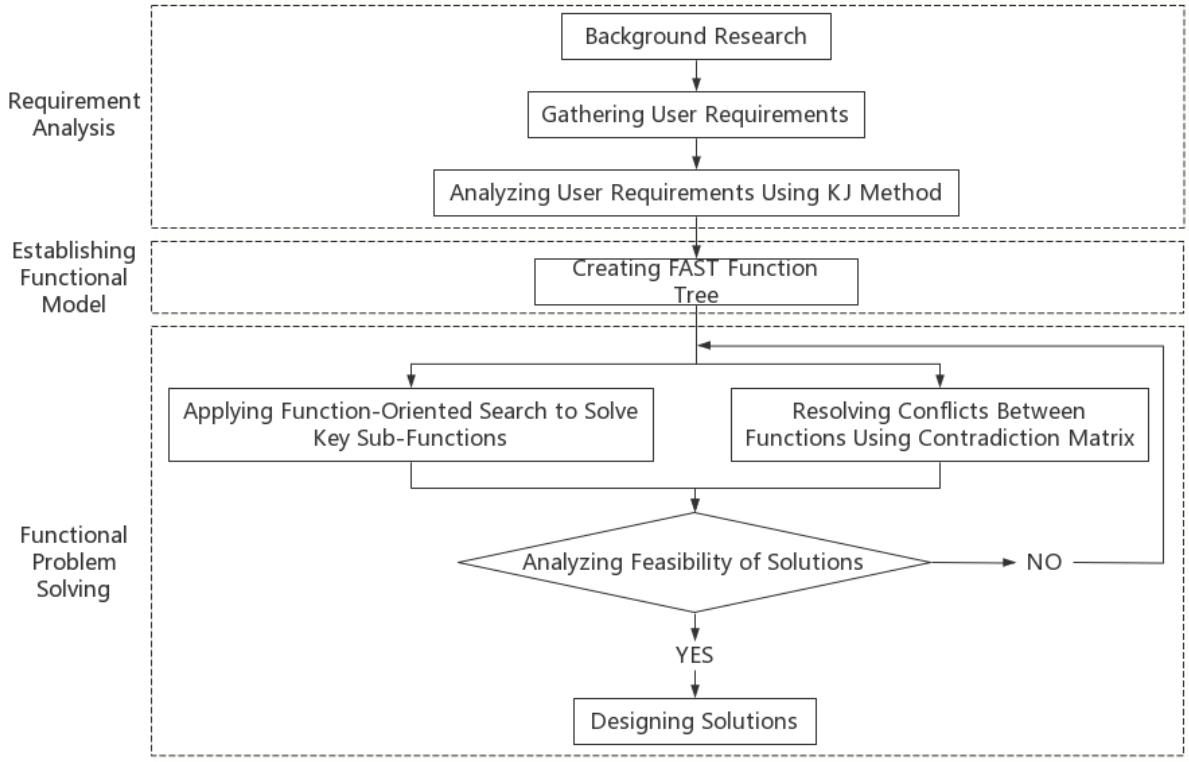



**Fig 1.** Innovative design process.

**Table 1.** Original requirement description.

| No. | Description of Requirements | No. | Description of Requirements |
|---|---|---|---|
| 1 | Daily Access to Items | 7 | Emergency SOS |
| 2 | Easy Retrieval | 8 | User-Adjustable Variability |
| 3 | Daily Task Reminders | 9 | Clear Item Categorization |
| 4 | Simple Operation | 10 | Emergency Medication Box |
| 5 | Environmentally Friendly | 11 | Space Saving |
| 6 | Entertainment and Decoration | 12 | Automatic Tracking |

## 3.2 User Requirements Analysis

User requirements serve as the driving force for product design, permeating the entire design process. Initially, students are tasked with conducting surveys among typical users, Alzheimer's disease patients, and individuals with lower limb disabilities. Subsequently, they are expected to conduct interviews with representative users from each category to gather the original user requirements. Based on the survey results, it is evident that the three groups share a fundamental requirement for convenient storage. However, variations exist in terms of other auxiliary functionalities based on their specific requirements. In summary, for the household intelligent storage cabinet, it is imperative to meet the basic storage requirements of all users. Additionally, it should incorporate human-machine interaction to assist Alzheimer's disease patients in storing items. For individuals with lower limb disabilities, mechanisms for adjusting the cabinet's height and outlet positions should be integrated to facilitate storage. Moreover, efficient classification attributes and user-friendly operations should aid individuals with poor storage habits in accessing and organizing items. Based on the consolidation, a total of 12 original requirements from the special user groups for household storage have been identified, as shown in **Table 1.**

KJ Analysis Method, also known as Affinity Diagram, primarily relies on the interrelation of information to categorize cluttered information into different hierarchies, clarifying the corresponding relationships among the information. When applying the KJ method in the analysis of user requirements for the household intelligent storage cabinet, user needs can be categorized with mutual affinity. This approach is presented through a combination of graphics and text,

using a visual format to reveal the essence of issues within the intricate web of user requirements (He, 2022). As illustrated in **Figure 2**, the three categories of the primary requirement indicator diagram for the household intelligent storage cabinet can be obtained. These categories are considered from the perspectives of storage demands, safety and health, and auxiliary adjustment.

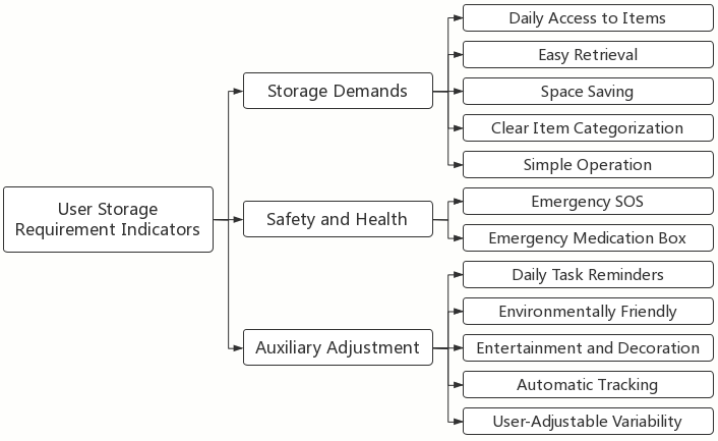



**Fig 2.** Three-level demand indicator chart.

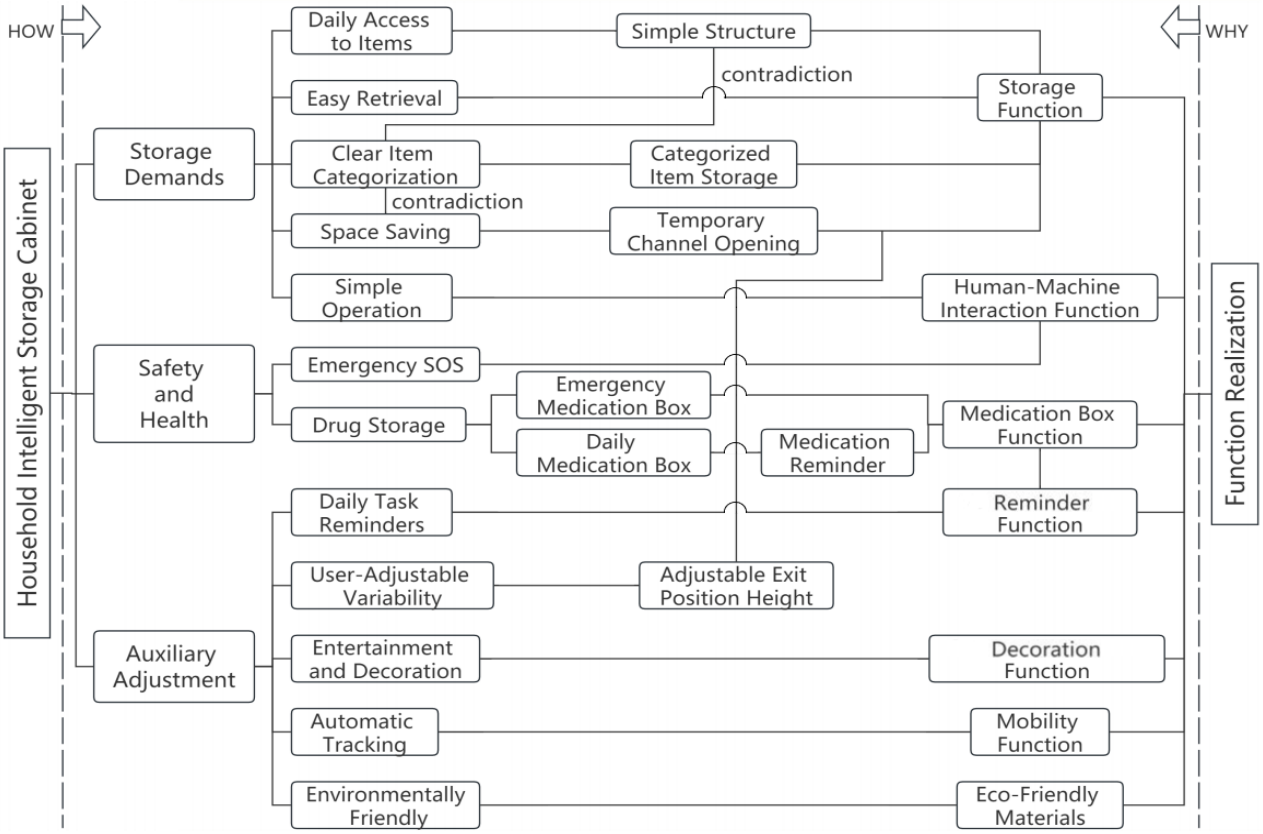



**Fig 3.** The FAST function tree of household intelligent storage cabinets.

## 3.3 Constructing a Product Functional Model

FAST method, also known as Function Analysis System Technique, is a top-down approach that analyzes the interrelationships between functions from a systemic perspective. It employs the construction of a FAST function tree to illustrate these relationships (Du et al., 2022). For the user requirements established in this case, a FAST function tree is constructed based on the perspectives of storage demands, safety and health, and auxiliary adjustment. This tree models the functionalities of the household intelligent storage cabinet, progressively analyzing the interrelationships between functions, as shown in Figure 3 below. Ultimately, the primary function of the household intelligent storage cabinet is determined to be storage, catering to users' efficient storage needs. Secondary functions include human-machine interaction, medication box functionality, reminder capability, decoration, and mobility, addressing specific requirements of distinct user groups. Moreover, a conflict is identified between the demand for clear item categorization and the requirements for a simple structure and space-saving design.

## 3.4 Functional Problem Solving

The general process of solving problems using the TRIZ theory involves simplifying the design-related issues into generic problems. Through the inventive principles, separation principles, scientific effects database, and other tools provided by TRIZ theory, universal

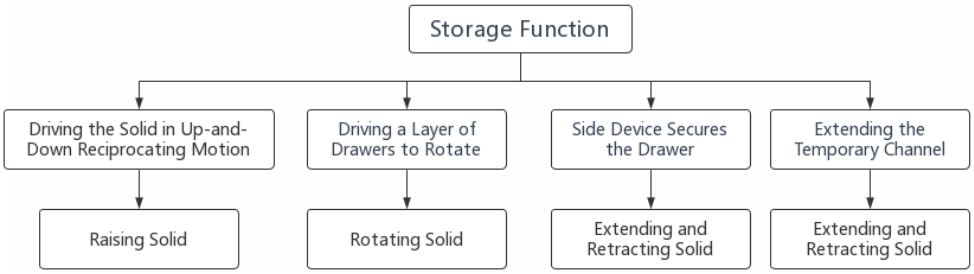



**Fig 4.** A generalized description of the storage function.

solutions for these general problems can be found (Cheng et al., 2021). Given that the storage function involves a larger number of target users and user requirements, represents a pivotal sub-function within the FAST function tree. The storage function will be taken as an example and TRIZ theory's function-oriented search and contradiction matrix is applied to solve and optimize it.

(1) Functional Problem Solving Based on Function-Oriented Search

Function-oriented search is a problem-solving tool based on the analysis of existing mature technologies worldwide. In the process of functional problem solving, utilizing TRIZ theory's function-oriented search, benchmarking analysis, and feature transfer for design can reduce the randomness in the design process and enhance design efficiency.

a. Generalized Description of Function

Combining the analysis of the usage process, the storage function can be broken down into four secondary sub-functions. Firstly, the up-and-down reciprocating movement of the central rotating device can position it at the same height as the drawers storing the desired items. Secondly, the rotation of the drawer containing the required items is propelled by the rotating device, and rotating the drawer containing the desired items to a reachable position on the side. Thirdly, a side fixation mechanism secures the drawer containing the required items. Finally, by extending a temporary channel, the drawer containing the required items is secured by the side fixation mechanism and then extends alongside the temporary channel simultaneously. It is then moved by the up-and-down reciprocating movement device to the upper exit, enabling users to retrieve the desired items.

As depicted in Figure 4 below, the four-step motion sequence of the storage function can be generalized as follows.

b. Function-Oriented Search

Function-oriented search is a tool that is used to find functional solutions in mature technological domains. It begins with behavior and objects and is used to seek reliable technical solutions in specific knowledge databases through a generalized description of functions. These solutions are then transferred to handle technical issues that arise during the design of new products. Specialized commercial software tools like Invention Tool or Goldfire can be utilized for this purpose, and searches can also be conducted through search engines and patent databases. An example of function-oriented search results for the functions 'Rotating Solid' and 'Extending and Retracting Solid' is presented in Figures 5 and 6 below.

c. Benchmarking Analysis and Feature Transfer

Benchmarking analysis is centered around the primary values within the engineering system, further analyzing and comparing the results selected through function-oriented search. The main principle is to achieve product functionality with low cost and ease of operation. Feature transfer is an analytical tool that extracts relevant characteristics from alternative products, which means transferring solutions or principles from more mature areas in other technological developments that address similar issues to the product being designed, for resolving existing problems (Gui et al.,2016).

Taking the example of driving a specific layer of drawers to rotate, benchmarking analysis revealed that a gearbox is structurally more suitable, relatively cost-effective, and offers higher space utilization. Therefore, the gearbox's structure is applied to drive the rotation of a specific layer of drawers through feature transfer. Through feature transfer, it is discovered that the gearbox primarily relies on the meshing of gears to achieve the rotation of the solid being driven. When applied to drive a specific layer of drawers, gear rotation can also be

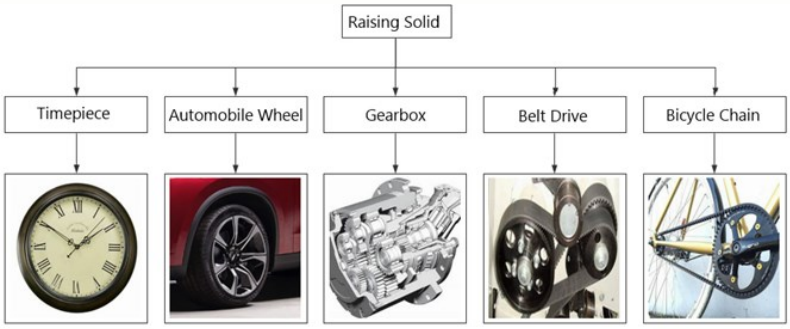



**Fig 5.** Feature oriented search for 'rotating solids'

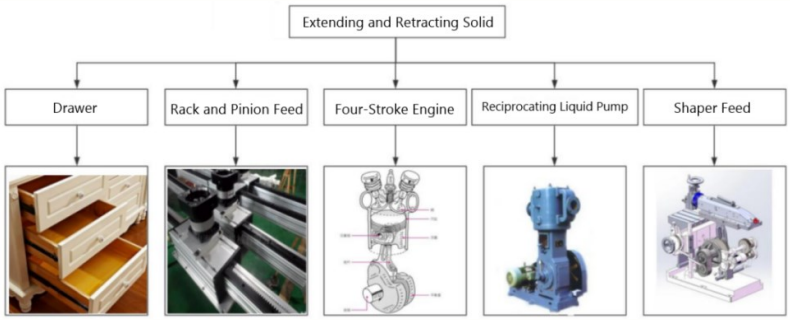



**Fig 6.** Function - oriented search for 'extend and pull back solids'

achieved through the intermeshing of gears, with a special cam fixed on a large gear to facilitate the rotation of the cam, which in turn rotates a specific layer of drawers.

Applying the same method to analyze other related sub-functions, the results of feature transfer are shown in Table 2 below, thus achieving the four secondary sub-functions of the household intelligent storage cabinet's storage function.

**(2).** Function Optimization Based on the Contradiction Matrix

The Contradiction Matrix of TRIZ theory establishes a correspondence between 39 engineering parameters and 40 inventive principles. After analyzing practical problems and consulting the Contradiction Matrix, corresponding inventive principles can be found through 'improved parameters' and 'worsened parameters.' These inventive principles provide a direction to help find specific solutions to the problem at hand (Huang et al., 2018).

Identifying contradictions through the utilization of the FAST function tree can sometimes present challenges. To address this issue, it is imperative to ensure that during the establishment of the FAST function tree, all functionalities are distinctly defined, specific, and comprehensible. Categorizing functionalities into primary and secondary categories and conducting an analysis of the relationships and dependencies among them through the depiction of function dependency diagrams is crucial. By discerning direct or indirect dependencies among functionalities, potential conflict points can be pinpointed. Subsequently, through brainstorming sessions and team discussions, potential functional conflicts can be identified.

Analyzing the FAST function tree and user requirements, it is evident that the household intelligent storage cabinet needs to ensure clear classification, which can be achieved by increasing the number of drawers and refining their categorization properties. However, increasing the number of drawers will enlarge the device's volume and complicate its structure, which falls under a technical contradiction. After considering the actual situation, a search in the contradiction matrix pinpointed the engineering parameter to be improved as 'the quantity of substance or transaction,' and the worsened engineering parameters as 'the volume of moving objects' and 'the complexity of the device.' By consulting the TRIZ Contradiction Matrix, corresponding inventive principles can be obtained, as shown in Table 3 below.;

| | Original Function Descriptions | Generalized Function Description | Benchmarking Analysis | Feature Transfer |
|---|---|---|---|---|
| Storage Function | Driving the Solid in Up-and-Down Reciprocating Motion | Raising Solid | Ball Screw Elevator | Ball Screw Transmission Principle |
| | Driving a Layer of Drawers to Rotate | Rotating Solid | Gearbox | Straight Spur Gear Transmission Principle |
| | Side Device Secures the Drawer | Extending and Retracting Solid | Gear Rack Feeding | Gear Rack Transmission Principle |
| | Extending the Temporary Channel | Extending and Retracting Solid | Four-Stroke Engine | Crank and Slider Mechanism Transmission Principle |

**Table 2.** Results of benchmarking analysis and feature transfer.

**Table 3.** The paradox Matrix of Household Intelligent storage cabinets.

| Improved Engineering Parameters | Worsened Engineering Parameters | |
|---|---|---|
| | 7 Volume of Moving Object | 36Complexity of the Apparatus |
| 26Quantity of Material or Events | 15，20，29 | 3，13，27，10 |

After analyzing the above-mentioned inventive principles, it is found that Inventive Principles 3, 15, and 20 are particularly helpful in optimizing the functionality of the household intelligent storage cabinet. Applying Inventive Principles 3, 15, and 20 to address the conflicts, as shown in Table 4 below.

**Table 4.** Application of invention principle.

| No. | Inventive Principle | Apply |
|---|---|---|
| 3 | Local Quality | It is possible to adjust the shape of the drawers to accommodate various sizes and forms of different household items. |
| 15 | Dynamics | Users can create their own designs for drawer shapes, adapting them to suit the storage needs of different items, utilizing rapid assembly and straightforward installation methods. |
| 20 | Continuity of Useful Action | Enabling the interlocking of drawers with varying sizes and shapes enhances spatial utilization. Ensuring a more seamless motion flow, reducing gaps between the motion phases of distinct secondary sub-functions, thereby elevating efficiency. |

The improved drawer shape, as a result of applying the inventive principles, is illustrated in Figure 7 below.

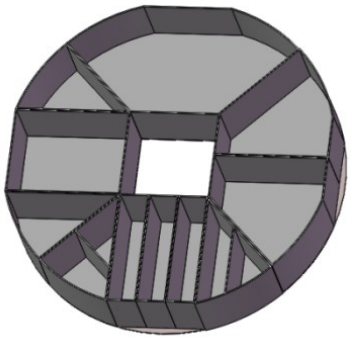
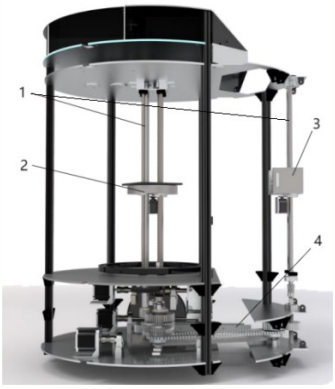


**Fig 7**. A cross-section of a drawer.

## 3.5 Integrated Design Solution

For the other sub-functions related to the storage functionality of the household intelligent storage cabinet, we applied a function-oriented search to find suitable methods from mature domains and integrated them into the design of the product. The resulting internal structure of the storage cabinet is depicted in Figure 8. When using the product, users can initially input the items they need through methods such as voice or touch screen. The intelligent storage cabinet will then, based on the user's requirements, perform a series of movements to move the storage space containing the desired item to a convenient location for user access.

The movement of the intelligent storage cabinet is realized through its structural design. First, through user input, the specific location where the user needs to retrieve the item is determined, which is drawer A. Mechanism 2 moves vertically alongside Mechanism 1's rotation to the same height as drawer A. Mechanism 2 rotates to move drawer A to a lateral position. Mechanism 3 locks drawer A in place, and then Mechanism 4 begins to move, creating a temporary passage. Drawer A is fixed by Mechanism 3 and simultaneously moved into the temporary passage. Finally, through the movement of Mechanism 1, drawer A is lifted to the exit position for user access.




1. Ball screw, 2. A middle rotating device, 3. Side-fixed device, 4. Crank-slider mechanism

**Fig 8**. Internal structure of storage function based on intelligent storage cabinet.

For other important sub-functions in the FAST function tree, a similar process can be applied for design as described above. When combined with conventional design practices, it eventually results in a comprehensive design solution for the household intelligent storage cabinet, as illustrated in Figure 9. The household intelligent storage cabinet designed in this case adds features such as mobility, human-computer interaction, reminder, and medicine box functions compared to traditional storage

cabinets. Through intelligent input, intelligent positioning movement, and intelligent daily reminders, the intelligent storage cabinet is characterized by clear classification, convenient storage, simple operation, and a high level of intelligence.

Additionally, the product offers a range of auxiliary functions to meet the daily needs of special user groups. It assists Alzheimer's patients in storing items through human-computer interaction, helps lower limb-disabled individuals store items by adjusting the mobility and height of the outlet, and aids those without good storage habits with effective categorization attributes and convenient operations. Overall, it effectively satisfies user requirements.

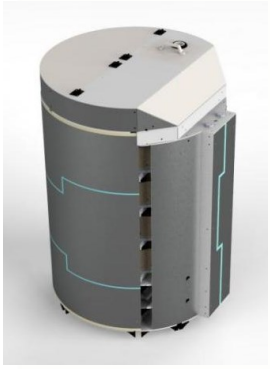


**Fig 9.** Comprehensive design of household intelligent storage cabinets.

Taking into account that the project topic originates from real-life scenarios, possesses openness and societal relevance, and its implementation process necessitates an in-depth analysis of engineering principles, encompassing multiple interrelated sub-problems, and showcasing creativity, it also strongly emphasizes the development of abilities in addressing complex engineering issues and fostering higher-order thinking.

## 4. Conclusion

A project-based teaching method for innovative product design based on KJ/FAST/TRIZ methodologies is proposed, to address the requirements for fostering students' abilities to solve complex engineering problems and their higher-order thinking skills. A case study of a household intelligent storage cabinet project completed by students is also presented. The application of integrated project-based teaching using KJ/FAST/TRIZ methodologies can cultivate students' attention to societal needs, implementing a step-by-step progression according to cognitive patterns throughout the entire conceptual design process. This approach enhances the feasibility of innovative design, gradually fostering students' ability to solve complex engineering problems through a gradual deepening of their understanding. Throughout this process, problem-oriented learning encourages students to integrate knowledge from various fields such as engineering, mathematics, and physics to formulate systematic solutions. Simultaneously, project-driven learning through teamwork fosters the development of comprehensive skills including communication, coordination, teamwork, and project management. Establishing regular feedback and reflection mechanisms throughout the learning process aids students in continuously improving their individual thinking and solutions, thereby enhancing self-awareness and self-improvement capabilities.

Compared to existing research, this paper proposes an integrated application of comprehensive innovation methods, presenting a product innovation design process based on KJ/FAST/TRIZ and substantiating its rationality. Oriented toward societal needs, this approach aims to cultivate students' abilities to solve complex engineering problems and meet high-level thinking demands.

This study will provide a certain reference value for promoting engineering certification in related fields and for the construction of First-class courses. In terms of how to better leverage the selection of project themes, the guidance during project implementation, and the assessment of project completion, further practical exploration is still required in the future.

## Acknowledgements

This work was supported by the 2023 Annual University-level 'Undergraduate Teaching Engineering' Project (Guanggongda Jiaozhi [2023] No. 51) at Guangdong University of Technology and the Grant for the Construction of Guangzhou Social Development and Technological Innovation Center (SL2023B04J00003).

## References

Cano-Moreno, J.D., Arenas Reina, J.M., & Sánchez Martínez, F.V. (2022). Using TRIZ10 for enhancing creativity in engineering design education. International Journal of Technology and Design Education, (32): 2749–2774.

Cheng, S.Y., Zhou, J.P., & Yang, J. (2021). Technological Innovation Approach — TRIZ Theory and Application, 2nd ed., Tsinghua University Press, Beijing.

Du, R.X., Du, H., & Fu, X.L. (2022). Design of Multi-Joint Primary and Passive Training Instrument Based on Fast Method. Industrial Design, (09):29-31.

Fiorineschi, L., Frillici, F.S., & Rotini, F., Hsu et al. (2018). Exploiting TRIZ Tools for enhancing systematic conceptual design activities. Journal of Engineering Design, 29(06):259-290.

Geng, J.H., Tian, X.T., & Ma, B.H. (2018). Professional Accreditation-oriented Education Reform of Engineering Practice Courses from Perspective of Industry. Research in Higher Education of Engineering, (02):136-141.

Gui, K., Liu, J.N., & Zhang, W.B., Hsu et al. (2016). Innovative Design of a Multipurpose Desk Based on Modern TRIZ Tools. Packaging Engineering, 37(14):34-37.

He, Y.L. (2022). Functional Design of Gun & Ammunition Cabinet Based on KJ Method and KANO Model. Packaging Engineering,43(14):82-89.

Huang, Z.F., Xie, S.S., & Wang, K., Hsu et al. (2018). Innovative Design for Wire Winding Drums of Reciprocating Wire Cutting Machine Based on TRIZ. Machine Tool & Hydraulics, 46(07):97-101.

Lin, J. (2016). Cultivating the Competence of Solving Complex Engineering Problems by Using PBL. Research in Higher Education of Engineering, (5):17-26.

Liu, F.Q.& Li, L.F. (2022). Product conceptual design based on KJ/Kano/FAST model. Journal of Machine Design, 39(06):149-154.

Mayda, M. & Börklü, H.R. (2014). An integration of TRIZ and the systematic approach of Pahl and Beitz for innovative conceptual design process. International Journal of Technology and Design Education, (36):859–870.

Shi, Y.W.& Pan, S.Y. (2018). Design of Elderly Intelligent Companion Care Robot Based on FAST Theory and TRIZ Theory. Packaging Engineering, 39(16):129-135.

Wu, A.Q., Han, Y.H., & Ye, W.T., Hsu et al. (2022). Research on innovative design of scenic scooter based on AHP / QFD / TRIZ. Packaging Engineering, 43(S1):151-160.

Yang, W.D., Cao, G.Z., & Peng, Q.J. Hsu et al. (2021). Effective radical innovations using integrated QFD and TRIZ. Computers & Industrial Engineering, (162):1–17.

Yuan, X.P., Chen, Y., & Cai, L. (2022). Research on the construction of national offline first-class undergraduate courses based on output orientation. China University Teaching, (09):67-73.

Zhang, F., Wang, L.Q., & Zhang, C.H., Hsu et al. (2023). Teaching research and practice of mechanical design emphasizing students'competencecultivation in the context of new engineering and technical disciplines. Journal of Machine Design, 40(05):156-160.

Zhao, Y.S., Liu, C., & Zhao, C.M. (2019). Cultivating Higher-order Thinking Skills by Project-- based Learning. Research in Higher Education of Engineering, (06):145-148.

## AUTHOR BIOGRAPHIES

Mr. Cheng Siyuan received his doctor's degree in Solid Mechanics from Chong Qing University. He has more than 20 years of experience in teaching. Currently, he is working as Professor at School of Electro-mechanical Engineering, Guangdong University of Technology. His areas of interest include Technological innovation method, Reverse engineering technology, and Mechanical design and simulation.

Dong Jie is a master's student at School of Electro-mechanical Engineering, Guangdong University of Technology, working under the supervision of Professor Cheng Siyuan. His research interests include innovation methods and patent avoidance design.

Mrs. Yang Xuerong received her doctor's degree in mechatronic engineering from Guangdong University of Technology. She has more than 15 years of experience in teaching. Currently, she is working as associate professor at School of Electro-mechanical Engineering, Guangdong University of Technology. Her areas of interest include Technological innovation method, Reverse engineering technology, and machine vision technology.

DOI: 10.6977/IJoSI.202409_8(3).0002
F. Junior, H. Navas/Int. J. Systematic Innovation, 8(3),12-27 (2024)

# Innovative potential for improvements in pellet production: from the perspective of TRIZ and Axiomatic Design

Flavio Numata Junior[1]*, Helena V. G. Navas[2,3]

[1] Department of Mechanical and Industrial Engineering, Nova School of Science and Technology, Universidade NOVA de Lisboa, Portugal

[2] UNIDEMI - Department of Mechanical and Industrial Engineering, Nova School of Science and Technology, Universidade NOVA de Lisboa, Portugal

[3] LASI - Intelligent Systens Associate Laboratory, Universidade do Minho, Portugal

* Corresponding author E-mail: flavio.numata@gmail.com

(Received 22 May 2023; Final version received 01 July 2024; Accepted 16 July 2024)

## Abstract

Improving the efficiency of an energy pipeline industry is complex. Therefore, this article explores the production of pine wood pellets with the application of the Axiomatic Design (AD) and Inventive Problem-solving Theory (TRIZ) techniques, to develop solutions to increase the productive and environmental performance of a pellet industry of Brazil. The results indicate an increase in environmental energy efficiency, in the quality of pellets and possible trade-off situations, creating opportunities for future improvements. This aspect demonstrates that the solutions lead to the evolution of the technical system with potential for innovation.

*Keywords: Axiomatic* Design, Environmental Performance, Pellets, TRIZ.

## 1. Introduction

Forest biomass is considered a renewable energy source that indirectly takes advantage of sunlight, transforming it into carbohydrates through the biochemical processes of photosynthesis (Narodoslawsky, 2010). These carbohydrates are transformed into solid, liquid, and gaseous fuels to produce other types of energy, including electrical energy (Goldemberg, 2009). The total annual global consumption of biomass was estimated at 55 EJ, which represents 10.2% of all primary energy consumed in the world (Edenhofer et al., 2011). Of the total of 4 billion m³ of wood consumed annually in the world, about 55% is used in the form of firewood or charcoal. Unfortunately, the energy efficiency of this form of biomass use is very low, with unused heat loss and high atmospheric emissions. Added to other polluting sources, these environmental effects have continued to expand, requiring improvement in the power energy system.

Accordingly, the European Council on the Environment adopted the “2030 Energy-Climate Framework”, which has three objectives: to reduce greenhouse gas (GHG) emissions by 40% compared to 1990; placing renewable energies at the level of 27% of the European Union's energy consumption and improving energy efficiency by 30% (European Commission, 2020). A great example of sustainable energy products is wood pellets. The source is a type of granulated biofuel of lignocellulosic biomass with excellent energy potential (Li et al., 2013; Kaliyan & Morey, 2009). The raw material is derived from several vegetable sources, such as cereals straw, sugar cane (Nunes et al., 2013; Cortez et al., 2020) and mainly, forest residues, such as tree bark, sawdust. or wood shavings from by-products derived from the timber industries (Chen et al., 2015; Liu et al., 2013). For many years, wood chips have been the most used products in energy systems (Alizadeh & Lund, 2020), but in recent years, pellets have increased their field of application due to their high energy potential (Bioenergy Europe, 2019; European Commission, 2019, Stelle et al., 2011). The pellet production system is also an example of green manufacturing due to the use of raw materials and the process of recovering industrial waste. From this scenario, how could the efficiency of the energy pellet industry be increased in productive and environmental aspects? This problem was addressed in the form of applied research to generate deeper knowledge than a theoretical approach. An exploration in a real context has shown results that enable future applications to improve the efficiency of the industrial plant.

## 2. Industrial Sustainability in Closed-Loop Production of Pellet

In addition to the properties of pellets as renewable materials, other factors can be improved in the phase of their manufacture. A prominent example is the Closed-Loop Production (CLP) system. This manufacturing model makes it possible to recover and insert by-products of the raw material in manufacturing, to reduce the consumption of materials and disposals, and to act systemically in operations and industrial costs (Gu et al., 2016). The CLP system increases the energy and productive efficiency of the manufacturing plant. But this model does not have wide application in industrial areas. The studies also highlight the importance of increasing the energy management of a factory from the stage of energy supply or feedback, to potentiate industrialization (Alizadeh & Lund, 2020; EUROPEAN COMMISSION, 2019; IEA, 2019). By integrating production and energy concepts in the same plant, there are specific variables that make the control and optimization of this type of industry more complex. Within this context, this paper proposes to identify and evaluate solutions that allow to increase the productive and environmental performance of CLP to increase the values of industrial sustainability. The complexity of this subject requires a theoretical approach to the conceptual foundation to identify conditions for improvement. The application of the proposal under operating conditions can define the model.

The pellet production system represents the context described. **Figure** 1 shows a typical pellet production chai, and considers the references of the database called "Industrial wood waste pelletizing, wood pellets A1", which considers industrial waste processing in accordance with DIN EN 14961-2 (2010) European Pellet Council (EPC, 2015). The power plant phase represents of Linear Production (PL) of the pellets system, and input/output flow in the first stage, and goods to the warehouse in the last phase.

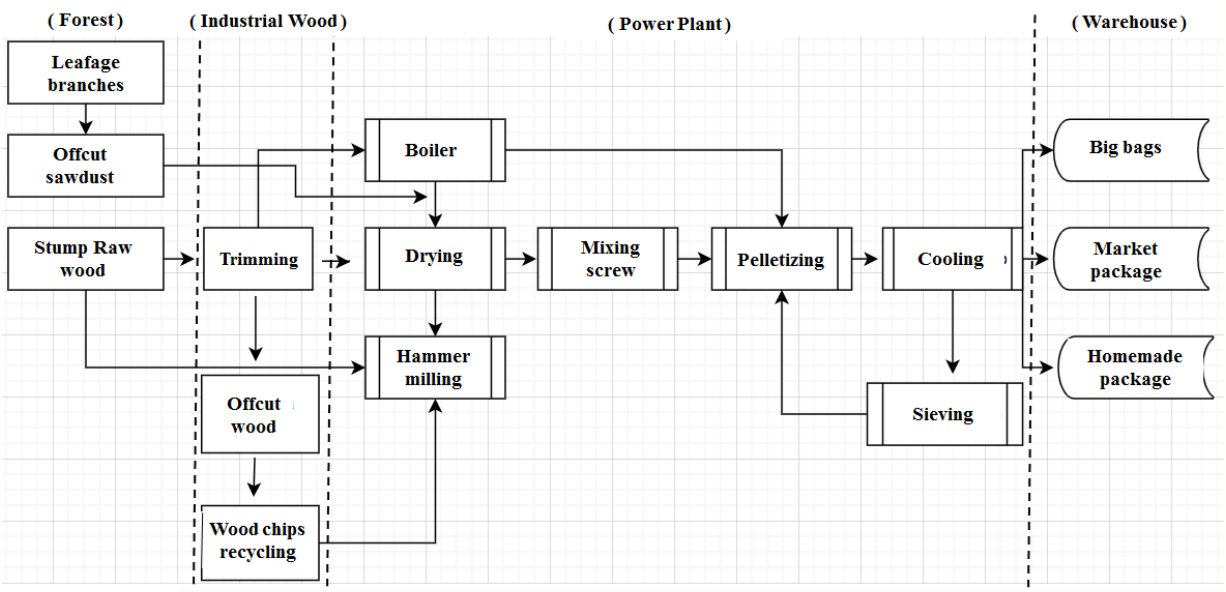



**Fig 1.** Wood pellet production chain.

## 2.1 Axiomatic Design and TRIZ applied in a pellet production system

Some industrial systems for the generation of energy products have an advanced technological level. The CLP model is an example because it reduces environmental effects in the primary phase (raw material) and also in the manufacturing stage by reducing the volume of by-products. Developing improvements to increase production efficiency is very important because it does not require high investments in new projects (greenfield plant), which requires large infrastructure involved and generates process stoppages for the implementation of changes. On the other hand, the improvements developed by remodeling variables of the existing technology, promote the condition of evolution to the industrial system itself, using its resources and process parameters to increase its efficiency. Considering this context, the TRIZ methodology was applied, using the system's PE for evaluation and generation of PI for improvements. In connection with this tool, the concepts of AD were applied, which makes interactions between system operators to improve the structure of the solution for processes and their functions.

For the detailed evaluation of each process, the Axiomatic Design (AD) was applied, which operates as a tool for systematic analysis of complex systems. Initially, AD axioms were considered in manufacturing systems (Suh, 2001) to structure the parameters and information about the product system. Axioms assume the relationship between domains and design elements in the duality between User and User Requirement (RU); Functional and functional requirement (RF); Physical and Design Parameter (PP) and Process and Process Variable (VP) (Harutunian et al., 1996). User Requirements (RUs) present the needs and expectations of users during the product life cycle, so they interact with Functional Requirements (RFs) to form the elementary structure of the basic requirements of the project. The Design Parameters (PPs) are the elements of the design solution in the physical domain, being associated with the Process Variables (VPs) that characterize the manufacturing methods and processes and meet the specified PP, being also established independently of the solution. The axiomatic approach is based on the interaction between RF and physical attributes. Thus, RF is achieved by the interaction between how to develop solutions and conflict or contradictory situations, which are applied by the TRIZ method. However, using this procedure is complex because it involves detailed evaluation in splitting stages of conflicting characteristics in time, space, and the system, and administrative, technical, and physical contradictions (Altshuller, 1998). These forms of contradictions structure a context of ideality and principles of split to develop solutions in the TRIZ model. The separation procedure delimits the physical aspects, the functional requirements and the contradictions to be eliminated, which at TRIZ are worked on through the evolution of the technical system that also involves environmental variables (Altshuller, 1989). For this characteristic, TRIZ's ARIZ 85A technique, analyzes sections to understand the whole, operating by cyclical and continuous assessments to explore the relationships and their hierarchies, as provided by the AD. This process involves studying analogous or non-analogous projects and heuristics in order to develop more creative and less complex ideas (Kwon, Lee & Kim, 2015). TRIZ enables a trend of multidisciplinary evolution, systematically exploring technological evolution for the generation of new solutions (Althsuller, 1999). In this sense, the correlation between AD and TRIZ is presented in table 1:

**Table 1.** Relationship AD and TRIZ.

| Domain specifications | | TRIZ | | | |
|---|---|---|---|---|---|
| DC | Customers | OS | Systems operators | Heuristics | |
| RF | Functional (Functional Requirements) | PE | Engineering Parameters | | |
| PP | Physical (Parameter Project) | ST | Technical Systems | PI | Inventive principles |
| VP | Process (Processes Variables) | PE | Engineering Parameters | | |

The table shows the AD domains for the operators of the analysis system and the variables that lead to the development of heuristics. This logic is consistent with the chained and systematic approach to the development of solutions generated from the Altshuller PI (1998), the Savransky Principles (2003), and the Combined Principles proposed by Mann et al. (2010). Heuristics based on evidence of utility in projects help to develop

solutions (Daly et al., 2012; Tessari & De Carvalho, 2015) according to functions and substances that exist or are absent, to develop the resolution of problems. In this context, the connection between TRIZ's PE and PI with the AA's RF and VP is shown in **Figure 2:**

| Engineering Parameters TRIZ |
|---|
| 1. Weight of moving object |
| 2. Weight of stationary object |
| 3. Lenght of moving object |
| 4. Lenght of stationary object |
| 5. Area of moving object |
| 6. Area of stationary object |
| 7. Volum ov moving object |
| 8. Volum of stationary object |
| 9. Speed |
| 10. Force |
| 11. Stress or pressure |
| 12. Shape |
| 13. Stability of the object's composition |
| 14. Streght |
| 15. Duration of action by a moving object |
| 16. Duration of action by a stationary object |
| 17. Temperature |
| 18. Illumination intensity |
| 19. Use of energy by moving object |
| 20. Use of energy by stationary object |
| 21. Power |
| 22. Loss of energy |
| 23. Loss of substance |
| 24. Loss of information |
| 25. Loss of time |
| 26. Quantity of substance (matter) |
| 27. Reliability |
| 28. Measurement accuracy |
| 29. Manufacturing precision |
| 30. External harm affects the object |
| 31. Object generated harmful factors |
| 32. Ease of manufacture |
| 33. Ease of operation |
| 34. Ease of repair |
| 35. Adaptability or versatility |
| 36. Device complexity |
| 37. Dificulty of detecting and measuring |
| 38. Extent of automation |
| 39. Productivity |

⟺ **Functional Requirements Processes Variables** ⟺

| Inventive Principles TRIZ |
|---|
| 1. Segmentation |
| 2. Extraction |
| 3. Local Quality |
| 4. Asymmetry |
| 5. Consolidation |
| 6. Universality |
| 7. Nesting |
| 8. Counterweight |
| 9. Prior Counteraction |
| 10. Prior Action |
| 11. Cushion in Advance |
| 12. Equipotentiality |
| 13. Do it in Reverse |
| 14. Spheroidality |
| 15. Dynamicity |
| 16. Partial or Excessive Action |
| 17. Transition into a new dimension |
| 18. Mechanical Vibration |
| 19. Periodic Action |
| 20. Continuity of Useful Action |
| 21. Rushing Through |
| 22. Convert Harm into Benefit |
| 23. Feedback |
| 24. Mediator |
| 25. Self-service |
| 26. Copying |
| 27. Dispose |
| 28. Replacement of Mechanical System |
| 29. Pneumatic or Hydraulic Constructions |
| 30. Flexible Membranes or Thin Films |
| 31. Porous Material |
| 32. Changing the color |
| 33. Homogeneity |
| 34. Rejecting and Regenerating Parts |
| 35. Transformation of Properties |
| 36. Phase Transition |
| 37. Thermal Expansion |
| 38. Accelerated Oxidation |
| 39. Inert Environment |
| 40. Composite Materials |

**Fig 2.** TRIZ's PE and PI relationship

## 3. Methodology

The study has an exploratory approach applied to a real experience, to generate a more comprehensive and rodusta investigation of an industrial context. In this way, this work is an applied research that was carried out in the energy products industry. The methodological organization was based on a theoretical basis to demonstrate tools for approaching and recognizing the object of solid study. In the next sections of research, the context and its variables were explored for the application and evaluation of its results. This methodological procedure is designed with the TRIZ technique because it evaluates a condition (problem situation) in order to develop improvements, in a structured and systematic way. Being an applied research, the industrial system is the object of exploration, and TRIZ and AD are the instrumental methods for generating improvements that will be evaluated in the stages of results and conclusion. In summary, the methodological scope associates the flow object of investigation in the following sections of the paper, material, and methods => results => discussions conclusion.

## 4. Material and Methods

### 4.1 Evaluation of the pellet production system

The industrial plant under analysis is located in the south of Brazil, and being one of the main exporters to the European market. The production meets the international standards of the ENplus A1 standard, which certifies high-quality standards throughout the pellet production chain. The production system can adapt according to the technological level and the quality standard of the product to be produced. Brazilian companies adapt their industrialization process to the European standards of the European Committee for Standardization (NREL, 2010) in order to serve the international market with the DIN and ISO quality standards in the ENplus® standard. The pellets are produced from pine wood residues in accordance with the ENplus® standard (EPC, 2015) presented in table 2:

**Table 2.** ENplus® data.

| Parameter | Data reference (units) |
|---|---|
| Diameter | 6 ± 1 ou 8 ± 1mm |
| Length | 3,15 < L = 40mm |
| Umidity | 10% moisture |
| Ash | 0,7% of dry mass |
| Mechanical durability | 98,0 % moisture |
| Temperature | < 40ºC |
| Higher heating value | > 4,6 kWh/kg (≥ 16,5 MJ/kg) |
| Bulk Density (BD) | 600 ≤ BD ≤ 750 kg/m³ |

The evaluation productive system is delimited from "gate to gate", considering only the processes of industrialization of pellets, as shown in **Figure 3**:

The input flow considers the raw material wood residues, energy in the form of electricity, and other substances used in the industrialization process. The outflow is waste or by-products generated. Emissions are also elements of outputs, receiving special observation due to the characteristics of production. As the analysis system considers closed-loop production, emissions from primary processes and production reuse processes are considered. The information data was collected from the Bioenergy Life Cycle Inventory (ICV) of the Swiss Center Ecoinvent, version 2.0 (2007) and technical data on substances from the Pellets Fuel Institute (PFI) (NREL, 2012), from the Bioenergy Europe statistical report on pellets (2019), the European Pellet Council (EPC, 2015), the Handbook on pellets (Obernberger & Thek, 2010) to be able to apply in the analysis considerations.

The general description of the domains in AD om productive system is presented as:

- RF in the functions that the product must fulfill;
- PP in the product design within the specifications of the ENPlus standard padronization;
- VP of production process parameters.

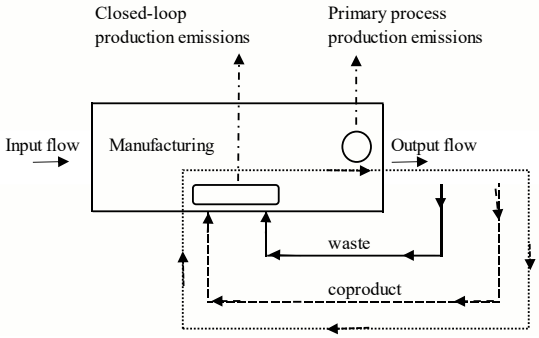



**Fig 3.** Productive system.

The specification of domains originating in:

The PP related to the physical parameters of the product design with its characteristics and properties regulated in the ENPlus standards and in the Handbook on pellets "The production and thermal use of biomass pellets" (Obernberger & Thek, 2010).

For the parameter RF, they refer to the properties and functions of the product and is represented in calorific value, mechanical resistance, smoke emission, odor emission, waste emission, dimensions (length, diameter, and mass), and appearance (shape and colors). These technical data represent the factors that characterize the RF of substances.

The VP comprises the parameters used in industrial devices and processes and refers to the operations of preparing the mixture, grinding, drying, pelletizing, cooling, and dimensional analysis of the product and packaging. These data represent the specific factors of the PV. From this information, with the outline of the factors that associate the RF or VP, the corresponding PIs are identified, forming the correlation matrix between AD and TRIZ to identify the generic solution associated with each element of analysis, shown in **Table 3**:

**Table 3.** AD and TRIZ correlation and generic solutions

| Domaine | RF or VP description | Factors | PI | Description | Generic Solution (based on industrial process data) |
|---|---|---|---|---|---|
| RF | Net calorif value | All substances | 5. Consolidation | Amount of blend components | "Poka-Yoke" system to recipe of components |
| RF | Smoke emissions (-) | Wood | (all substances) | Homogeneous blend of substances | Mixer machine with variable and regenerative cycles of operation |
| RF | Odor emissions (-) | Sawdust | | Recovery emissions | Recovery particulate substances (filter, burn, remanufacturing) |
| RF | Waste emissions (-) | Fluid | | | |
| RF | Dimension (lenght/diameter/weight) | Water | | Dimensional and profile evaluation | Comparision with quality standard |
| RF | Profile (shape and color) | | | | |
| RF | Net calorif value | Substances | 5. Consolidation | Amount of blend components | Mixer machine with variable and regenerative operating cycles |
| RF | Mechanical durability | Blended | 33. Homogeneity | Density bulk | Increase pressing force |
| RF | Profile (shape and color) | Homogeneity | | Homogeneus blend of substances | Increase potential for compaction of substances |
| RF | Net calorif value | Moisture | 29. Pneumatic or Hydraulic Constructions | Mixing process | Mixer machine with variable and regenerative operating cycles |
| RF | Mechanical durability | Moisture | | Density bulk | Increase pressing force |
| | | | | Drying | Increase drying (hot and cool matching (contradiction)) |
| RF | Mechanical durability | Durability | 5. Consolidation | Homogeneus blend of substances | Mixer machine with variable and regenerative operating cycles |
| RF | Dimension (lenght/diameter/weight) | Dimension | 36. Phase Transition | | Increase drying (hot and cool matching (contradiction)) |
| RF | Profile (shape and color) | Aspect | 33. Homogeneity | | |
| RF | Profile (shape and color) | Color | 32. Changing the color | Homogeneus blend of substances | Mixer machine with variable and regenerative operating cycles |
| | | | | | Standardized setup (apply of automation of processes) |
| VP | Setup | Wood | 5. Consolidation | Homogeneus blend of substances | Mixer machine with variable and regenerative operating cycles |
| | | Sawdust | (all substances) | | Standardized setup (apply of automation of processes) |
| | | Fluid | | | |
| | | Water | | | |
| | | Moisture | 35. Physical or chemical transformation | Drying | Increase drying (hot and cool matching (contradiction)) |
| | | Dimension | 17. Transition into a new dimension | Measure and weigh blend of substances | "Poka-Yoke" system to recipe of components |
| | | Homogeneity | 33. Homogeneity | Homogeneus blend of substances | Mixer machine with variable and regenerative operating cycles |
| VP | Milling | Pressure | 36. Phase Transition | Process parameters | Standardized setup (apply of automation of processes) |
| | | Temperature | 36. Phase Transition | Process parameters | Standardized setup (apply of automation of processes) |
| | | Homogeneity | 33. Homogeneity | Homogeneus blend of substances | Mixer machine with variable and regenerative operating cycles |
| VP | Drying | Temperature | 36. Phase Transition | Process parameters | Standardized setup (apply of automation of processes) |
| | | Moisture | 35. Physical or chemical transformation | Drying | Increase drying (hot and cool matching (contradiction)) |
| VP | Pelletization | Pressure | 36. Phase Transition | Process parameters | Standardized setup (apply of automation of processes) |
| | | Temperature | 36. Phase Transition | Process parameters | Standardized setup (apply of automation of processes) |
| VP | Cooling | Temperature | 36. Phase Transition | Process parameters | Standardized setup (apply of automation of processes) |
| | | Moisture | 35. Physical or chemical transformation | Drying | Increase drying (hot and cool matching (contradiction)) |
| VP | Dimensional evaluation (fines) | Dimension | 17. Transition into a new dimension | Fines recovery | Closed-loop pelletizing |
| VP | Packing | Dimension | 17. Transition into a new dimension | Check the weight of the packages | "Poka-Yoke" system to recipe of components |

To understand the potential of the solutions presented, the classification proposed by Altshuller (1998) from TRIZ is considered, with a score that qualifies solutions according to the level of complexity, creativity, and innovation (Navas et al., 2015), which is demonstrated in the following taxonomy of solutions presented in **Table 4**:

**Table 4.** TRIZ level of generical solutions.

| Generic Solution | Level |
|---|---|
| "Poka-Yoke" system to recipe of components | 2 |
| Mixer machine with variable and regenerative operating cycles | 3 |
| Increase drying (hot and cool matching (contradiction)) | 2 |
| Standardized setup (apply of automation of processes) | 2 |
| Closed-loop pelletizing | 3 |
| Increase pressing force | 1 |
| Recovery particulate substances (filter, burn, remanufacturing) | 3 |
| Comparising with quality standard | 1 |

| Engineering Parameters (PE) | Contradictions (C) | Inventive Principles (PI) |
|---|---|---|
| 11. Stress or pressure | Pressure, temperature, shape, energy Restrictive factors | 1, 2, 19, 35 |
| 17. Temperature | Pressure, temperature, shape, energy Restrictive factors | 15, 19, 35, 36 |
| 19. Use of energy by moving object | Blend stability Complexity of control | 3, 10, 13, 36 |
| 20. Use of energy by stationary object | Blend stability Quantity of substance | 35 |
| 26. Quantity of substance | Loss of energy or substance Reliability | 5, 10, 23, 35 |
| 32. Ease of manufacture | Quantity of substance, maintenance Reability | 19 |
| 38. Extent of automation | Adaptability, maintenance | 28 |

**Table 5.** Inventive Principles.

The classification also seeks to stimulate the development of more complete and innovative solutions, in order to achieve the evolution of the system, that is, to seek ideality, which is one of the most important principles of TRIZ (Zlotin et al., 2011). Looking at **Table 2**, there is a greater concentration in levels 2 and 3, therefore, there is an opportunity to develop level 4 solutions, that is, those that are developed from the application of new scientific concepts and are limited to about only 4% of the contingent of solutions generated (Navas et al., 2015). Applying the contradiction matrix, which relates the PE with the contradictory parameters, the following PI are identified in **Table 5:**

From the identification of the PI, the description of the improvements to be applied in the processes is presented:

PI 1: Segmentation

Segmentation suggests a separate analysis of the components of the pellet-forming mixture. The objective is to find out what is the level of emissions of each of the components, to identify which component has the greatest contribution. From this data, investigate which material could be replaced with the same energy potential. The materials are: forest waste; green sawdust and water.

PI 2: Extraction

PI 2 is related to PI 1. Segmentation analyzes the influence of each material and PI 2 acts to remove or reduce the composition of the impacting substance. Observing the composition of the pellet materials (table 5), A1 class it is recommended to reduce to a lower rate than the material composition, main pollutants substances, suggested by the standard are azote: ≤ 0,3 (%), chlorine: ≤ 0,02 (%) and ash: ≤ 0,7 (%). Data presented in table 6:

**Table 6.** Pellets compositions ENplus Standard.

| Substances - EnPlus Classes | | A1 | A2 | A3 |
|---|---|---|---|---|
| Moisture | % | ≤ 10 | ≤ 10 | ≤ 10 |
| Ash | % | ≤ 0.7 | ≤ 1.5 | ≤ 3.0 |
| Bulk Density | kg/ $m^3$ | ≥ 600 | ≥ 600 | ≥ 600 |
| Mechanical Durability | % | ≥ 97.5 | ≥ 97,5 | ≥ 96.5 |
| Net Calorific Value | MJ/kg | 16.5 a 19 | 16.3 a 19 | 16 a 19 |
| Fines | % | ≤ 1 | ≤ 1 | ≤ 1 |
| Sulfur | % | ≤ 0.03 | ≤ 0.03 | ≤ 0.04 |
| Azote | % | ≤ 0.3 | ≤ 0.5 | ≤ 1 |
| Chlorine | % | ≤ 0.02 | ≤ 0.02 | ≤ 0.03 |
| Arsenic | mg/kg | ≤ 1 | ≤ 1 | ≤ 1 |
| Cadmium | mg/kg | ≤ 0.5 | ≤ 0.5 | ≤ 0.5 |
| Chromium | mg/kg | ≤ 10 | ≤ 10 | ≤ 10 |
| Copper | mg/kg | ≤ 10 | ≤ 10 | ≤ 10 |
| Lead | mg/kg | ≤ 10 | ≤ 10 | ≤ 10 |
| Mercury | mg/kg | ≤ 0.05 | ≤ 0.05 | ≤ 0.05 |
| Nickel | mg/kg | ≤ 10 | ≤ 10 | ≤ 10 |
| Zinc | mg/kg | ≤ 100 | ≤ 100 | ≤ 100 |

PI 3: Quality Assurance

The quality of the pellet product and its production process is related to the lower environmental effect generated. Starting from the substances responsible for the discharges identified in PI 1 and 2, the categories of environmental impact classified in the dimension "pollution" are identified according to the standard ISO / TR14047 (US EPA, 2006). The pollution group generates global warming and the depletion of the ozone layer by the CH4, NH3 and CO2 emissions that occur in the processes, under conditions of direct potential impact "+" and indirect potential impact "(+)" is presented in **table 7**:

PI 5: Combination

This PI suggests the combination of operations in time and space with other industrial processes. This solution increases production efficiency by taking advantage of the system's own unused energy sources. It is about identifying passive energy generating processes, in kinetic, mechanical or thermodynamic form. The installation of an energy capture and conversion device is the solution to combine operational efficiency with sustainability. The use of these sources of cogeneration can replace part of the consumption of the original energy

**Table 7.** LCA impact categories

| Categories of impact | General areas of protection | | |
|---|---|---|---|
| | Natural Resources | Human Health | Environmental Health |
| a) Depletion of resources | | | |
| Abiotic resources | + | | |
| Biotic resources | + | | |
| b) Pollution | | | |
| Global warning | | (+) | + |
| Stratospheric ozone depletion | | (+) | (+) |
| Photo-oxidant formation | | + | + |
| Acidification | | (+) | + |
| Eutrophication | | | + |
| Human toxicity | | + | |
| Ecotoxicity | | (+) | + |
| c) Damage to ecosystem | | | |
| Land use | | | + |

unit and reduce the emissions generated, mainly, derived from electricity. The main sources for conversion are mass; heating specific value; power; speed; rotative devices (turbines) or linear motion equipment. In all situations, it is necessary to identify the conversion coefficient to apply to the energy combination system.

PI 10: Prior action

This PI can be applied in two dimensions. A simplified proposal is linked to the "End of Pipe" principle of cleaner production (Kong et al., 2013), with the introduction of filters in the system to avoid discharges and environmental effects. The installation can be applied in the following ways: filters after the process (retention of atmospheric emissions); and filters in the replenishment flow of material recovered from the process.

Another form of prior action would be the specification of standardized methods and processes for the efficient preparation of the mixture, generating a balanced composition of components as the materials are presented in relation to their granulometry or moisture so that the mixture's homogenization or substance addition is increased binder.

PI 13: Inversion

The process of reintroducing materials for reprocessing characterizes the inversion of the material flow because the outlets become the inputs in the process. This form of operation designates the production of closed-loop production, reducing the consumption of primary resources in its material design and also of energy. Recovery involves co-products or materials that have not reached specified quality standards and avoids unnecessary disposal, thus minimizing discharges and emissions during pellet industrialization.

PI 15: Dynamics

The application of PI 15 can be performed through a new specification of the material grinding process, seeking to optimize the use of the primary resource. Right after the grinding process, the materials are inspected using a sieve with the smallest allowed specification of the diameter of the particles to generate dynamism in the use of the materials. This procedure will streamline the frequency of selection of materials and also make better use of the raw material. The processes use mesh for the 8 mm diameter of the particles.

Another possibility would be to recalibrate the operation of the metering valve of the pre-pelletizing mixture, so that the matrix filling is adequate and equivalent to the level of inlet pressure of the pelletizer, as the process is responsible for the fusion of lignin and the union of wood.

PI 19: Periodic action

The variables acting on this inventive solution are temperature and pressure. These parameters are fundamental for productive efficiency in quality and productivity. Installing sensors for remote and periodic control of production conditions enables continuous monitoring of operations. The monitoring will serve to reduce the loss of energy load and keep the process stabilized. Supervisory production systems are complete and automated devices to control and act on production processes.

PI 23: Feedback

Reusing input waste can mitigate waste and pollution. This PI seeks to raise the level of environmental management in accordance with ISO 14000 (ABNT: ISO 14000). The solution is to recover the fines of pellets with a dimension smaller than specified by the standard, reintroducing in the pelletization process what characterizes the feedback principle. The recovery goes through a period of partial reserve of the material and later replenishment. It is important to highlight the possible interference on the property of lignin in agglutinating the material, due to its secondary use, after the material undergoes a mechanical process and at an elevated temperature. The expansion of the pellet by changing the temperature must also be checked, which must be readjusted to homogenize the mixture. The standard ENplus A1 specifies fines content less than 1% (EPC, 2015) of the pellet mass, expressed by the amount of powder present in the sample related to the initial mass (DIN EN 15210-1, 2010) or less than 3, 15 mm (ISO 3310-2: 2010). The granulometric classification sieve must be used to separate the 3.15 mm grains to avoid the destination of very small particles and to generate more ash in the boilers, that is, to avoid the emission of particles in the atmosphere.

PI 28: Replacement of Mechanical Devices

This PI is related to PI 19, which suggests the implementation of automated systems for monitoring and controlling production by supervisory systems. The automated means replace mechanical devices for parameterizing process data, such as pressure and temperature settings. This inventive solution allows greater accuracy in regulation data, real-time monitoring, and remote action if necessary.

PI 35: Change of Physical or Chemical State and PI 36: Phase Change

The change in state is associated with the state of aggregation, concentration, or consistency of the condition of the material. In this sense, this PI is associated with the way to improve the results of productivity and sustainability of production. The first case is related to improving the consistency of the mixture with the introduction of additives for this purpose. Lignin is the main substance responsible for increasing the consistency and energetic power of the pellet, but this substance is already present in the primary raw material. An alternative is to apply the starches derived from corn, potatoes, rice, wheat, or manioc because they are organic derivatives with the presence of natural oil that acts as a lubricant for the pressing matrix (Unpinit et al. 2015). However, it is necessary to perform heating of the additive with water vapor to occur the gelatinization reaction, which consists of a chemical reaction so that the starch granules adhere to the components of the mixture. Another inventive solution would be to cool the pellets by thermal exchange for the mixture. After the pelletization process, the pellets are at a temperature close to 90oC and require cooling. This calorie would be transmitted, through induction, to the mixture, for example, the incorporation of additives.

## 4.2 Trends of Evolution

For Altshuller (1998) the trends in the evolution of technical systems allow to increase the predictability of the technology evolution process. In other words, the analysis of the technological standards of existing products helps to predict the next steps of evolution and to make strategic decisions in the development of other projects. Therefore Altshuller (1998) established eight trends for the evolution of technical systems (Lavengin, 2013):

1. Technology follows a cycle of birth, growth, maturity, and decline;
2. Increasing ideality;
3. Unbalanced development of subsystems, resulting in contradictions;
4. Increased dynamism and controllability;
5. Increasing complexity, followed by simplicity;
6. Joining and splitting system components;
7. Transition from microsystems to macrosystems using energy fields to improve performance and control;
8. Reducing human involvement on devices with increasing automation.

According to the principles of TRIZ systems evolve to increase ideality, being directly proportional to positive attributes and inversely proportional to negative factors and costs (Nakagawa, 2001). In addition to the eight evolution trends mentioned, there are twenty other evolution patterns, called Property Spectrum (SP) that show the trajectory of properties that are altered or transformed when passing from one product to another (Mann et al. 2010): 1. State; 2. Heartbeat; 3. Fragmentation; 4. Surface; 5. Porosity; 6. Automation; 7. Form; 8. Transparency; 9. Color; 10. Coordination; 11. Flexibility; 12. Direction; 13. Integration; 14. Opposites; 15. Market; 16. Information; 17. Symmetry; 18. Fibers; 19. From liquid to spray; 20. From liquid to foam. Following this vector of systematic improvement, Mann (2002) listed 31 Trends of Evolution (ET), presented in **Table** 8 to demonstrate the evolutionary potential of each of the trends:

**Table 8.** 31 Evolution Trends (31 TE) (MANN, 2002)

| Evolutionary trend | | Evolutionary potential |
|---|---|---|
| 1 | Smart materials | 1.1 Passive material 1.2 Adaptable material (one-way) 1.3 Adaptable material (two-way) 1.4 Adaptable material (full) |
| 2 | Space segmentation | 2.1 Solid monolithic 2.2 Hollow structure 2.3 Multiple hollow structure 2.4 Porous structure 2.5 Porous structure with active additives |
| 3 | Surface segmentation | 3.1 Smooth surface 3.2 Surface with rib protrusions 3.3 3D roughened surface 3.4 Roughened surface + active pores |
| 4 | Object segmentation | 4.1 Solid monolithic 4.2 Segmented solid 4.3 Particulate solid 4.4 Fluid 4.5 Segmented fluid 4.6 Gas 4.7 Plasma 4.8 Field 4.9 Vacuum |
| 5 | Evolution macro to nano scale (space) | 5.1 Continuous (at all levels): $10^{2}$ => $10^{1}$ => $10^{0}$ => $10^{-1}$ => $10^{-2}$ => $10^{-3}$ = $10^{-4}$ => $10^{-5}$ ... |
| 6 | Webs and fibres | 6.1 Homogeneos sheet structure 6.2 2D surface regular mesh structure 6.3 3D fibers alignment according load conditions 6.4 Addition of active elements |
| 7 | Decreasing density | 7.1 Continuous (at all levels): $10^{2}$ => $10^{1}$ => $10^{0}$ => $10^{-1}$ => $10^{-2}$ => $10^{-3}$ = $10^{-4}$ => $10^{-5}$ ... |
| 8 | Increasing asymmetry (to match external asymmetries) | 8.1 Symmetrical system 8.2 Partial asymmetry 8.3 Matched asymmetry |
| 9 | Boundary breakdown | 9.1 Many boundaries 9.2 Few boundaries 9.3 No boundaries |
| 10 | Geometric evolution (linear) | 10.1 Point 10.2 1D line 10.3 2D plane 10.4 3D surface |
| 11 | Geometric evolution (volumetric) | 11.1 Planar structure 11.2 2D structure 11.3 Axisymmetric structure 11.4 3D structure |
| 12 | Dinamization | 12.1 Immobile system 12.2 Jointed system 12.3Full flexible system 12.4 Pneumatic system 12.5 Modular system (field) |
| 13 | Action coordination | 13.1 Non coordinated action 13.2 Partially coordinated action 13.3 Fully coordinated action 13.4 Different actions during intervals |
| 14 | Rhythm coordination | 14.1 Continuos actions 14.2 Periodic action 14.3 Use of resonance 14.4 Travelling wave |
| 15 | No linearities (matching to external) | 15.1 Linear consideration of system 15.2 Partial accouting of non linearities 15.3 Full accommodation of non linearities |
| 16 | Mono/Bi/Poly (similar) | 16.1 Mono-system 16.2 Bi-system 16.3 Tri-system 16.4 Poly-system |
| 17 | Mono/Bi/Poli (various) | 17.1 Mono-system 17.2 Bi-system 17.3 Tri-system 17.4 Poly-system |
| 18 | Mono/Bi/Poli (increasing diferences) | 18.1 Similar components 18.2 Components with biased characteristics 18.3 Component plus negative component 18.4 Different components |

| | | |
|---|---|---|
| 19 | Reduced damping | 19.1 Heavy damping 19.2 Critical damping 19.3 Light damping 19.4 Un damped |
| 20 | Increased use of senses | 20.1 1 sense 20.2 2 senses 20.3 3 senses 20.4 senses 20.5 5 senses |
| 21 | Increasing color use | 21.1 No use of color (monochromatic) 21.2 Multicolor 21.3 Visible spectro 21.4 Full spectro |
| 22 | Increasing transparency (seamless) | 22.1 Opaque construction 22.2 Partially transparent 22.3 Transparent 22.4 Active transparente elements |
| 23 | Customer purchase focus | 23.1 Performance 23.2 Reliability 23.3 Convenience 23.4 Price |
| 24 | Market evolution | 24.1 Commodity 24.2 Product 24.3 Services 24.4 Experience 24.5 Transformation |
| 25 | Design point (milestone) | 25.1 Design optimized for single operation 25.2 Design optimized for two operations 25.3 Design optimized for several discrete operations 25.4 Design reoptimized continuosly |
| 26 | Degrees of freedom (DF) | 26.1 1 DF 26.2 2 DF 26.3 3 DF 26.4 4 DF 26.5 5 DF 26.6 6 DF |
| 27 | Trimming | 27.1 Complex system 27.2 Elimination of no key components 27.3 Elimination of non-key sub-systems 27.4 Trimmed system |
| 28 | Controllability | 28.1 Direct control action 28.2 Action through intermediar 28.3 Addition of feedback 28.4 Intelligent feedback |
| 29 | Reducing human involvement | 29.1 Human 29.2 Human + tools 29.3 Human + powered tools + 29.4 Human + semi automated tools 29.5 Human + automated tools 29.6 Automated tools |
| 30 | Design methodology | 30.1 Cut and try 30.2 Steady state design 30.3 Transient effects included 30.4 Slow degradation effects included 30.5 Cross coupling effects 30.6 Design for murphy |
| 31 | Reducing number of energy conversions (zero goal) | 31.1 3 energy conversions 31.2 2 energy conversions 30.3 1 energy conversion 30.4 No energy conversion |

Correlating the 8 Evolution Trends (8 TE) with the 20 Property Spectrum (20 SP) and the 31 Evolution Trends (31 TE), it is possible to verify the innovation potential that may be provided by the evolution process of the technical system with the PI, is presented in table 9:

**Table 9.** 31 Mapping of technical system evolution.

| 8 TE | PI | 20 EP | PI | 31 TE | PI |
|---|---|---|---|---|---|
| 1. Technology follows a birth cycle, growth, maturity and decline | | 1. Physical matter | 35 | 1. Smart materials | |
| 2. Continuous growth of ideality | 15 , 19 , 23 | 2. Pulsation | 3 | 2. Space segmentation | 5 |
| 3. Inequality in the development of sub-systems, resulting in contradictions | 13 , 23 | 3. Fragmentation | 1 , 2 | 3. Surface segmentation | |
| 4. Increasing dynamism and controllability | 15 | 4. Surface | | 4. Object segmentation | 1 |
| 5. Increasing complexity, followed by simplicity | 5 , 10 | 5. Porosity | 5 | 5. Evolution macro to nano scale | |
| 6. Integrating and separating parts of system | 1, 2 | 6. Automation | 28 | 6. Webs and fibres | |
| 7. Transition from microsystems to macrosystems using energy fields to improve performance | 1 , 2 , 35 , 36 | 7. Shape | | 7. Decreasing density | 5 |
| 8. Reduction of human involvement and increasing automation | 19 , 28 | 8. Transparency | | 8. Increasing asymmetry (matching to external) | 10 |
| | | 9. Color | | 9. Boundary breakdown | |
| | | 10. Controllability | 10 | 10. Geometric evolution (linear) | |
| | | 11. Flexibility | 3 , 15 | 11. Geometric evolution (volumetric) | 3 |
| | | 12. Senses | 19 , 23 | 12. Dinamization | 15 |
| | | 13. Integration | | 13. Action coordination | 13 |
| | | 14. Opposites | 2 , 13 , 35 | 14. Rhythm coordination | 10 , 19 |
| | | 15. Market | | 15. No linearities (matching to external) | |
| | | 16. Information | | 16. Mono/Bi/Poly (similar) | |
| | | 17. Symmetry | 36 | 17. Mono-bi-poli (various) | 23 |
| | | 18. Fibres | | 18. Mono/Bi/Poli (increasing diferences) | |
| | | 19. From liquid to spray | | 19. Reduced damping | |
| | | 20. From liquid to foam | | 20. Increased use of senses | |
| | | | | 21.Increasing color use | |
| | | | | 22. Increasing transparency | |
| | | | | 23. Customer purchase focus | |
| | | | | 24. Market evolution | |
| | | | | 25. Design point (milestone) | |
| | | | | 26. Degrees of freedom (DF) | |
| | | | | 27. Trimming | 2 |
| | | | | 28. Controllability | 23 |
| | | | | 29. Reducing human involvement | 28 |
| | | | | 30. Design methodology | |
| | | | | 31. Reducing number of energy conversions | 35 , 36 |

Analyzing the spreadsheet, it is observed that the highest frequency of occurrence of the following evolution trends by evolutionary models are:

- 8 TE: 4 times TE 2 (Ideality) and TE 7 (Transition from microsystems to macrosystems);
- 20 SP: 3 times the EP 11 (Flexibility);
- 31 TE: 2 times the TE 14 (Coordination of rhythms) and TE 31 (Reduction of energy use).

## 5. Discussions and Conclusion

The results show that there are no correlations with the same evolution trends, however, there is a multidisciplinary transformation involving the evolution of the technical system. The changes must start from modifications of specific points to project to larger contexts (micro to macro system), bringing greater capacity for adaptations to processes (flexibility and pace) to minimize industrial interferences and, mainly, environmental effects with greater economy of energy. This set of transformations seeks ideality, which is one of the fundamental principles of TRIZ. Regarding the technical aspects, it is possible that some of the inventive solutions created have specific characteristics, and may generate limitations in other processes (trade-off), according to each type of PI applied. For example, studies indicate that the introduction of waste, related to PI 23, can improve the density of the pellet, but does not increase its mechanical hardness, and may even increase the level of emissions if the reuse content exceeds more

than 50% of reused waste (Miranda et al., 2012; Nunes, 2013).

The PI 35 and PI 36 solutions aimed at the control and calibration of pressure and temperature parameters are determining factors for the production of pellets with quality standards. However, it is important to check the notes in the "Handbook of pellets" (Obernberger & Thek, 2004) which demonstrates the interference of parameters on solid biofuels, in terms of apparent density, moisture and ash contents, calorific value, abrasion, the use of starch additives, the composition of different chemical elements and the presence of heavy metals. The amount of ash is related to the chemical composition of the pellets, which have interference from PI 35 and 36 as well. If there is a high content of ash emission, there will be nitrogen, chlorine, and potassium, and these chemical residues are responsible for the problems of corrosion and accumulation in the flue gas elimination pipes of the process equipment. In this case, the installation of filters reduces the emission of pollutants, as indicated in PI 10. Researches show that pellets derived from eucalyptus wood have an ash content of 0.93%, while pine pellets have an ash content ranging between 0.33 % up to 0.59% (Garcia et al., 2016). Therefore, solutions directed by PI 15 and PI 19 must have control as to the type of material to be replaced.

The moisture content of the lignocellulosic material at the entrance of the pelletizing process directly influences the amount of energy required for the machines. The higher moisture content reduces the energy required but produces pellets with lower density and durability. Moisture decreases friction when the material passes through the pelletizing die holes and also at the back pressure. On the other hand, low moisture content increases the back pressure, resulting in high bulk density and higher energy consumption of the press. The optimum moisture content for the raw material is determined between 10% to 15%, depending on the type of material, as well as the amount of energy consumed during the pelletization process, which should not exceed 4% of the energy contained in the material raw material (Garcia-Maraver & Carpio, 2015). These factors are important for the closed-loop production process due to specific characteristics of the recovery of by-products, mentioned in PI 23. The temperature of the raw material used influences the performance of the production process, therefore, they can be controlled by PI19 and PI 28 The higher the temperature at the entrance of the process, the better result will be achieved in the pressing, resulting in a biofuel with higher density and durability, and less lower energy consumption (Nielsen et al., 2009).

Opportunities for future work

One of the evolutions in pellet production can be made by the torrefaction processes (Pirraglia et al., 2012; SCI, 2018). This process presents level 4 of TRIZ, as the solution is generated from new scientific concepts. The torrefaction is a slow thermochemical process that lasts from 30 to 90 minutes, in an inert atmosphere and temperature ranging between 200 ° C and 300 ° C. The electric power consumption for the grinding of torrefied wood raw material decreases with applied torrefact as a fuel (or additive combustible) in boilers. This process generates volatilization of hemicellulose and changes the properties of biomass, increasing the yield of roasted wood between 66% and 75% (Zwart, 2006; Sklar, 2009). Thus, the torrefied biomass pellet may increase its energy potential by about 1.3 times (Sklar, 2009). In addition to a higher energy density, close to that of mineral coal, between 20-23 GJ per ton, the roasting produces a pellet with low moisture content and low risk of biological degradation. For this reason, the production of roasted wood pellets becomes the object of research and important investments, in the hope of replacing coal in the production of electricity (Pirraglia et al., 2013). Regarding industrialization, torrefaction can be carried out before or after the biomass pelletization process (Ghiasi et al., 2014). When carried out before pressing the mixture, the resulting biomass presents greater water loss and drier, with the possibility of breaking and appearing a darker color. Torrified biomass facilitates pressing, reducing energy consumption by about 70% to 90% during this phase in the pelletization process (Shang et al., 2012; Liu et al.; 2013).

A certain progress is observed in this process, however, the available data on the industrial use of torrefaction technology is very limited. In relation to the economic context, the most effective method of utilization of wood raw material depends on the large-scale production to impact their market value. Even with new industrialization technologies, the production of wood pellets is the leading and most rapidly developing area of utilization of renewable biological sources of energy.

## Acknowledgment

The authors acknowledge the Portuguese Fundação para a Ciência e a Tecnologia (FCT - MCTES) for its financial support via the project UIDB/00667/2020 and UIDP/00667/2020 (UNIDEMI).

## References

Altshuller, G. (1998). 40 Principles: TRIZ, Keys to Technical Innovation. Worcester, MA: Technical Innovation Center.

Altshuller, G., Zlotin, B., Zusman, A. & Philatov, V. (1989). Searching for New Ideas: From Insight to Methodology - The Theory and Practice of Inventive Problem Solving. Kishinev: Kartya Moldovenyaska.

Altshuller, G., Philatov, V., Zlotin, B. & Zusman, A. (1999). Tools of classical TRIZ. Detroit: Ideation International Inc, v. 266.

DALY, S. R., YILMAZ, S., CHRISTIAN, J.L., SEIFERT, C.M. & GONZALEZ, R. (2012). Design Heuristics in Engineering Concept Generation. Journal of Engineering Education. v. 1, n. 4, 601–629, 2012. Available on line at: <doi:10.1002/j.2168-9830.2012.tb01121.x>

EUROPEAN PELLET COUNCIL (EPC). (2015). ENplus - Quality Certification Scheme for wood pellets. ENplus Handbook Part 3: Pellet Quality Requirements. Version 3.0. European Biomass Association (AEBIOM) Brussels, Belgium.

EUROPEAN COMMISSION (2019). Competitiveness of the heating and cooling industry and services. ENER/C2/2016-501. Brussels, Jun, 2019.

GARCIA, D. P., CARASCHI, J. C., VENTORIM, G. & VIEIRA, F.H.A. (2016). Trends and challenges of origin brazilian agroforestry pellets industry. Cerne, v. 22, n. 3, 233-240.

GHIASI, B. ET AL. (2014). Densified biocoal from woodchips: Is it better to do torrefaction before or after densification? Applied Energy, Amsterdam, v. 134, n. 1, 133-142.

GOETZL, A. (2015) Developments in the global trade of wood pellets. U.S. International Trade Commission.

HARUTUNIAN, V. ET AL. Decision Making and Software Tools for Product Development Based on Axiomatic Design Theory. The 1996 CIRP General Assembly in Como, Italy. August, 25-31, 1996. Vol 45/1.

INTERNATIONAL STANDARD ORGANISATION (ISO) ISO 14040. (1997). Environmental management - Life Cycle Assessment—principles and framework; 1997.

KONG, L., XIONG, Y., TIAN, S., LI, Z., LIU, T. & LUO, R. (2013). Intertwining action of additional fiber in preparation of waste sawdust for biofuel pellets. Biomass and Bioenergy, v. 59, p. 151–157.

LIU, Z. ET AL. (2013). Of carbonization conditions on properties of bamboo pellets. Renewable Energy, Amsterdam, v. 51, p. 1-6.

MANN, D. (2002). Innovation, Hands-On Systematic. Creax Press. Ieper, Belgium.

MANN, D. L., DEWULF, S., ZLOTIN, B. & ZUSMAN, A. (2010). Matrix 2010: Updating the TRIZ Contradiction Matrix. Belgium: CREAX Press, 2010.

MIRANDA, T., ARRANZ, J. I., MONTERO, I., ROMÁN, S., ROJAS, C. V. & NOGALES, S. (2012). Characterization and combustion of olive pomace and forest residue pellets, Fuel Processing Technology, 103, 91–96.

NAKAGAWA, T. (2001). Introduction to TRIZ (Theory of Inventive Problem Solving): A technological Philosophy for Creative Problem Solving. In: The Twenty third Annual Symposium of Japan Creativity Society, Japão: Tokyo University, 2001.

NARODOSLAWSKY, M. (2010). Structural prospects and challenges for bio commodity processes. Food Technology and Biotechnology.

NAVAS, H. V., CRUZ-MACHADO, V. & TENERA, A. M. B. R. (2015). Integration TRIZ in Project management processes: na ARIZ contribuction. Procedia Engineering, 131 (2015) 224-231. World Conference: TRIZ FUTURE, TF 2011-2014.

NATIONAL RENEWABLE ENERGY LABORATORY (NREL). (2012). Biomass Resource Allocation among Competing End Uses. Technical Report. NREL/TP-6ª20-5417, May, 2012.

NATIONAL RENEWABLE ENERGY LABORATORY (NREL). 2010. U.S. life-cycle inventory database. https://www.lcacommons.gov/nrel/search Pellet Fuels Institute (PFI). 2011. PFI standards. http://pelletheat.org/ pfi-standards/.

NUNES, L. J. R., MATIAS, J. C. O. & CATALÃO, J. P. S. (2013). Energy recovery from cork industrial waste: Production and characterisation of cork pellets, Fuel, 113 (2013), 24–30.

OBERNBERGER, I.; THEK, G. (2010). The Pellet Handbook – The Production and Thermal Utilisation of Biomass Pellets. London, Routledg, 2010.

PIRRAGLIA, A. ET AL. (2013). Technical and

economic assessment for the production of torrefied lignocellulosic biomass pellets in the US. Energy Conversion and Management, Amsterdam, v. 66, 153-164.

SAVRANSKY, S. D. (2000). Engineering of Creativity: Introduction to TRIZ Methodology of Inventive Problem Solving. Boca Raton: CRC Press.

SHANG, L. ET AL. (2012). Quality effects caused by torrefaction of pellets made from Scots pine. Fuel Processing Technology, Amsterdam, v. 101, 23-28.

SKLAR, T. (2015). Torrefied wood, a bio-energy option that is ready to go": a biomass digest special report. [s. l.: s. n.], 2009. Available online at: <http://biomassdigest.net/blog/2009/12/31/torrefiedwood- a-bio-energy-option-that-is-ready-to-go-a-biomass-digest-special-report/>.

STELLE, W., HOLM, J. K., SANADI, A. R., BARSBERG S., AHRENFELDT, J. & HENRIKSEN, U. B. (2011). Fuel pellets from biomass: the importance of the pelletizing pressure and its dependency on the processing conditions. Fuel, vol. 90, 3285-3290.

SOCIETY OF CHEMICAL INDUSTRY (SCI) (2018). On the Map: The dynamics of the global wood pellet markets and trade – key regions, developments and impact factors. SCI and John Wiley & Songs, Ltd. Biofuels, Bioproducts and Biorefining. (2018), 10.1002/bbb.

Tessari, R. K. & De Carvalho, M. A. (2015). Compilation of heuristics for inventive problem solving. Procedia Engineering, v. 131, 50-70.

Unpinit, T., Poblarp, T., Sailoon N., Wongwicha, P. & Thabuot, M. (2015). Fuel properties of bio-pellets produced from selected materials under various compacting pressure. Energy Procedia, vol. 79, 2015, 657-662.

US Environmental Protection Agency (US EPA). 2006. Life cycle assessment: Principles and practice. US EPA, Washington, D.C. http://www.epa.gov/nrmrl/std/lca/lca.html

Zlotin, B., Zusman, A. & Hallfell, F. (2011). TRIZ to Invent your Future Utilizing Directed Evolution Methodology. TRIZ Future Conference. Procedia Engineering, v. 9, 126-134, 2011. Available on line at:< doi:10.1016/j.proeng.2011.03.106 >

Zwart, R., Boerriger, H. & Van Der Drift, A. (2006). The impact of biomass pre-treatment on the feasibility of overseas biomass conversion to Fischer-Tropsch products. Energy & Fuels, Washington, v. 20, 2192-2197.

## AUTHOR BIOGRAPHIES

**Flavio Numata Junior** is a Researcher at University Center of UniSENAI Blumenau campus, Santa Catarina State, Brazil. Before then, he has 13 years of industrial experience in the companies with leadership in the automotive sector with Renault and Denso. Flavio received his Master of Science degree in Technology from UTFPR, MBA degree from INPG&UNIVALE and graduate Industrial Engineering from UNIMAR (BR). He is currently the Director of University Center of UniSENAI - Blumenau. He also is a Doctoral Industrial Engineering student in the international programme at Nova University of Lisbon (UNL). His research focuses on environmental impacts of industrial process power systems. He is author of research papers in international journals and conference proceedings. His areas of interest include Systematic Innovation including TRIZ, Design & Manufacturing Management and Sustainable Industrial Development.

**Helena Navas** is a Professor at the Department of Mechanical and Industrial Engineering of the NOVA School of Science and Technology - Portugal and researcher at UNIDEMI. She represents the Portuguese Association for Quality (APQ) in the Portuguese Technical Standardization Committee on Research Activities, Development, and Innovation (RDI) and represents FCT NOVA in the Portuguese Technical Standardization Committee on Project Management. She is a Lean Leader and Lean Six Sigma Black Belt Manufacturing for The Lean Six Sigma Company. She is a researcher, consultant and trainer in Innovation, Systematic Innovation, TRIZ Methodology, Lean and Continuous Improvement. Helena Navas has been a guest speaker at various seminars, round tables and events dedicated to innovation. She is a columnist in the Newsletter Innovation & Entreprene

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

# Implementation of lean techniques to reduce mudas in smart tone horn assembly

Thirumurugaveerakumar S[1], Aswin Baalaje R[2]
[1]Associate Professor, Kumaraguru College of Technology, Coimbatore
[2]UG Scholar, Kumaraguru College of Technology, Coimbatore
(Received 22 May 2023; Final version received 01 July 2024; Accepted 04 July 2024.)

## Abstract

In today's competitive manufacturing environment, companies are constantly looking for ways to improve in their production process. A new redesigned assembly line is proposed for bottleneck problems for continuous flow type horn assembly line. We present the improvement of the production rate and balance loss ratio of the manual assembly line. In order to improve the production rate, four different methodologies were proposed and the best one is chosen based on the output.

*Keywords: Lean Thinking, Productivity improvement, Value Stream Mapping.*

## Introduction:

A horn is a sound-making device that can be equipped for motor vehicles, buses, bicycles, trains, trams, and other types of vehicles. The sound made usually resembles a "honk" (older vehicles) or a "beep" (modern vehicles). The driver uses the horn to warn others of the vehicle's approach or presence or to call attention to some hazard. Motor vehicles, ships, and trains are required by law in some countries to have horns. Like trams, trolley cars, and streetcars, bicycles are also legally required to have an audible warning device in many areas, but not universally, and not always a horn. In the horn assembly, a number of components are involved. A core principle in lean methodology is the removal of waste within an operation. And in any business, one of the heaviest drains on profitability is waste. Lean waste can come in the form of time, material, and labor. But it may also be related to the utilization of skill- sets as well as poor planning. In lean manufacturing, there are seven types of waste. In general, customers are not willing to pay for these activities because they do not benefit from them. The lean management and continuous improvement philosophy (Kaizen) attempts to decrease as much waste as possible. Our main objective is elimination of these seven kinds of waste which can help companies reduce costs, increase employee engagement and customer happiness, and increase profits.

## Literature Review:

Lean and agile manufacturing plays a vital role to uplift the production process by lowering manufacturing cost minimal inventory and reducing the cycle time between the processes (C. Hemalatha et. al., 2021). Continuous improvement in the assembly sector can take two different approaches, depending on the investment that one wants or can make, and the depth of the actions that one intends to implement (P. Dias et. al. 2019). Value stream mapping helps in the implementation of new production management concepts - in particular in SMEs (D. Klimecka-Tatar et. al. 2022). The combination of VSM and Computer simulation have a significant effect on decreasing lead time and waste in the color industry ( J. M. Rohani et. al. 2015). VSM used seven tools and concluded that this approach was not only limited to automobile industry but can also be applied for different industries as well (P. Hines et. al.1997). Among all of the methods, VSM extension supports the production planners to choose supportive Smart manufacturing Solutions (N. L. Martin els. 2020). As part of managing the internal environment, senior managers offer support and encouragement. Internal management is a key success factor for companies adopting green practices. (Tharun J et. al. 2022) Event based framework approach helps to tackle the problem of a heterogeneous and incomplete data landscape in the digitalization of value stream mapping

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

(T. Teriete et. al. 2022). The most advantageous aspect of using VSM is that it establishes the possibilities for future advancements (S. M. Zahraee et. al. 2020). The Digital Value Stream Mapping concept is intended to improve the original Lean Value stream method by adding a practical way to assess "digital improvement," which is said to have been overlooked in the original VSM approach (D. Arey et. al. 2021).

## Methodology:

From the literature survey mentioned above, smart tone horn assembly line can be balanced by two methods

1. Traditional method of line balancing
2. Value stream mapping (VSM)

Line balancing was done by these two methods to produce 1350 horns. The results of these methods were compared and analyzed by arena software.

## TRADITIONAL METHOD OF LINE BALANCING

The traditional method of line balancing is to reduce the number of workstations by combining the workstations in a way that will reduce the WIP time. Work In Process (WIP) Work In Progress (WIP), goods in process, or in-process inventory are a company's partially finished goods waiting for completion and eventual sale or the value of these items. These items are either just being fabricated or waiting for further processing in a queue or buffer storage. **Table 1** shows work in progress time is the amount of time the part needs to be waited before operation.

**Table 1.** WIP-Current

| Process no | Process name | Process Time In Secs | Current WIP time in sec |
|---|---|---|---|
| 1 | Spool holder riveting | 16 | 0 |
| 2 | Terminal base diode point plate and point holder assembly | 10 | 16 |
| 3 | Terminal riveting and tuning screwinserting | 26 | 0 |
| 4 | Diaphragm assembly Riveting & height measuring | 14 | 9 |
| 5 | Gasket assembly pre crimpling & final crimpling | 23 | 5 |
| 6 | Air gap measuring and adjusting | 29 | 0 |
| 7 | Pre tuning and mounting bracket assembly | 17 | 0 |
| 8 | Horn tuning and testing | 10 | 13 |
| 9 | Measurement of tuning range and sealant application | 23 | 0 |
| Total Time | | 168 | 43 |
| Total Lead Time | | 211 sec | |

WIP time of 43 sec is non-value-added time.

- Combining process 1 and process 2 will reduce the WIP time of 16 seconds.
- Combining process 7 and process 8 will reduce the WIP time of 13 seconds to 5 seconds.

Implementing the method will eliminate the WIP time of 24 seconds. Total, it will reduce the WIP time form 43 seconds to 19 seconds. It will also reduce the manpower from 13 to 11.The lead time is 211 sec for manufacturing 1150 horns per shift.

**Table 2** Show the WIP Future implementing method and lead time will reduce to considerable amount in seconds.

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

**Table 2.** WIP-Future

| Process no | Process name | Process time in sec | Future WIP time in sec |
|---|---|---|---|
| 1 | Spool holder riveting & terminal base diode point plate and point holder assembly | 26 | 0 |
| 2 | Terminal riveting and tuning screwinserting | 26 | 0 |
| 3 | Diaphragm assembly riveting & height measuring | 14 | 9 |
| 4 | Gasket assembly pre crimpling & final crimpling | 23 | 5 |
| 5 | Air gap measuring and adjusting | 29 | 0 |
| 6 | Pre tuning and mounting bracketassembly & Horn tuning and testing | 17 | 5 |
| 7 | Measurement of tuning range and sealant application | 23 | 0 |
| **Total time** | | **168** | **19** |
| **Total Lead Time** | | **181 sec** | |

By implementing the method, the lead time will reduce to 181 seconds. With a lead time of 181 seconds, it has capacity to produce 1350 horns per shift. The main disadvantages of this method are that the size of the horn testing station will be increased, and the workers will find it too hard to perform the operations.

## Value Stream Mapping (VSM)

Value stream mapping is a method of lean manufacturing which uses symbols, metrics and arrows to show and improve the flow of inventory and information required to produce a product or service which is delivered to a consumer. It is a visual representation which enables one to determine where the waste occurs.

Value stream maps are utilized to assess current manufacturing processes and create ideal and future state processes. It used as a tool which enables a company to map the process flow that helps in identifying various factors like:

- Value added time (time taken for producing the product),
- Non-Value-added time (time taken which do not contribute to the production of end product),
- Cycle time (time required to perform a process) and
- Changeover time (time required to change tool and programming etc.).

This helps in identifying and eliminating mudas (wastes), thereby implementing lean principles. After identifying the non-value-added steps in the current state, a future state value stream map is developed which acts as blueprint for lean activities. The future state value stream map often represents a significant change compared to the way the company currently operates.

The future state value stream map often represents a significant change compared to the way the company currently operates. The value stream map team thus develops a step-by-step implementation strategy to make the Future state a reality.

The key elements of the value stream map are shown:

- The Customer and his requirements.
- Process steps.
- Process Metrics.
- Inventory.
- Supplier with material flows.
- Information and Physical flows.
- Total lead time and Takt time

**Takt time=** Total available production time / Average customer demand.

Current Takt time = (27000/1150) = 23 seconds.

Future Takt time = (27000/1350) = 20 seconds

Future Takt time = (20*0.90) = 18 seconds

Table 3 show the Process Table Current and future TAKT time with manpower.

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

**Table 3**. Process Table - Current

| S.no | Process operations | Process time | WIP time | TAKT time | | Man power |
|---|---|---|---|---|---|---|
| | | | | Current | Future | |
| 1 | Spool holder riveting | 16 | 0 | 23 | 18 | 1 |
| 2 | Terminal base diode point plate and point holder assembly | 10 | 16 | 23 | 18 | 1 |
| 3 | Terminal riveting and tuning screw inserting | 26 | 0 | 23 | 18 | 1 |
| 4 | Diaphragm assembly Riveting & height measuring | 14 | 9 | 23 | 18 | 1 |
| 5 | Gasket assembly pre crimpling and final Crimpling. | 23 | 5 | 23 | 18 | 1 |
| 6 | Air gap measuring and adjusting. | 29 | 0 | 23 | 18 | 1 |
| 7 | Pre tuning and mounting bracket assembly | 17 | 0 | 23 | 18 | 1 |
| 8 | Horn tuning and testing | 10 | 13 | 23 | 18 | 1 |
| 9 | Measurement of tuning range and sealant application | 23 | 0 | 23 | 18 | 5 |

**Figures 1&2** show the Current Process chart and Current Pareto Chart with appropriate numerical values.

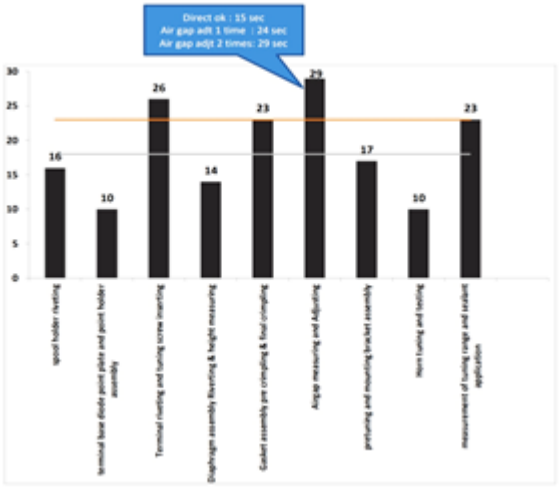



**Fig 1.** Current Process Chart

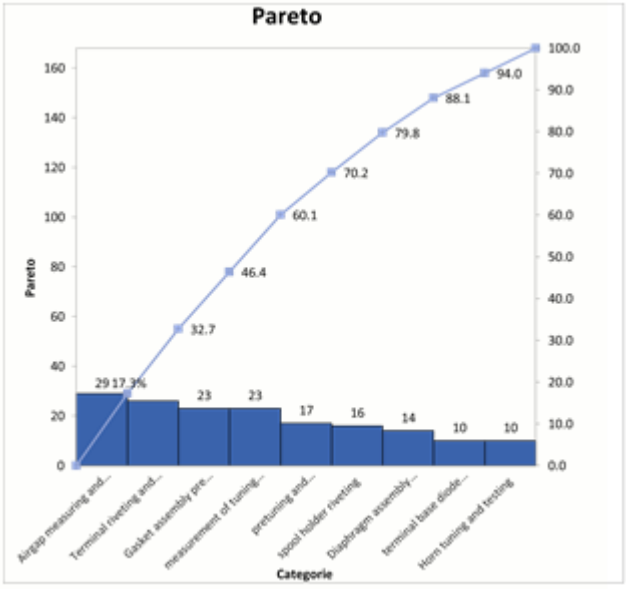



**Fig 2.** Current Pareto Chart

**Figure 3** show the Value Stream Mapping Layout Before Implementation with cycle time

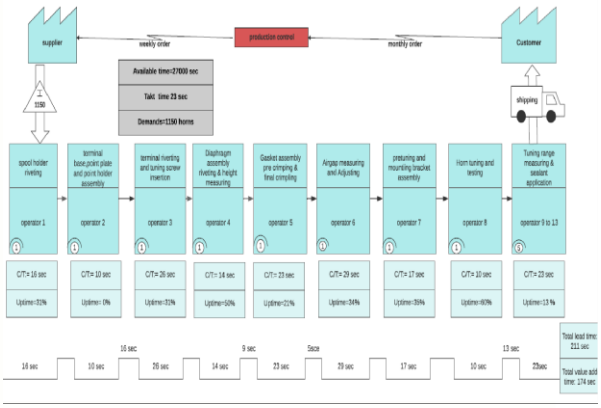


**Fig 3.** Value Stream Mapping Layout Before Implementation.

## Problem identification in VSM

There are four processes that have cycle times greater than that of TAKT Time. They are,

- Terminal riveting and tuning screw inserting.
- Air gap measuring and adjusting.
- Gasket assembly pre crimpling & final crimpling.
- Measurement of tuning range and sealant application.

**Tables 4** & **5** show the ECRS principle and process

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

table in the future. The elemental time was reduced in the following way

**Table 4.** ECRS Table

| Operation | Element | Method |
|---|---|---|
| Eliminate | Process number =3; Element =5; Double green button is pressed by both hands and work piece in right hand is swapped to left hand. | 4 sec can be eliminated by usage of proximity sensors |
| | Process number =6; Element=3; Press the double green button. | 3 sec can be eliminated by usage of proximity sensors |
| | Process number =6; Element 8,10; Using the right hand rotate the rod so air gap will reduce. | Element 8=3 sec and 10=2 sec can be eliminated by operating the machine using foot |
| | Operate the machine using right hand. | |
| | Process number =5; Element= 8;<br>Press double green button using both hands. | The element 8=3 sec can be eliminated by using sensor. |
| | Process number =2; Element=1; Housing is taken in right hand and it changed to left hand. | Interchanging the side can eliminate this element. |
| **Combine** | Process number =3; Element= 3,4;<br>Previous workpiece is removed by right hand.<br>Work piece in left hand is fixed in fixture. | For the elements 3=3 sec & 4=4 sec by combining both the process we can reduce 3 sec in the cycle time. |
| | Process number =3; Element= 7,8;<br>Tuning screw is inserted on screwing machine by right hand.<br>Riveted work piece is inserted<br>in the screwing machine by left hand | For the elements 7=2 sec & 8=4 sec by combining both the process we can reduce 2 sec in the cycle time. |
| | Process number =6; Element= 5,6;<br>Put the mounting washer using right hand.<br>Place it in the conveyer using left hand | Combine the element 5=3sec and 6=2sec so the time can be reduced to<br>2 sec |
| | Process number =5; Element= 5,6,7;<br>Remove the horn placed in pre- crimpling using right hand.<br>Place the new assembly horn in the pre-crimpling machine using left hand.<br>Then remove the horn placed in final crimpling using L hand and place it in conveyer | Combining pre crimpling and final crimpling in a single machine can reduce element time to 5 sec |
| | process number=4; Element =3,4;<br><br>2 SHIM (ss) is taken in right hand and placed over it. Washer is placed over it using left hand | Combing the element 3 and 4 will reduce the element time from 5 sec to 2 sec |
| **Rearrange** | Process number=9; Element=1 to 19;<br><br>1. Measure the tuning range<br>2. Paint the blue ink and apply thread locker<br>3. Paint black ink over the diaphragm<br>4. Imprint batch code<br>5. Test the horn and apply green sticker | This process is done by 5 workers and the times taken are 3.06, 5.32, 5.06, 3.15 and 7.23 seconds.<br><br>After the action analysis of these processes, we found the above processes can be done by 3 persons. It will not become the new bottleneck processes.<br><br>It is done by combining the operation 2 and operation 3 and also operation 4 and operation 5. |
| **Simplify** | Process number=9; element=20;<br>Push the empty tray towards the 1st operator. | Use of gravity conveyer to pass the tray to initial position. |
| | process number=6; Element=1; Horn is taken form conveyer using left hand and changed to right hand. | Element can be simplified by interchanging the side. |

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

**Table 5.** Process Table - Future

| *S.no | Process operations | Process time | IP time | TAKT time | | Manpower | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Current | Future | | | |
| 1 | Spool holder riveting | 16 | 0 | 23 | 18 | 1 | | |
| 2 | Terminal base diode point plate and point holder assembly | 9 | 0 | 23 | 18 | 1 | | |
| 3 | Terminal riveting and tuning screw inserting | 18 | 9 | 23 | 18 | 1 | | |
| 4 | Diaphragm assembly riverting & height measuring | 11 | 0 | 23 | 18 | 1 | | |
| 5 | Gasket assembly pre crimpling & final crimpling. | 18 | 7 | 23 | 18 | 1 | | |
| 6 | Airgap measuring and adjusting. | 21 | 3 | 23 | 18 | 1 | | |
| 7 | Pre tuning and mounting bracketas-sembly | 17 | 0 | 23 | 18 | 1 | | |
| 8 | Horn tuning and test-ing | 10 | 0 | 23 | ]]18 | 1 | | |
| 9 | Measurement of tun-ingrange and sealant application | 19 | 9 | 23 | 18 | 3 | | |

**Figures 4 & 5** show the Future Process Chart and Future Pareto Chart with appropriate numerical values.

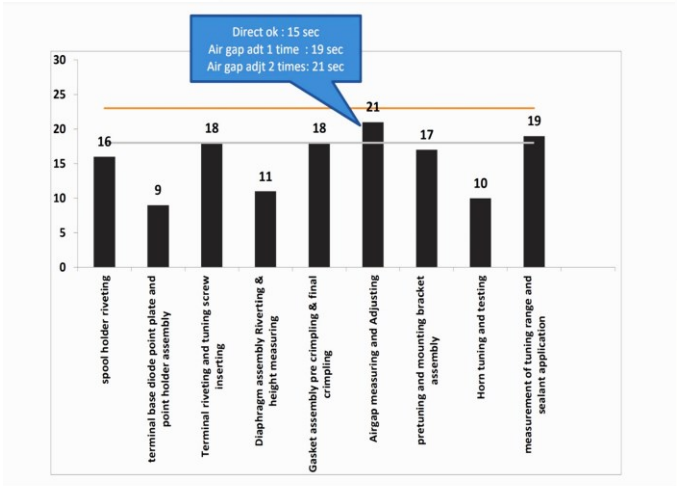



**Fig 4.** Future Process Chart

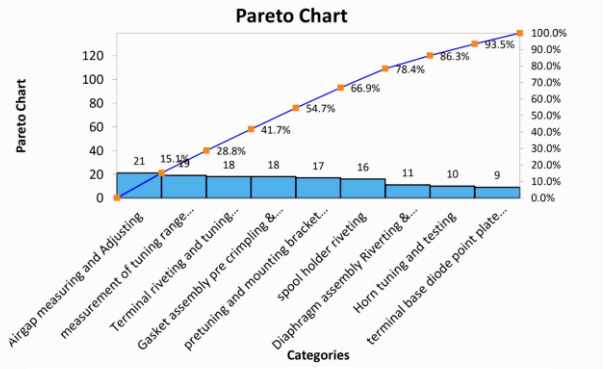



**Fig 5.** Future Pareto Chart

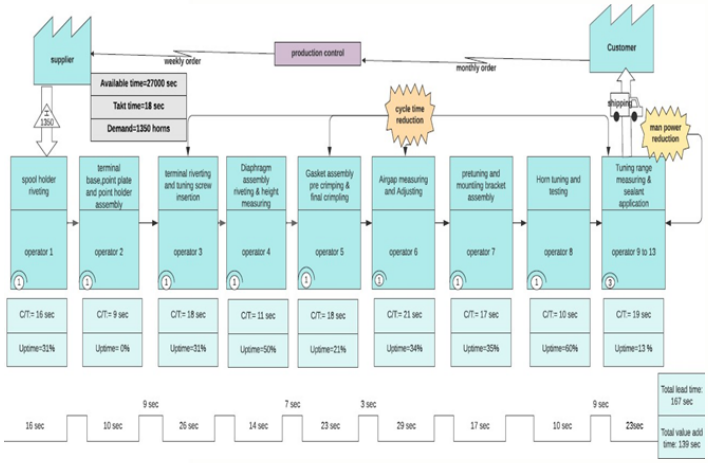



**Fig 6.** Value Stream Mapping Layout After Implementation

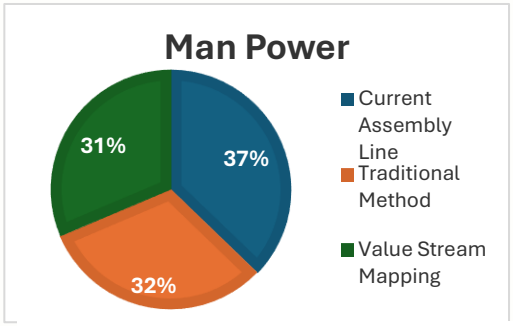



**Fig 7.** Pie Chart Representation of Manpower

## Comparison of four solutions with current assembly line

Here the comparison between the different processes, manpower, lead time, and cycle time are discussed in the **Table 6**, **Figures 7,8,9** respectively to get accurate results regarding the most optimized process to be implemented in the horn industry.

**Table 6.** Comparison of Four Solutions with Current Assembly Line

| Category | Cycle time | WIP time | Lead time | Work-stations | Manpower |
|---|---|---|---|---|---|
| **Current assembly line** | 168 | 43 | 211 | 9 | 13 |
| **Traditionalmethod** | 168 | 19 | 181 | 7 | 11 |
| **Value StreamMapping** | 139 | 28 | 167 | 9 | 11 |


**Fig 8.** Comparison between the Lead time

DOI: 10.6977/IJoSI.202409_8(3).0003
T. S., A. B. R./Int. J. Systematic Innovation, 8(3),28-35 (2024)

## Results:

From the above table and chart, by comparing the two methodologies, it was proven that adapting Value Stream Mapping (VSM) has higher benefits compared to all other methods in smart tone horn assembly line. By implementing this process, time-consuming activities in the assembly line can be reduced. This process not only improves horn production in the assembly line and human intervention is also minimized. If the system is modified and used in assembly line leads to less consumption of time. Combining these two processes reduces the labor cost.

## Conclusions:

In this globalizing era, lean thinking and high productivity are essential for the successful operations of manufacturing firms. Manufacturing firms are in search of various means and ways to reduce resource wastage to enhance operation efficiency and productivity. Even though many articles discuss theories of work-study techniques, few of them deal with the application a implementation of work-study. Even though the proposals and solutions developed in this project are about productivity enhancement in a horn company, the methodology adopted for the problem identification and solution development can be customized to other assembly line-based manufacturing firms to improve operation efficiency and productivity

## References:

Arey, C., Le, D., C. H., & Gao, J. (2021). Lean industry 4.0: A digital value stream approach to process improvement, in Procedia Manufacturing, 2020, vol. 54, pp. 19–24. doi: 10.1016/j.promfg.2021.07.004.

Dias, P., Silva, F. J. G., Campilho, R. D. S. G., Ferreira, L. P., & Santos, T. (2019). Analysis and improvement of an assembly line in the automotive industry, in Procedia Manufacturing, 2019, vol. 38, pp. 1444–1452. doi: 10.1016/j.promfg.2020.01.143.

Hemalatha, C., Sankaranarayanasamy, K., & Durai-raaj, N. (2020). Lean and agile manufacturing for work-in-process (WIP) control, in Materials Today: Proceedings, 2021, vol. 46, pp. 10334–10338. doi: 10.1016/j.matpr.2020.12.473.

Hines, P., & Rich, N. (1997). The seven value stream mapping tools, International Journal of Operations and Production Management, vol. 17, no. 1, pp. 46–64, 1997. doi: 10.1108/01443579710157989.

Klimecka-Tatar, D., & Ingaldi, M. (2022). Digitization of processes in manufacturing SMEs - Value stream mapping and OEE analysis, in Procedia Computer Science, 2022, vol. 200, pp. 660–668. doi: 10.1016/j.procs.2022.01.264.

Martin, N. L., Dér, A., Herrmann, C., & Thiede, S. (2020). Assessment of smart manufacturing solutions based on extended value stream mapping, in Procedia CIRP, 2020, vol. 93, pp. 371–376. doi: 10.1016/j.procir.2020.04.019.

Rohani, J. M., & Zahraee, S. M. (2015). Production Line Analysis via Value Stream Mapping: A Lean Manufacturing Process of Color Industry, Procedia Manufacturing, vol. 2, pp. 6–10, 2015. doi: 10.1016/j.promfg.2015.07.002.

Stasinopoulos, P., Tolooie, A., Abrishami, S. J., Shiwakoti, N., & Zahraee, S. M. (2020). Lean manufacturing analysis of a Heater industry based on value stream mapping and computer simulation, in Procedia Manufacturing, 2020, vol. 51, pp. 1379–1386. doi: 10.1016/j.promfg.2020.10.192.

Teriete, T., Böhm, M., Sai, B. K., Erlach, K., & Bauernhansl, T. (2022). Event-based Framework for Digitalization of Value Stream Mapping, in Procedia CIRP, 2022, vol. 107, pp. 481–486. doi: 10.1016/j.procir.2022.05.012.

Tharun, J., Aswin Baalaje, R., Raj Kumar, R., & Sreeharan, B. N. (2022). A Review On Green Supply Chain Management: Driving And Barrier Factors, International Journal of Engineering Development and Research, vol. 10, no. 3, pp. 95–103, 2022. [Online]. Available: www.ijedr.org

DOI: 10.6977/IJoSI.202409_8(3).0004
S. M. S. Rasoul, F. haghighirad, etc/Int. J. Systematic Innovation, 8(3),36-55 (2024)

# Identification of challenges and solutions for smartphone mobile application development using quality function deployment (QFD): The case of Iran ecosystem

Seyed Mahdi Sadat Rasoul*[1], Farzad haghighirad[2], Seyed Ali Entezar[3]
[1]Department of Operations and Information Technology Management, Kharazmi University, Tehran, Iran
* Corresponding author Email: msadatrasoul@khu.ac.ir

(Received 27 March 2024; Final version received 20 June 2024; Accepted 16 July 2024.)

## Abstract

The growing prevalence of smartphones has led to the emergence of numerous individuals and small companies venturing into the realm of mobile application development (MAD). The MAD ecosystems are highly dynamic, facing constant challenges stemming from financial, social, cultural, and economic factors. Addressing and resolving these challenges is of paramount importance, particularly when considering the involvement of innovative, diverse, and youthful talents in application development.

This paper introduces a novel approach by applying the Quality Function Deployment (QFD) matrix to connect and prioritize challenges and their corresponding solutions within the context of mobile application development in the Iranian mobile application development ecosystem.

The study reveals three primary challenges: "insufficient government financial support," "lack of knowledge in management and deficient management skills," and "absence of meaningful learning from previous experiences and other MAD ecosystems in the international markets." These challenges are identified as crucial hindrances to the effective enhancement of the MAD ecosystem. Consequently, tailored solutions are proposed for each challenge, aiming to address and overcome these obstacles for the sustainable growth of the mobile application development landscape in Iran.

*Keywords: Mobile Application Development, Quality Function Deployment (QFD), Challenges, Solutions, Iranian MAD ecosystem*

## 1. Introduction

The widespread use and accessibility of smartphones, coupled with internet penetration and the adoption of wearable devices, have spurred numerous companies to venture into the development of various types of mobile applications (MAs). The attractiveness of this market is underscored by its current substantial size and a projected high Compound Annual Growth Rate (CAGR) of 14.4% annually in the future. Despite the presence of 11 key players in the market, the diverse needs of users create opportunities for new entrants. Governments can play a pivotal role in supporting these new players through the provision of smart development support packages. However, the intricate challenges and the interdependence of solutions in this market pose a significant risk. Many companies and government sectors, in their pursuit of market share and technological advancements, grapple with erroneous decisions, resulting in setbacks and losses in their strategic plans and investments (Apoorv, 2023).

Statista's projections indicate that the total revenue generated in Mobile Application Development (MAD) is anticipated to reach around 613 billion dollars by 2025. This substantial revenue is derived from diverse sources, including advertising, non-free prepaid downloads, subscriptions, and in-app purchases. In 2021, the estimated number of available applications in various digital markets surpassed six million. Moreover, the projected revenue is expected to exceed 1000 billion dollars by 2023, reflecting the robust growth trajectory of the MAD industry (Statista, 2021, 2023). Globally,

the field of MAD has contributed significantly to employment, with an estimated 30 million individuals engaged in this sector in 2019 (Statista, 2021). To provide a localized perspective, reports from Iran, particularly from platforms like Café Bazaar, reveal that the industry experienced a turnover of more than 350 billion Tomans in 1998. During this period, over 127 thousand applications were available, developed by a total of 67,000 MAD developers since the industry's inception. Of these, 29,000 developers remain active today (Bazaar Café, 2022). To gain a comprehensive understanding of Iran's standing in the MAD industry, the paper explores the ratio of value-added by MAD to the country's GDP. Comparing this ratio with the global average, the study finds that the average mobile application development value added to the GDP ratio index for the world in 2017 was approximately 0.004, whereas Iran reported an index of 0.0003. Despite the existing gap, it's crucial to note that Iran has witnessed exponential growth over the last five years in comparison to the global average. In this period, global application sales increased by 350%, while Iran experienced a remarkable surge of 900%, making it nearly three times faster than the global growth rate (Bazaar Café, 2022).

The primary aim of this paper is to delineate the challenges impeding the success of Mobile Application Development (MAD) in Iran, leading to a shortfall in achieving the targeted financial turnover. Uniquely, this study employs Quality Function Deployment (QFD) to correlate these challenges with tailored solutions. QFD is a robust tool for translating qualitative customer requirements (What) into quantitative engineering characteristics (How). Widely utilized in decision-making for product design and production, QFD facilitates the alignment of customer expectations (What's) with product features (How's) during the design and development process (Pedarpour et al., 2022; Aydarov et al., 2018). The challenges within the MAD sector (What's) are identified through an extensive literature review and interviews. Simultaneously, solutions to these challenges (How's) are determined using the same methodologies. The QFD method is then applied to establish the intricate relationships between the challenges and their corresponding solutions. Careful ranking of these solutions is conducted, providing a structured approach to addressing and mitigating the identified challenges in the MAD landscape in Iran. This methodology allows for a comprehensive understanding of the issues and offers a systematic framework for devising effective solutions to enhance the overall performance and success of the MAD industry in the Iranian context.

The structure of this article is as follows: Section 2 provides the background information on challenges and solutions in Mobile Application Development (MAD). Section 3 outlines the research methodology employed in this study. Moving on to Section 4, the research results and achievements are detailed. Finally, Section 6 concludes the study by summarizing key findings and insights obtained throughout the research process.

## 2. A Brief Review on Research Background & Theoretical background

### 2.1. MAD challenges and solutions

The central focus of this paper is the identification of challenges impeding the growth of Mobile Application Development (MAD) and the provision of effective solutions to overcome these obstacles. The research background section is bifurcated into two key components: challenges and solutions. Furthermore, these challenges and solutions are categorized into distinct groups, encompassing financial aspects, regulatory frameworks, environmental factors, market dynamics, and internal variables within companies, including production and technical or management skills. By comprehensively addressing challenges and proposing specific solutions within these categories, the research aims to contribute to the strategic advancement of the MAD sector, fostering sustainable growth and innovation. This structured approach facilitates a nuanced understanding of the multifaceted challenges faced by the industry and offers targeted solutions for each identified impediment.

Financing has consistently emerged as a pervasive obstacle for various organizations; nonetheless, its significance is notably amplified for companies harboring new and innovative ideas. Investors often grapple with the decision to invest in fields characterized by substantial market ambiguity. Governments, recognizing the inherent risks in such investments, commonly undertake a share of this risk by instituting rules and regulations and allocating funds in these circumstances. This critical interplay between financing, innovation, and risk mitigation has been a focal point in the research endeavors of several scholars. For instance, studies by (Davila, Foster, and Gupta,2003) as well as (Nielsen and Keuschning,2004) have delved into this very issue, shedding light on the challenges and strategies associated with financing innovation in the corporate landscape. These references underscore the importance of navigating financial challenges, especially in industries marked by high levels of uncertainty, and provide

valuable insights for organizations and policymakers grappling with similar concerns.

Marketing and sales are widely recognized as significant hurdles in the development of smartphone apps. Numerous research studies have been conducted to explore potential solutions to the challenges associated with effectively marketing and selling mobile applications. In a study conducted by (Gutierrez and Xu,2006), the critical variables and factors contributing to the success of mobile-related products, services, and apps were examined. The study identified several key factors that play pivotal roles in determining the success of mobile applications: Understanding and addressing these critical variables can significantly contribute to overcoming marketing and sales challenges in the smartphone app development landscape. By emphasizing factors such as ease of access, customer confidence, and usability, developers can create applications that not only meet user needs but also stand out in the competitive market. The findings of studies like the one conducted by (Gutierrez and Xu,2006) provide valuable insights for practitioners seeking effective strategies to navigate the complexities of marketing and sales in the mobile app industry.

The paper also delves into the examination of environmental factors that lie beyond the control of entrepreneurs, posing crucial obstacles to their success. In a comprehensive study, macroeconomic variables, such as bank loan profits, inflation, and potentially various other factors, were identified as influential elements that can either directly or indirectly impact the overall success of entrepreneurs and their projects (Arasti, Zandi, and Talebi, 2012). Furthermore, another study explored the most impactful factors contributing to the success of entrepreneurs in this context. Independence in the country, the availability of fundamental infrastructures and platforms for technology-based product developers, and the absence of teamwork and group activities, as well as a lack of organizational culture, were identified as significant variables (Sadatrasoul et al., 2022; Chitsazan et al., 2017). These findings highlight the complex interplay between external environmental factors and internal organizational dynamics, shedding light on the multifaceted challenges that entrepreneurs face in the pursuit of success. Understanding these factors is instrumental for entrepreneurs and policymakers alike, as it enables the development of strategies that address both external macroeconomic influences and internal organizational aspects to foster a conducive environment for entrepreneurial endeavors.

In a study focused on the mobile entrepreneurship ecosystem in Southern Africa, three primary challenges—Funding, Commercial, and Technical Support—were identified as significant hurdles (Nyamaka, Botha & Biljon, 2018). This underscores the multifaceted nature of challenges faced by entrepreneurs in the region. Another study conducted a comprehensive examination of technical challenges affecting the success of mobile application development. Nine crucial challenges were identified through a literature review, and an additional four challenges were revealed through interviews. These challenges were categorized across three types of mobile applications—native, web, and hybrid—including issues such as fragmentation, testing, reuse of codes, lack of tools support, lack of expertise, compatibility, and security (Ahamd, Feng, Asim & Yousif, 2018). This highlights the diverse and technical nature of obstacles faced in the mobile application development landscape. Additionally, rules, regulations, and legal considerations pose significant obstacles to MAD entrepreneurial activities. A study by Bamshad, (Talebi & Yazdani,2021) emphasized that the lack or inefficiency of precise rules and regulations in new fields in Iran is a major predicament for entrepreneurs. This is exacerbated by the need for multifaceted licenses from various public and government organizations in the MAD ecosystem. Interestingly, this issue is not as prevalent in many developed countries, indicating a unique challenge in emerging markets. Furthermore, (Fartash et al.,2022) extracted five growth challenges specifically for Pioneer Knowledge-based ICT Firms, contributing additional insights into the intricate challenges faced by entrepreneurs in the information and communication technology sector. These collective findings underscore the diverse and complex landscape of challenges in mobile entrepreneurship, spanning financial, technical, regulatory, and growth-related domains. Understanding and addressing these challenges are critical for fostering a supportive environment for mobile application development and entrepreneurship.

Table 1 in the appendix summarizes external and non-technical application development challenges. In the following sections, a list of challenges that have been extracted in literature are presented in **Table 1**. Last four columns of Table 1, represents the four professional experts which have been asked to add and also comment on the necessity or lack of necessity of the challenges in Iran considering that they are extracted from literature studies for the world. It should be emphasized that some challenges have been added by experts, which have not been mentioned in the literature. Considering that this research is exploratory in nature,

and therefore challenges whose existence in Iran was not make sense by experts were removed.

Table 2 in the appendix provides a list of solutions discussed in the literature. In this table, four experts have been asked about the efficiency of the relevant strategy in Iran and add more if they see new solutions by their own ideas and experience. In other words, the evaluation of the challenges and solutions raised in the literature by taking into account the situation in Iran, has been done by four active experts in the industry in **Table 1** and **2**. Finally, the challenges and solutions are mapped using QFD and final results are found.

## 2.2. Quality function deployment (QFD)

Quality Function Deployment (QFD) was originally formulated by Shigeru Mizuno and Yoji Akao in the 1960s as a quality system designed to efficiently deliver services and goods to customers by attentively addressing their needs and ensuring satisfaction (Vinayak and Kodali, 2013). Traditionally, QFD has been employed to systematically translate customer requirements into technical specifications, essentially mapping needs to features (John et al., 2014). Notably, QFD has found application in the domain of mobile government services, playing a role in mapping these services to citizens' needs (Zheng and Pulli, 2012). In studies focusing on government services, such as the work conducted by Alsaadi et al. (2018), QFD has been utilized to identify citizens' requirements through focus group discussions. These requirements were then categorized into distinct classifications, and technical specifications were developed to effectively map them.

The implementation of QFD serves three primary goals:

1.**Identifying Customer/Citizen/Client Needs and Wants:**
The initial step involves comprehensively understanding and articulating the needs and desires of the customers, citizens, or clients.

2.**Translating Needs into Characteristics and Specific Factors:**
QFD facilitates the translation of identified needs into specific technical characteristics and factors.

3. **Building and Delivering a Quality Product or Service:**
The ultimate objective is to create and provide a high-quality product or service by concentrating on customer, citizen, or client satisfaction.

By adhering to these goals, QFD acts as a robust methodology for ensuring that products and services are aligned with the expectations and requirements of end-users. In the context of mobile application development, QFD proves to be a valuable tool for systematically linking challenges to corresponding solutions and prioritizing them based on their impact on customer satisfaction and overall success.

## 3. Research Methodology

In this section, a combination of descriptive, exploratory, and quantitative approaches is employed, leveraging Quality Function Deployment (QFD) with the House of Quality (HOQ) matrix as the primary tool. The HOQ matrix is a widely utilized QFD tool in the literature and serves as the main matrix for QFD methodologies (Alsaadi et al., 2018). The general steps of the HOQ method involve capturing the "Whats" (customer requirements) and linking them to the "Hows" (technical requirements) to address these customer needs.

1. **Phase 1: Develop the Set of Customer Needs (What's):**
The initial phase involves identifying and articulating the set of customer needs or requirements.

2. **Phase 2: Measure Customer Importance/Priority:**
Assign priorities to customer needs, gauging their relative importance or significance.

3. **Phase 3: Identify the Technical Requirements (How's):**
Determine the technical requirements necessary to address and fulfill the identified customer needs.

**4. Phase 4: Create a Correlation Matrix of the Technical Requirements:**
Develop a matrix that correlates the technical requirements, establishing relationships and dependencies.

5. **Phase 5: Create a Relationship Matrix Between Customer Requirements and Technical Requirements:**
Construct a matrix that illustrates the relationships between customer requirements and the identified technical requirements.

6. **Phase 6: Perform Competitive Benchmarking:**
Evaluate and compare the performance of the identified technical requirements against those of competitors in the industry.

**7. Phase 7: Prioritize the Technical Requirements:**

Prioritize the technical requirements based on their correlation with customer needs and competitive benchmarking.

**8. Phase 8: Determine Which Technical Requirement to Deploy:**
Make informed decisions on which technical requirements to prioritize and deploy, ensuring alignment with customer priorities and industry benchmarks.

Furthermore, the methodology of this study benefits from the design of 5 questionnaires aimed at enhancing its argumentative power and credibility, which are respectively:

1. **Phase** 2: Closed-ended format interview questionnaire;

**Phase** 3 Two distinct questionnaires for Delphi method:
"Challenges Questionnaire" comprising 37 factors;

2. "Solutions Questionnaire" comprising 24 factors;

3. **Phase** 4 Pairwise comparison between the identified challenges and solutions questionnaire;

4. **Phase** 5 AHP solution addresses the challenges questionnaire.

By applying the HOQ matrix and systematically progressing through these steps, the research methodology aims to establish a robust framework for linking and prioritizing challenges and solutions in the context of mobile application development in Iran. This approach ensures a comprehensive understanding of customer needs and facilitates the identification of tailored technical requirements to address these needs effectively.

## 3.1. Appling QFD to MAD sector in IRAN

This section customizes and details the eight HOQ steps followed:

**Phase-1**: **Obtain challenges and solutions Through Systematic Literature Review (SLR)**

The Systematic Literature Review (SLR) is a rigorous process designed to systematically identify, select, and critically appraise relevant research (Dewey, A. & Drahota, A. 2016). This study uses this method on challenges and solutions in Mobile Application Development (MAD) before the review is conducted. It entails devising a meticulously planned search strategy with a clear focus or in response to the defined question: "What are the main challenges and solutions in Mobile Application Development (MAD)".

Therefore, Suitable search terms are selected in order to ensure no appropriate article is missed. The search terms used are a variety of combinations of below mentioned words, which are almost equivalent, were searched. In fact, 3 by 5 by 3 (total 45) different search scenarios were explored: (Mobile Application Development = Application Development, Mobile apps development); (Challenges = problems, issues, limitations, obstacles); (Solutions = action, initiative). Boolean operators are used to combine major search terms. Exclusion criteria are articles before 2005 and articles which investigate the technical challenges of MAD. Then the abstract of the articles are read. Those removed articles are which they are entirely irrelevant from their abstract. Finally, a selection from the initially selected shortlist of papers, the decision is made upon reading the entire articles and selecting the ones that meet the selection criteria.

**Phase-2: Data Collection and Finalization and Completion**
**Improvement Via Semi Structured Interviews**

Semi-structured interviews with Mobile Application Development (MAD) experts play a crucial role in gathering insights, validating challenges, and extracting additional information. Semi-structured interviews with experts are held for two primary purposes at this stage. First, these challenges and obstacles of development in the industry are discussed from their perspective, and then the appropriate challenges are confirmed through further studies. Closed-ended format interview questionnaire is used to extract additional challenges/solutions as a data collection tool for gathering data from MAD market/economy experts. Carefully a panel of 18 MAD experts representing diverse roles within the industry is selected. This includes participants from the IRAN Ministry of ICT, the IRAN Vice Presidency for Technology and Knowledge-Based Firms, as well as business development managers and CEOs in the MAD sector. Based on their availability four of them are interviewed scheduled 30–60 minutes.

**Phase-3: Combination of the SLR and the Interview Results Using Delphi Method**

The Delphi process, incorporating insights from new high professional experts along with challenges and solutions from Systematic Literature Review (SLR) and interviews, is employed to iteratively refine the understanding of challenges and solutions in Mobile

Application Development (MAD). Delphi process with the involvement of new high professional experts is initiated. Given pragmatic considerations such as time, access, and expenses, a panel of four experts is chosen. This decision aims for a balance between empirical viability and the richness of expert insights. Two distinct questionnaires are developed, for the first initial iteration of the Delphi process. These questionnaires are crafted to elicit expert opinions and ratings on the identified challenges and solutions in MAD. Then the experts are asked to rate the importance of factors using a Likert scale ranging from one to five. This rating system enables a nuanced evaluation of factors, with a focus on their significance in the context of MAD. Conducting several iterations of the Delphi process, experts are allowed to provide feedback, reassess factors, and refine their opinions. The iterative nature of this approach facilitates consensus-building and refines the understanding of critical factors in MAD. A threshold for factor selection, considering factors with a rating score of more than three as significant is set. This criterion ensures the identification and prioritization of factors that attain a collective consensus of importance from the expert panel. Justifying the decision about the panel size as empirical and pragmatic aligns with the constraints of time, access, and expenses. The selection of four highly professional experts reflects a balance between resource limitations and the need for diverse expertise. By employing the Delphi process with an empirical decision on panel size and iterative refinement, this methodological approach ensures a systematic and consensus-driven exploration of challenges and solutions in the dynamic field of Mobile Application Development.

**Phase-4: Establishing pairwise Comparison of Challenges Using Correlation Matrix**

In this phase, a Likert scale ranging from 1 ('Strongly Disagree') to 9 ('Strongly Agree') is employed to assess and rank each Challenge/Solution. This process involves administering a questionnaire to individuals closely associated with the Mobile Application Development (MAD) industry. The primary objectives are to gauge the perceived importance of challenges and the potential efficacy of solutions in addressing each challenge. Then averaging and then normalization is done one the results. Due to the lack of sufficient space, the relevant table could not be presented in the article.

**Phase-5: Establishing Pairwise Comparison of Solutions Using Correlation Matrix of AHP Method and Extracting the Solution/Challenge Influence**

To gauge the effectiveness of each solution in addressing specific challenges within Mobile Application Development (MAD), the Analytic Hierarchy Process (AHP) is employed. The AHP correlation matrix, utilizing a Likert scale of nine, facilitates a comprehensive evaluation. A Likert scale ranging from one to nine for the evaluation process is employed. This scale allows for a nuanced and detailed assessment of the extent to which each solution addresses specific challenges. Experts assign scores on the Likert scale based on their judgment and experience. Scores provided by experts for each solution is aggregated, then averaging the Likert scale ratings is done. Averaging helps in obtaining a representative and consolidated measure of the perceived effectiveness of each solution. Normalization of the averaged scores for variations in individual expert scoring tendencies, bring scores to a common scale for meaningful comparisons. Then the correlation matrix to quantitatively assess the impact of each solution on specific challenges is utilized. The matrix serves as a valuable tool for decision-makers to prioritize solutions based on their efficacy in addressing critical challenges.

**Phase-6: Creating Pairwise Comparison of Solutions**

During this stage of the methodology, the focus is on conducting pairwise comparisons among proposed solutions within the Mobile Application Development (MAD) context. The objective is to elucidate the extent to which these solutions can interact with each other post-implementation. This comparative analysis provides insights into the potential synergies or conflicts that may arise when multiple solutions are concurrently employed. Respondents are tasked with selecting scores from the range of (-1, 0, +1), each indicative of a specific relationship between two solutions:

- **+1: Alignment and Positive Impact**

Represents an alignment relationship where the two provided solutions exhibit a positive impact on each other. This suggests that implementing both solutions simultaneously results in a mutually beneficial outcome.

- **0: Irrelevancy and Unaffected Solutions**

Indicates that the two solutions are irrelevant to each other, and implementing one does not affect the other. There is no discernible positive or negative impact when these solutions are conducted simultaneously.

- **-1: Inverse Relationship and Potential Negative Impact**

This signifies an inverse relationship where conducting two strategies simultaneously may lead to a potential negative impact. This highlights the importance of considering potential conflicts or counterproductive outcomes when implementing certain combinations of solutions.

The pairwise comparison exercise provides a nu-

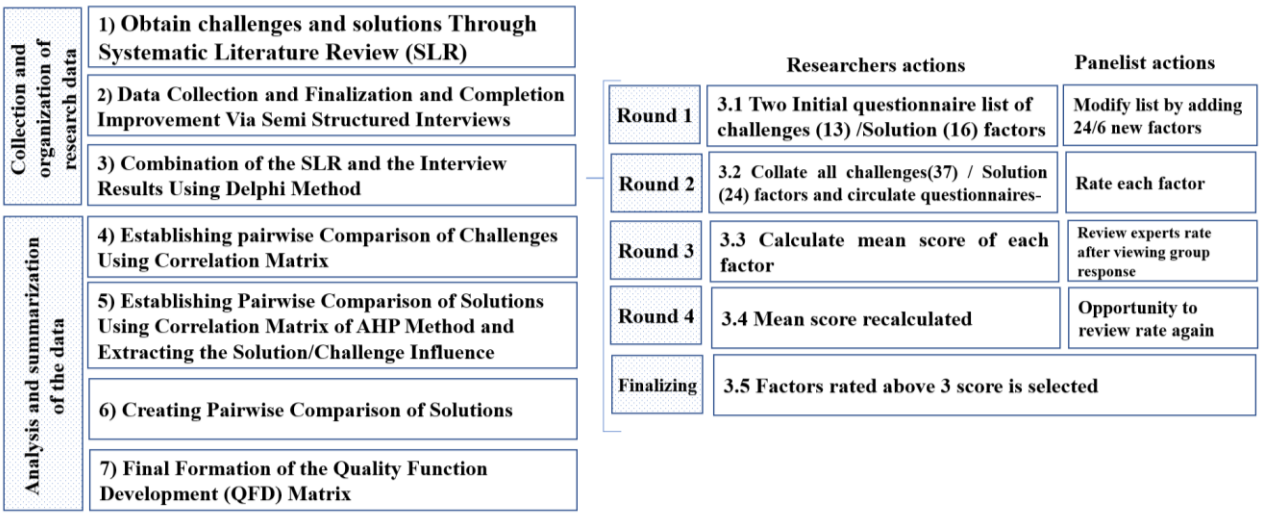



**Fig 1.** The Different Phases of Conducting a Research

anced understanding of the interplay between solutions, enabling stakeholders to make informed decisions about the strategic deployment of multiple solutions within the MAD ecosystem. The resulting scores contribute to a comprehensive view of how different solutions may complement or conflict with each other, guiding decision-makers in optimizing their approach to address challenges effectively. This stage serves as a crucial step in enhancing the strategic planning and coordination of solutions for sustained positive outcomes in MAD.

**Phase-7: Final Formation of the Quality Function Development (QFD) Matrix**

After the normalization process and establishing relationships between solutions and challenges, the next step involves matrix multiplication. This involves multiplying the normalized matrix by the challenge matrix. The outcome of this multiplication provides a precise ranking and prioritization of solutions. From this result, the priority solutions are extracted.

## 4 . Research Results & Discussion

The outcomes of the implementation of the seven phases, as detailed in the previous section, are summarized in this segment.

**Phase-1:** which was comprehensively addressed in the research background section, focusing on Mobile Application Development (MAD) as an official entrepreneurial activity, has yielded a consolidated list of challenges and solutions. The following challenges and corresponding solutions are proposed, approved by experts, and aligned with the comprehensive tables 1 & 2 provided from the background and extracted from the experts can be found in the paper's appendices.

1. The absence of government financing and support (the development of useful and effective financing programs in the form of facilities, grants, and investments for this particular sector);
2. The absence of financial incentives and bonuses from the government (the lack of helpful financial incentives, including taxes, export bonuses, etc.)
3. The absence of collateral (which is often due to their individuality or micro-enterprise nature, they do not have sufficient collateral to use the current support provided);
4. Deficiency in raising capital;
5. The absence of venture capitalists;
6. Significantly high interest rates of bank facilities;
7. Sales & marketing issues;
8. The lack of proper knowledge of the current market (the lack of effective market research reports, etc.);
9. Inappropriate selection of target markets;
10. The absence of transparent and proper rules & regulations (particularly in-app store, applications, etc.);
11. The absence of a proper grading system and the lack of official registration of innovative ideas and copyrights;

12. Non-skilled labor and the absence of sufficient workforce;
13. The lack of experience in developing success tales in the industry;

The proposed solutions derived from the research background segment are further elaborated upon in the following section. It's important to note that these solutions are also documented in **Table 2** in appendix 2, providing a comprehensive reference for readers. Each solution is detailed to provide a clearer understanding of the recommended approaches to address challenges in Mobile Application Development (MAD).

1. Intellectual policymaking in terms of government fundings;
2. Intellectual policymaking by the government in the form of financial bonuses and taxes;
3. Entering and captivating venture capitalists;
4. Independency and internal financing (using the ownership capital for work advancements;
5. Acquiring capital earning skills;
6. The development of facility packages appropriate to the cash flow of field operators by banks (the design of proper interest rates and repayment schedules commensurate with the revenue stream of the developers);
7. Financial government support for high-risk venture capitalists;
8. Lowering the interest rates and establishing equitable interest rates by banks;
9. The establishment of a market recognition system prior to entering product design and development;
10. The development of appropriate copyright laws to protect innovative ideas;
11. Adjustment and adaptation of the current rules & regulations of the system;
12. The development of a mechanism for creating, training, and maintaining specialized workforce and personnel;
13. Acquiring sufficient managerial and entrepreneurial skills before the launch;
14. The attempt to create a networking system as well as a team spirit amongst the staff members and activists in the industry;
15. The documentation and organization of the past activities conducted in the field of entrepreneurship and the obtainment of experience from previous successful and failed projects in terms of counseling;
16. The re-establishment of macroeconomic indicators (such as inflation, exchange rate, interest rates on bank deposits, etc.);

**Phase-2:** To ensure the robustness of the identified challenges and solutions from the initial phase, the outcomes were presented to a panel of four expert interviewees. These experts, possessing both academic and managerial backgrounds, were selected based on their extensive experience in relevant fields:

- **Expert No1**: Over 10 years in the venture capital industry;
- **Expert No2**:15-year experience in the field of IT and Knowledge management;
- **Expert No3**: 15-year experience in the field of IT and Mobile applications;
- **Expert No4**: 12 Years of professional activity in IT industry.

In order to identify challenges and solutions, interviews with experts have been continued until we reach convergence in the answers.

Additionally, the conducted extraction and coding procedures in the interview resulted in newly identified challenges for MAD, which will be further described on the following page:

1. The perspective of maximum potential return on capital in the country and the current parallel markets such as gold, stock exchanges, and many more, in which investments develop significantly higher returns in a short time;
2. High return on capital expectancy in a short period, in addition to high liquidity in such a way that if another suitable investment opportunity is presented, the capital will be directed towards that investment by the investors;
3. Traditional and classical investments in the current well-known industries and the lack of risk-taking to invest in new fields;
4. The limited size of digital domestic markets as well as Persian-speaking market in the entire world (which is approximately estimated at 200 million users), in addition to the absence of reputable digital content publishers with sufficient knowledge on the market and a history of international marketing of digital content and sales of such products;
5. The absence of recognition of most developers of the field such as private individuals, in support programs provided by government agencies, including the Ministry of Information & Communication Technology, the Ministry of Islamic Guidance and Culture, in addition to the Innovation & Prosperity Fund and the National Development Fund and more;
6. The absence of proper laws to protect the rights of the developers and the unbalanced entry of

application distribution network firms, which may potentially target the markets for smartphone app stores (for example Miket, Bazaar Café, etc.), as well as mobile and virtual operators, value-added service companies, in addition to also having a significant impact on both reduction in profit margins as well as the overall motivation of developers to continue their operation in this industry;

7. The collection of excessive taxation from content-based product exporters in the country in case of complete formality and transparent activities;
8. The absence of digital borders and the inability to collect income tax and value-added tax from foreign developers and opportunistic integration of distribution networks with foreign developers;
9. Sanction-related issues, which potentially include the absence of resource obtainment in terms of either currency or branding in the digital space of external mobile apps, including Google App & App Store, in addition to providing open-source software to the publishers;
10. The possibility of a recession;
11. The lack of online payment culture in society for using applications, particularly content-based apps that are used amongst the members of the community;
12. Lack of communications and specialized events such as teamwork events that can significantly aid in the transfer and exchange of experiences, etc.;
13. The absence of proper pricing rules & regulations;
14. The absence of proper knowledge in management on experiences and achievements of activists in Iran and the rest of the world;
15. The lack of creative and innovative individuals of primary products in the following stages;
16. Excessive recognition of the technological layer and neglect in the business part;
17. The process of coding and extraction of the results of interviews has also led to the addition of the following potential solutions for the mobile app industry;
18. The establishment of capital injection and professionalize investors, to create a portfolio of various investments;
19. Financing through crowds based on the type of products available in the industry;
20. Entering more optimal and extensive markets by increasing the overall knowledge and awareness on the targeted markets;
21. The authorization of publishers and directing them to detect crucial and significant exporting target markets;
22. Recognition of individuals and micro-enterprises that produce apps in government and public sector programs;
23. The development of digital rules & regulations to collect revenue and value-added tax from smartphone app developers and their publishers;
24. Assembling specialized events constantly by entrusting them to an expert on teamwork, etc.;

**Phase-3:** The Delphi process initiated with the administration of two distinct questionnaires: the "Challenges Questionnaire" comprising 37 factors and the "Solutions Questionnaire" comprising 24 factors. The participating experts were tasked with reviewing and rating these factors. The mean scores were then calculated, and experts were given the opportunity to review and rate new factors that emerged during this process. An iterative approach was adopted, with mean scores recalculated after each round of expert input. Experts had multiple opportunities to reassess and provide further ratings. The process continued until a final set of factors was identified based on predefined criteria. Finally, the factors' mean scores are calculated and the ones which have the score above 3 are selected. At the end 15 challenges and 15 solutions are selected. Due to the divergent perspectives of experts, their consensus required 4 stages in the Delphi process in order to converge. The W value in Kendall's test or Kendall's W equal to 0.67 indicates a relatively strong agreement between the ratings given by experts. The value of 0.67 indicates that the level of agreement between the judges is generally high, but there are some significant differences in the ratings. This value indicates that the experts agree to a large extent on the evaluation of the items, but this agreement is not absolute and there are still differences of opinion.

**Phase-4:** In this stage of the article, a supplementary questionnaire was administered to facilitate a comprehensive comparison between the identified challenges and solutions. The insights gathered from twenty staff members and activists in the field of Mobile Application Development (MAD) across different age groups, genders, occupations, ethnicities, geographical locations, and income levels in order to lend greater credibility to the research findings are selected. They are questioned crucial in evaluating the perceived importance and effectiveness of each factor. Next, the process of averaging and normalization are also implemented, after the collection of data is complete, resulting in a 15*20-dimension (factors*experts) matrix. **Table 3** shows the final results which are extracted by weighted averaging of the aforementioned matrix. The table encapsulates the

weighted average scores for each challenge and solution, providing a synthesized perspective on their perceived importance and effectiveness. This comparison contributes to a nuanced understanding of the factors, offering insights into their relative significance as perceived by industry stakeholders.

**Table 3.** The Complete & Comprehensive Table of Pairwise Comparisons

| No | Challenge Name | Normal weight | Questionnaire sequence number |
|---|---|---|---|
| 1 | Lack of government financial support | 0.0755 | 7 |
| 2 | Lack of knowledge management implementation | 0.0708 | 14 |
| 3 | Weak managerial skills | 0.0702 | 6 |
| 4 | Uneconomical plans | 0.0681 | 5 |
| 5 | Lack of competitive market | 0.0678 | 4 |
| 6 | Weaknesses in capital raising skills | 0.0675 | 3 |
| 7 | Lack of risky investors and lack of support for them | 0.0669 | 9 |
| 8 | Lack of comprehensive team building events in the field of content | 0.0666 | 13 |
| 9 | Lack of clear rules and regulations | 0.0661 | 11 |
| 10 | No collateral | 0.0654 | 1 |
| 11 | Lack of proper knowledge of the market | 0.0649 | 15 |
| 12 | Not so big domestic markets and Persian language markets | 0.0644 | 2 |
| 13 | High bank rates | 0.0632 | 10 |
| 14 | Lack of government financial incentives | 0.0628 | 8 |
| 15 | Problems of sanctions for entering the foreign market and payment sphere | 0.0599 | 12 |

**Phase-5:** In this stage of the study, the Analytic Hierarchy Process (AHP) correlation matrix was employed to assess the extent to which each solution addresses the identified challenges. The Likert scale ranging from 1 to 9 was utilized for this purpose, and the results were subjected to averaging and normalization. Considering that the numbers of the AHP questionnaire scale are between 1 and 9 and this part of the QFD matrix has three states (high influence, normal influence, and no influence), it was decided that in the final HOQ, the numbers between 0 and 3 as the lower limit, the numbers between 3 and 6 as the middle limit and numbers between 6 and 9 as the upper limit of the QFD matrix.
Also, table (4) provides valuable information on the precise impact of each challenge and its appointed solutions, to help determine the capability of the solution against challenges; besides, all the interviewees are also asked to purposefully fill in the required cells in the table using this method. Additionally, the pairwise comparison approach is once again implemented in Table 4, which functions as an abstracted version of the results.

**Phase-6:** In this phase of the study, an essential table was crafted to unravel the relationships among solutions. Participants, including experts and respondents, were solicited to assign a grade between (-1, 0 & +1) based on their perceived interactions between different solutions. The culmination of these responses was averaged, and a triple interval between (-1 & +1) was established to denote the relationships. Furthermore, the conclusion is formed by calculating the average of the obtained responses and conducting a triple interval between (-1 & +1), as summarized and shown in the following Table 5. Additionally, the Inconsistency Ratio is also calculated to be 0.06, which is lower than 0.1 and can potentially indicate a slight incompatibility amongst the implemented comparisons

**Table 4.** The Prioritization of the proposed solution

| No | Solution name | Normal weight | Questionnaire sequence number |
|---|---|---|---|
| 1 | Creating a knowledge management system | 0.079258 | 12 |
| 2 | Government support for venture capitalists | 0.074518 | 4 |
| 3 | Develop the necessary rules and regulations to register ideas | 0.073616 | 8 |
| 4 | Develop appropriate laws and regulations and amend laws | 0.072311 | 5 |
| 5 | Acquire the necessary managerial and entrepreneurial skills | 0.072306 | 6 |
| 6 | Acquire the necessary skills in raising capital | 0.071215 | 11 |
| 7 | Reducing interest rates on loans by banks | 0.070077 | 9 |
| 8 | Detailed knowledge of the market before entering | 0.069380 | 10 |
| 9 | Internal financing | 0.068877 | 7 |
| 10 | Perform feasibility studies before starting work | 0.063858 | 14 |
| 11 | Entry of venture capitalists | 0.062758 | 1 |
| 12 | Government protection policies | 0.062488 | 2 |
| 13 | Government financial incentive policies | 0.059367 | 3 |
| 14 | Reform of the prevailing climate (economic, sanctions) | 0.051085 | 15 |
| 15 | Moving towards team building in the field of content | 0.048888 | 13 |

**Table 5.** The correlation matrix of Solutions

| The Relationship Amongst Solutions | Entry of venture capitalists | Government protection policies | Government financial incentive policies | Government support for venture capitalists | Develop appropriate laws and regulations and amend laws | Acquire the necessary managerial and entrepreneurial skills | Internal financing (owner capital) | Develop the necessary rules and regulations to register ideas | Reducing interest rates on loans by banks | Detailed knowledge of the market before entering | Acquire the necessary skills in raising capital | Creating a knowledge management system | Moving towards team building in the field of content | Perform feasibility studies before starting work | Reform of the prevailing climate (economic, sanctions) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Entry of venture capitalists | ۱ | | | | | | | | | | | | | | |
| Government protection policies | ۱ | ۱ | | | | | | | | | | | | | |
| Government financial incentive policies | ۱ | ۱ | ۱ | | | | | | | | | | | | |
| Government support for venture capitalists | ۰ | ۰ | ۰ | ۰ | | | | | | | | | | | |
| Develop appropriate laws and regulations and amend laws | ۰ | ۰ | ۰ | ۰ | ۰ | | | | | | | | | | |
| Acquire the necessary managerial and entrepreneurial skills | -۱ | -۱ | -۱ | -۱ | ۰ | ۱ | | | | | | | | | |
| Internal financing (owner capital) | ۰ | ۰ | ۰ | ۰ | ۱ | ۰ | ۰ | | | | | | | | |
| Develop the necessary rules and regulations to register ideas | ۰ | ۰ | ۰ | ۱ | ۰ | ۰ | ۰ | ۰ | | | | | | | |
| Reducing interest rates on loans by banks | ۰ | ۰ | ۰ | ۰ | ۱ | ۰ | ۰ | ۰ | ۰ | | | | | | |
| Detailed knowledge of the market before entering | ۱ | ۰ | ۰ | ۱ | ۰ | ۱ | ۰ | ۰ | ۰ | ۰ | | | | | |
| Acquire the necessary skills in raising capital | ۰ | ۰ | ۰ | ۰ | ۰ | ۱ | ۰ | ۱ | ۰ | ۰ | ۱ | | | | |
| Creating a knowledge management system | ۰ | ۰ | ۰ | ۰ | ۰ | ۱ | ۰ | ۰ | ۰ | ۰ | ۰ | ۱ | | | |
| Moving towards team building in the field of content | ۰ | ۰ | ۰ | ۰ | ۰ | ۱ | ۰ | ۰ | ۰ | ۱ | ۰ | ۱ | ۰ | | |
| Perform feasibility studies before starting work | ۱ | ۰ | ۰ | ۰ | ۱ | ۰ | ۰ | ۰ | ۰ | ۰ | ۰ | ۰ | ۰ | ۰ | |
| Reform of the prevailing climate (economic, sanctions) | | | | | | | | | | | | | | | |

**Phase-7:** Quality Function Deployment (QFD) Matrix Analysis is used to achieve synergy. In this pivotal step, the culmination of calculations and analyses from preceding phases manifests in the establishment of the Quality Function Deployment (QFD) matrix, as showcased in Table 6. This matrix serves as a comprehensive representation of the research achievements, encapsulating the intricate relationships among challenges and proposed solutions.

**Key Elements:**

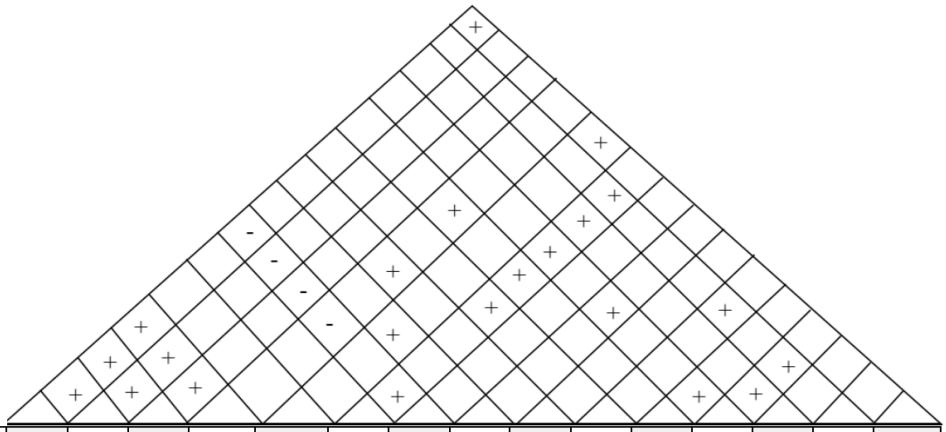
- **Challenges and Weights:** The left column of the middle matrix delineates the identified challenges, each accompanied by its respective weight. This weight signifies the significance of each challenge in the overall MAD ecosystem.
- **Solutions and Weights:** The upper row of the middle matrix presents the proposed solutions; each assigned an importance weight. These weights reflect the perceived efficacy and relevance of each solution in addressing challenges.
- **Triangular Matrix:** Positioned at the top, this matrix outlines the proposed solutions and their overlap weights. It sheds light on the potential

interchangeability of solutions, showcasing the dynamic relationships among them.

- **Central Matrix:** At the core of the QFD matrix, this section denotes the effectiveness of implementing solutions in tackling one or more challenges. The anticipation is that the application of specific solutions will lead to the mitigation of related challenges, fostering a holistic improvement in the MAD ecosystem.

This QFD matrix is a valuable tool for strategizing and implementing targeted interventions, ensuring a harmonized approach to fortify the Mobile Application Development sector in the face of diverse challenges.

**Table 6.** The Finalized Matrix of the Research & Quality Function Deployment (QFD) Describing the Relationships Between Solutions & Challenges



| QFD | Weight of challenges | Entry of venture capitalists | Government protection policies | Government financial incentive policies | Government support for venture capitalists | Develop appropriate laws and regulations and amend laws | Acquire the necessary managerial and entrepreneurial skills | Internal financing (owner capital) | Develop the necessary rules and regulations to register ideas | Reducing interest rates on loans by banks | Detailed knowledge of the market before entering | Acquire the necessary skills in raising capital | Creating a knowledge management system | Moving towards team building in the field of content | Perform feasibility studies before starting work | Reform of the prevailing climate (economic, sanctions) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No collateral | 0.0654 | 9 | 9 | 9 | 9 | 3 | 3 | 9 | 3 | 9 | 3 | 9 | 3 | 1 | 3 | 3 |
| Not so big domestic markets and Persian language market | 0.0644 | 1 | 3 | 3 | 3 | 1 | 1 | 3 | 3 | 3 | 9 | 3 | 3 | 1 | 3 | 1 |
| Weaknesses in capital raising skills | 0.0675 | 3 | 9 | 9 | 9 | 3 | 9 | 9 | 3 | 9 | 3 | 9 | 3 | 1 | 3 | 3 |
| Lack of competitive market | 0.0678 | 1 | 3 | 3 | 3 | 9 | 3 | 3 | 9 | 3 | 9 | 1 | 3 | 3 | 3 | 3 |
| Uneconomical plans | 0.0681 | 3 | 1 | 1 | 3 | 1 | 9 | 3 | 3 | 3 | 3 | 3 | 9 | 9 | 9 | 1 |
| Weak managerial skills | 0.0702 | 3 | 1 | 1 | 3 | 3 | 9 | 3 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 1 |
| Lack of government financial support | 0.0755 | 3 | 9 | 9 | 9 | 3 | 3 | 3 | 3 | 9 | 3 | 9 | 3 | 1 | 1 | 1 |
| Lack of | 0.0628 | 3 | 9 | 9 | 9 | 3 | 3 | 3 | 3 | 9 | 3 | 9 | 3 | 1 | 1 | 1 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| government financial incentives | | | | | | | | | | | | | | | | |
| Lack of risky investors and lack of support for them | 0.0669 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 3 | 9 | 3 | 1 | 9 | 9 |
| High bank rates | 0.0632 | 3 | 3 | 3 | 3 | 3 | 3 | 9 | 3 | 9 | 3 | 3 | 3 | 1 | 3 | 3 |
| Lack of clear rules and regulations | 0.0661 | 1 | 1 | 1 | 3 | 9 | 1 | 3 | 9 | 1 | 3 | 3 | 3 | 1 | 3 | 3 |
| Problems of sanctions for entering the foreign market and payment sphere | 0.0599 | 1 | 1 | 1 | 1 | 3 | 3 | 3 | 1 | 1 | 3 | 3 | 3 | 1 | 1 | 3 |
| Lack of comprehensive team building events in the field of content | 0.0666 | 1 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 3 | 9 | 1 | 1 |
| Lack of knowledge management implementation | 0.0708 | 3 | 1 | 1 | 3 | 3 | 9 | 1 | 3 | 1 | 3 | 1 | 9 | 3 | 3 | 1 |
| Lack of proper knowledge of the market | 0.0649 | 3 | 1 | 1 | 3 | 9 | 3 | 1 | 3 | 1 | 9 | 3 | 9 | 1 | 9 | 1 |
| Weight of solutions | | 0.06276 | 0.06249 | 0.05937 | 0.074178 | 0.0723108 | 0.07231 | 0.06888 | 0.07362 | 0.07008 | 0.06938 | 0.07121 | 0.07926 | 0.04889 | 0.06386 | 0.051085 |

**Discussion, Evaluations, and Results: Navigating the Landscape of Mobile Application Development (MAD) Challenges and Solutions**

In this comprehensive article, the landscape of non-technical challenges and solutions within Mobile Application Development (MAD) is meticulously explored. The findings provide valuable insights into key sectors where real-world experts and regulators can channel their efforts for substantial enhancements in the efficiency and success of MAD. The discussion revolves around addressing the research questions posed throughout the study:

**• RQ1) Most important challenges and solutions.**

The identification of challenges and solutions brings to light five paramount challenges: "The lack of proper management knowledge on the achievements and experiences of other activists in the industry", "the absence of sufficient financial support from the government", "deficiency in management skills", "inefficient projects due to various factors" and "the lack of a competitive market". On the other hand, the best performing solutions and strategies include: "government support for venture capitalists", "the establishment of a proper knowledge management mechanism", "generating the required rules & regulations to protect innovative ideas", "the construction of various programs for managerial and entrepreneurial skills acquirement", and "the adjustment of the current rules & regulations".

**• RQ2) Major challenges can be meaningfully separated from lesser important challenges and focused more.**

The examination of the weight importance of challenges, as illustrated in Table 3, reveals a nuanced landscape where the weights are closely aligned. This close proximity among the weights signifies that challenges cannot be easily separated into two distinct groups of more and less important challenges. The intricate interplay and interconnectedness of these challenges underscore the complexity of the Mobile Application Development (MAD) ecosystem, emphasizing the need for a comprehensive and holistic approach to address the array of challenges it presents.

**• RQ3) By focusing on a specific and limited number of effective solutions, it becomes feasible to exert influence over a substantial array of important challenges.**

Among the 15 solutions outlined, the top seven solutions, characterized by weights exceeding 0.07, emerge as pivotal. These solutions are strategically significant, offering pathways to meaningfully address and alleviate a range of challenges within the Mobile Application Development (MAD) ecosystem. This targeted approach emphasizes the potential for concentrated efforts to yield impactful outcomes across various facets of the MAD landscape.

**RQ4) There are transformational solutions, from which implementation of them can create in solving several challenges.**

There are not one nor few solutions in which they can solve the several challenges in a transformative manner.

**• RQ5) Balance must be observed to solve challenges and it is not necessarily possible to create great growth capabilities by solving some challenges in just one dimension.**

Analyzing the problems and the solutions shows that a balance is needed in order to enhance the IRAN MAD ecosystem.

**• RQ6) There is a sequence to run the solutions or they can be done in parallel and with a spiral and evolutionary method.**

Analyzing the solutions shows that most of the solutions can be run in parallel and therefore a major improvement in IRANs MAD ecosystem can take place in the years ahead.

## 5 – Conclusion

Current research contributes to the field of MAD stockholders and government organizations and administers with a comprehensive view of challenges and implementation solutions in real-world. It also contributes with an analysis framework that can be used to derive such insights in other cases. It can also contribute to be used as a basis for formulating effective policies in the field of MAD. The aim of this research was to identify challenges as well as appropriate solutions for MAD in Iran. Towards this aim, QFD is used as a matchmaking framework and the challenges and solutions of the MAD sector were identified with proper data gathering from previous papers, research, websites, news, and finally experts.

The research proposed an integrated framework, which takes a broad perspective on extracting challenges and solutions, integrates Financing, Market, Rules & Regulations, Environmental, Technical & Production and Internal elements and explicitly considers the need for alignment between these elements. Using the framework to analyze Iran MAD, we identified 15 challenges and 15 solutions

Additionally, in regard to the relationships amongst the proposed solutions, it is concluded that the implementation of the solutions, including government financing policies, government incentive policies, obtaining the required skills, the removal of obstacles such as sanctions, and the entry of venture capitalists all can potentially be significantly beneficial for the growth of the industry. Moreover, research showed that some challenges and solutions are interrelated across the different alignment types. This implies that, when proposing implementation of a solution to a specific challenge, it is important to also assess its effects on other parts.

### 5.1 Theoretical Contributions

The paper adds value to the QFD literature and provides new insight into government development plans which need challenge recognition and ranking and solution extraction and ordering their implementation importance. QFDs engineering approach is applied to translate the MAD industries activist as a voice of citizens to the government solution implementation. SLR is used to extract the challenges in the global and local MADs and the proper solutions are designed after ranking the challenges.

### 5.2 Practical Contributions

The QFD-AHP-Delphi method provides an effective approach for government authorities, other decision-makers, and MAD activist to recognize which factors require more attention and develop a plan to result in the highest possible MAD activist's satisfaction. In fact, the paper offers an understanding of citizens' perceptions of solutions to the challenges.

### 5.3 Limitations of the Study and Directions for Future Research

Understanding the limitations of this study is very important. The questionnaire respondents and challenge extraction convergence were one of the limitations of the research as there are different activists in the industry including individual developers which works for themselves or as freelancers, development teams which work together and have and probably will not register a company, small and midsized development companies and finally the digital publishers and local android markets. Expert panelists are also of challenge as the more interviewees the more efficient the results are. Future works could be done on specific sectors of the MAD market using new extensions of QFD including QFD and goal programming (GP), QFD and expert systems, fuzzy QFD, QFD under uncertainty, and dynamic QFD. Solutions projects and their implementation execution sequence could be a good topic for future research and considerations. Also, Technical benchmarking with countries like India, China, and Poland which are Top Countries for Offshore MAD could be done in future studies.

## Appendices
## Appendix (1)

**Table 1.** Challenges from Investigated Articles and The Outcome of The Interviewees and Their Opinions on Challenges are shown by stars (*)

| Categories | Challenges | Reviewed Articles | | | | | | | | | | | | | | | Interviewees | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Motevasseli | Bitab | Chit Sazan | Amiri | Karimi | Arasti | Arasti | Bargar | Rob | Schulten | Boot | Mayer | Davila | Dimo | Alexi | 1st Expert | 2nd Expert | 3rd Expert | 4th Expert |
| Financing | Lack of government financial support (development of effective financing programs in terms of facilities, grants & investments). | | * | | * | * | | | | | | | | | | | | * | | |
| | Lack of government financial bonuses (lack of effective financial incentives, including taxes, exported bonuses, etc.). | | * | | * | * | | | | | | | | | | | | * | | |
| | The lack of collateral (not having the required collateral to use the current insurances, due to their individuality or completion by micro-enterprises). | * | | | | * | | * | * | | | | * | | | | | * | | * |
| | Deficiency in capital raising skills. | | | | | * | | | * | | | * | | | | | | | * | * |
| | The lack of venture capitalists and mechanisms for sufficient support. | | | | | | | | | * | * | | | | * | | * | | | |
| | Significantly high interest rates on bank facilities. | | | | | | * | | | * | * | | | | | | | | | |
| | Maximum return on capital in-country and the current parallel markets such as gold, stock exchange, etc., which can potentially create higher returns as well as fluctuations and more. | | | | | | | | | | | | | | | | * | | * | |
| | High return expectancy in a short period and, in addition to enhanced liquidity in such a way that if another suitable investment opportunity was provided, the investor's capital will be directed towards that path. | | | | | | | | | | | | | | | | | * | | |
| | Using classical & traditional methods of investment instead of taking risks and investing in new and undetermined fields and areas. | | | | | | | | | | | | | | | | | * | | * |
| Market | Sales & Marketing issues. | | | | | | * | | | | | | | | | | | | | |
| | The lack of a competitive market. | * | * | | * | * | * | | | | | | | | | | | | | |
| | Inefficiency in the current size of the domestic market and the Persian-speaking markets (which is estimated at approximately less than 200 million users in the entire world). | * | * | | * | | * | | | | | | | | | | * | | | |
| | Inappropriate target market selection (improper targeting, improper selection, etc.) | | | | | | | | | | | | | | | | | * | * | |
| | The lack of sufficient knowledge for targeting the right markets | | | | | | | | | | | | | | | | | * | * | * |
| | The lack of authentic and powerful digital publishers with sufficient knowledge on domestics and international needs as well as the premium cost of access to reputable foreign publishers for apps distribution. | | | | | | | | | | | | | | | | * | | | |

| Categories | Challenges | Reviewed Articles | | | | | | | | | | | | | | | Interviewees | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Motevasseli | Bitab | Chit Sazan | Amiri | Karimi | Arasti | Arasti | Bargar | Rob | Schulten | Boot | Mayer | Davila | Dimo | Alexi | 1st Expert | 2nd Expert | 3rd Expert | 4th Expert |
| Rules & Regulations | The lack of transparency and proper rules & regulations (for digital stores, applications, etc.) | | * | | | | | * | | | | | | | | | | * | * | |
| | The lack of a proper grading system, in addition to the absence of an official registration system for new ideas and copyright laws. | | * | | | | | | | | | | | | | | | * | | |
| | The absence of recognition of many app developers including private intellectual individuals in support programs provided by government agencies, such as the Ministry of Information, Technology & Communications, the Ministry of Islamic Culture & Guidance, the innovation and prosperity fund, the National Development Fund, etc. | | | | | | | | | | | | | | | | * | | | |
| | The lack of proper laws to protect the interest of the developers and the unjust entry of application distribution network companies at times, which potentially include markets for mobile app stores (such as Bazaar Café, Miket, etc.), in addition to mobile operators, virtual smartphone operators and value-added service enterprises and a considerable reduction in profit margins and motivational level of developers to their continuous operation in this field. | | | | | | | | | | | | | | | | * | | | |
| | The income tax collection, value-added, and many more are the potential variables that can directly impact the exporters of domestic applications by the country of exported destination as well as the country of origin of production (Iran). | | | | | | | | | | | | | | | | * | | | |
| | The lack of proper supervision on the import of mobile apps and non-collection of income and value-added tax from the importers of such applications. | | | | | | | | | | | | | | | | * | | | |
| Environmental | Inadequate macroeconomic environment (macroeconomic indicators include inflation, exchange rate, interests on bank deposits and, etc.) | | | | | | * | | | | | | | | | | | | | |
| | Sanction-related issues that can potentially affect the obtainment of currency and branding in the space of stores from foreign smartphone applications and resources, including Google Play, App Store, in addition to also providing open-source apps to the publisher. | | | | | | | | | | | | | | | | * | * | * | * |
| | The lack of payment culture and mistrust for using apps, particularly content-based applications that are commonly used by the members of communities. | | | | | | | | | | | | | | | | | | | * |

| Categories | Challenges | Reviewed Articles | | | | | | | | | | | | | | | Interviewees | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Motevasseli | Bitab | Chit Sazan | Amiri | Karimi | Arasti | Arasti | Bargar | Rob | Schulten | Boot | Mayer | Davila | Dimo | Alexi | 1st Expert | 2nd Expert | 3rd Expert | 4th Expert |
| Technical & Production | Non-skilled labor and lack of sufficient workforce that is required. | | | | | | | | | | | | | | | | | | * | |
| | The lack of specialized events in the industry, including team building events, in addition to the transfer and exchange of experiences and more. | | | | | | | * | | | | | | | | | * | | * | * |
| | The lack of a proper system of knowledge management and sharing experiences and various achievements in this industry with the other activists from Iran and the rest of the world. | | | | | | * | | | | | | | | | | | * | * | * |
| | Non-economic nature of the implemented projects (due to improper project management, in addition to the lack of using new production methods, etc.) | * | | | | | | * | | | | * | | * | | | | * | * | |
| | Inappropriate evaluation of projects. | | | | | | | | | | | | | | | | | | | * |
| | The lack of proper pricing rules & regulations. | | | | | | | | | | | | | | | | | * | | * |
| | The absence of innovative individuals with ideas in the next stage and the conduct and implementation of the projects. | | | | | | | | | | | | | | | | | | * | |
| Internal | Internal deficiency of managerial skills. | | * | * | * | | * | | | | | | | | | * | | | | |
| | The lack of teamwork culture. | | | * | | | | | | | | | | | | | | | | |
| | The lack of experience in launching different types of apps or projects. | | | | | | * | | | | | | | | | * | | | | |
| | Excessive attention to the technological layer and neglect of the business part. | | | | | | | | | | | | | | | | | | * | |

## Appendix (2)

**Table 2.** Summary of Solutions investigated by reviewed articles and experts' opinions

| Categories | Solutions | Articles | | | | | | | | | | Experts | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Bitab | Bunj Shafiyi | Amiri | Rob | Burger | Walker | Mayer | Dimo | Lich | Alexi | 1st Expert | 2nd Expert | 3rd Expert | 4th Expert |
| Financing | A change of perspective in both capital and budget management. | | | | | | | | | | | | | * | |
| | Government economic policies and financing. | * | * | * | | | | | | | | * | * | | * |
| | Government financial incentive policies. | * | * | * | | | | | | | | * | * | | * |
| | Economic outscoring. | | | | | | | | | | | | * | | |
| | The entry of venture capitalists. | | * | * | | | | | * | * | * | | * | | * |
| | Internal financing (which often requires the use of the owner's capital). | | | | * | * | | | | | | | * | * | |
| | Acquiring sufficient capital raising skills. | | | | | | | | | | | | | * | |
| | Acquiring bank loans. | | | | | | * | * | | | | | | | |
| | The government support for venture capitalists. | | | | | | | | * | * | * | | * | | * |
| | Lowering interest rates on bank loans for such projects. | | | | | | * | * | | | | | * | | |

| Categories | Solutions | Articles | | | | | | | | | | Experts | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Bitab | Bunj Shafiyi | Amiri | Rob | Burger | Walker | Mayer | Dimo | Lich | Alexi | 1st Expert | 2nd Expert | 3rd Expert | 4th Expert |
| Market | Sufficient market knowledge before entering the market. | | | | | | | | | | | | * | * | * |
| | Constantly attempting to enter new and broader markets. | | | | | | | | | | | * | * | | |
| | The more optimal presence of publishers in the market and approaches of captivating international markets. | | | | | | | | | | | * | | | |
| Rulers & Regulations | The development of proper rules & regulations for the registration of innovative ideas. | | * | * | | | | | | | | * | * | | |
| | The development and adjustments of the current rules & regulations. | * | * | * | | | | | | | | * | | * | |
| | The complete recognition of app developers and content in the support programs of the Ministry of ICT & Postbank. | | | | | | | | | | | * | | | |
| | Approving laws of obtaining a guarantee from the publisher firm for the app developers, to not use the programmed source code. | | | | | | | | | | | * | | | |
| | Excessive taxation of content-based exporters at home, due to the previous taxes charged by the publishers and producers to which the application software and content were exported. | | | | | | | | | | | * | | | |
| | Sufficient supervision to value-added tax as well as income tax on foreign program publishers in Iran. | | | | | | | | | | | * | | | |
| | Elimination of unnecessary government interventions. | * | | | | | | | | | | | | | |
| Environmental | The reformation of environmental climate by the government and related institutions (sanctions & stability in the market) | | | * | | | | | | | | | | | |
| | Holding a variety of conferences to encourage such approaches and the introduction of these innovations. | | | | | | | | | | | | * | | |
| Technical & Production | Technical & production efforts to captivate and retain the specialized workforce. | | | | | | | | | | | | | * | |
| | Scientific pricing and researchers. | | | | | | | | | | | | | | * |
| | The conduct of feasibility before the launch. | | | | | | | | | | | | | | * |
| | The use of team development of content approaches. | | | | | | | | | | | * | | | |
| | The establishment of a knowledge management system in the firm. | | | | | | | | | | | | | * | |
| | Continues corporation with innovators until the last stages of production. | | | | | | | | | | | | | * | |
| Internal | Focusing on the corporate business layer and preparing a fundamental structure for teams. | | | | | | | | | | | | | * | |
| | Acquiring the required managerial and entrepreneurial skills before launch (training) | | | * | | | | | | | * | | * | * | * |
| | Creating a team spirit amongst the employees of the company. | | | | | | | | | | | | | * | |
| | Studying the past activities of entrepreneurship and gain experience from previously completed projects. | | | | | | | | | | * | | | | |

## References

Ahamd, A., Feng, C., Asim, S M., yousif, A. (2018). An Empirical Study of Investigating Mobile Applications Development Challenges, SPECIAL SECTION ON SOFTWARE STANDARDS AND THEIR IMPACT IN REDUCING SOFTWARE FAILURES ,6, 17711-17728.

Ahammeda, F., Azeem A., (2013). Selection of the most appropriate package of solar home system using analytic hierarchy process model in rural areas of Bangladesh. Renew Energy. 55: pp. 5-11.

Alsaadi, M. R., Ahmad, S. Z., & Hussain, M. (2018). A quality function deployment strategy for improving mobile-government service quality in the Gulf cooperation council countries. Benchmarking: An International Journal, 25(8), 3276-3295.

Apoorv P, Akanksha P , Vineet K , (2023). Mobile Application Market by Type, alliedmarketresearch, by segment report.

Arasti, Z., Zandi, F., & Talebi, K. (2012). Exploring the effect of individual factors on business *failure in Iranian new established small businesses. International Business Research, 5(4), 2.

Aydarov, D., Klochkov, Y., Ushanova, N., Frolova, E., & Ostapenko, M. (2018). Developing Plans for QFD Based Quality Enhancement. System Performance and Management Analytics, 241-249.

Beck, T., Demirgüç-Kunt, A., & Maksimovic, V. (2008). Financing patterns around the world: Are small firms different. Journal of Financial Economics, 89(3), 467-487.

Bamshad, V., talebi, k., yazdani, h., (2021). A Systematic Mapping Review and analysis of cognitive science in Opportunity Identification Literature: 1994-2018. Journal of Entrepreneurship Development, 14(1), 1-19.

Bunj Shafiyi. (2018). International confrance on Applied Research in Management & Industrial Engineering, Investing in small midsized ventrures, Tehran, IRAN.

Boot, A. W. (1999). European lessons on consolidation in banking. Journal of Banking & Finance, 23(2-4), 609-613.

Chavis, L. W., Klapper, L. F., & Love, I. (2011). The impact of the business environment on young firm financing. The world bank economic review, 25(3), 486-507.

Chitsazan, H., Talebi, K., & Mohebali, A. (2017). Identifying and Prioritizing the Most Common Methods of Bootstrap Financing in IT-based Start-ups. Journal of Entrepreneurship Development, 10(1), 61-79.

Davila, A., Foster, G., & Gupta, M. (2003). Venture capital financing and the growth of startup firms. Journal of business venturing, 18(6), 689-708.

Dewey, A. & Drahota, A. (2016) Introduction to systematic reviews: online learning module Cochrane Training https://training.cochrane.org/interactivelearning/module-1-introduction-conducting-systematic-reviews

Fartash, K., Baramaki, T., Saremi, M. S., & Sadabadi, A. A. (2022). The Growth Challenges of Pioneer Knowledge-based Firms of ICT. Journal of Science and Technology Policy, 15(3), 41-54(in persian).

John, R., Smith, A., Chotipanich, S. and Pitt, M. (2014), "Awareness and effectiveness of quality function deployment (QFD) in design and build projects in Nigeria", Journal of Facilities Management, 12 (1) 72-88.

Gutiérrez, J. A & Xu, G.(2006). An exploratory study of killer applications and critical success factors in m-commerce. Journal of Electronic Commerce in Organizations (JECO), 4(3), 63-79.

Karimi, A., & Bozarjomehri, S. (2014). Phenomenology of experiences and entrepreneurs' perceptions of entrepreneurship phenomenon; A qualitative study. Journal of Entrepreneurship Development, 7(3), 467-486.

Nyamaka , A., Botha, A., Biljon, J,. Marais, M. (2018). Challenges Botswana's Mobile Application Developers Encounter: Funding, Commercial and Technical Support, 2018 IST-Africa Week Conference (IST-Africa).

Pedarpour, A., Haghighi-Rad, F., Abtahi, A.R. (2022). A QFD-TRIZ Hybrid Method for a Hygiene Product. Int.J. Sysytematic Innovation, 7(1), 18-34.

Sadatrasoul, S. M., Asadifar, R., Shokri, H. R., (2022). Identify policy tools to develop technological collaboration of large industrial companies with knowledge-based companies, Innovation Management Journal, 80.

Statista Research department. (2021). Mobile app revenue worldwide 2017-2025, by segment report.

Szczepanski, Martin.(2018). European app economy, State of play, challenges and EU policy. From ERPS (European Parliamentary Research Service): http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2018)621894

Vinayak, K. and Kodali, R. (2013), Benchmarking the quality function deployment models, Benchmarking: An International Journal,20(6), 825-854.

Zheng, X. and Pulli, P. (2012), "Improving mobile services design: a QFD approach", Computing and Informatics, 26(4),369-381.

- **AUTHOR BIOGRAPHIES**

**Seyed Mahdi Sadatrasoul** is assistant professor at kharazmi university since 2014. He has more than 18 years of experience in Technology based SMEs business model innovations, venture capital and private equity investments. He has contributed more than 33 research papers to many international and national journals and conferences. He has also published two books by reputed publishers. His research interests include business model design and diagnose, financial investment models in Tech SMEs.

**Farzad Haghighi-Rad** is an Iranian researcher in the field of system thinking, system dynamics and management science. He received his doctorate in industrial engineering in 2010 from university of Tarbiat Modares. In the last 15 years, he has provided consulting and training services to more than 55 private companies and government agencies in the Iran. He is currently an assistant professor at the Faculty of Management of Kharazmi University.

**Seyyed Ali Entezar** has obtained his bachelor's degree in public sector management from Beheshti University and his master's degree in information technology management from Kharzami University. His interest is in innovative research of government and public sector policies in advancing information technology companies in developing countries.

DOI: 10.6977/IJoSI.202409_8(3).0005
D. D, A. O N/Int. J. Systematic Innovation, 8(3), 56-62 (2024)

# Predicting the impact of blockchain technology implementation in SMEs

Divya D[1*], Arunkumar O N[2],
[1]Assistant Professor, Symbiosis Centre for Management Studies (SCMS), Symbiosis International (Deemed University) (SIU), Bengaluru, Electronics City, Hosur Road, Bengaluru, Karnataka, India. Email: divya.d@scmsbengaluru.siu.edu.in
[2]Assistant Professor, Symbiosis Institute of Business Management (SIBM), Symbiosis International (Deemed University) (SIU), Bengaluru, Electronics City, Hosur Road, Bengaluru, Karnataka, India. Email: arunkumar@sibm.edu.in

(Received 30 August 2023; Final version received 18 November 2023; Accepted 05 July 2024.)

## Abstract

In the last few years, blockchain technology, or BCT, has gained much traction. Small and medium-sized businesses (SMEs) struggle more than their larger counterparts when it comes to technological adaptation because they lack the technology infrastructure required to implement blockchain technologies. The major contribution of this paper is to predict the impact of blockchain technology implementation on the performance of SMEs. A multiple-output regression model is utilized in this research to predict the impact of BCT on SMEs' performance. The cost of implementing and maintaining blockchain technology, IT project management, compatibility, benefit over other available technological options, and trialability are the independent variables that were considered in the analysis. Software revision, sophistication level, innovation complexity, and observability are the dependent variables. Researchers and industry professionals can use the study to comprehend how implementing blockchain technology affects SMEs.

*Keywords: Blockchain Technology, SMEs, multiple output regression*

## 1. Introduction

Small and medium-sized businesses are thought to form the foundation of the Indian economy. The world is witnessing a huge digital revolution as we enter Industry 4.0 and the age of digital transformation. As a result, organizations need to evolve their technological infrastructure to survive in the industry (Salim et al., 2022). The various technologies that can bring technological innovation to SMEs include AI, machine learning, blockchain technology, and IoT. Although SMEs are not organizationally ready to adapt to all these modern technologies, adopting this technology will enhance their productivity, performance, and security. This research work focuses on the impact of BCT on SME performance.

Blockchain is defined as "a distributed database, which is shared among and agreed upon as a peer-to-peer network. It consists of a linked sequence of blocks (a storage unit of the transaction), holding timestamped transactions that are secured by public-key cryptography (i.e., "hash") and verified by the network community. Once an element is appended to the Blockchain, it cannot be altered, turning a Blockchain into an immutable record of past activity" (Seebacher & Schuritz, 2017).

Previous research works in this area analyze the impact of BCT on risk management, considering the advantages of BCT, including transparency, security, and traceability (Chowdhury et al., 2022). However, identifying the impact of a technology is also very important to decide whether or not the organization has to go ahead with the new technology. This research work tries to identify the impact of BCT on SME performance. This paper works around the assumption that the implementation of BCT technology improves SME performance. This can be proved by predicting the efficiency of the SME after BCT implementation. The first step in this case is to identify a model that gives an accurate prediction of SME's performance before and after BCT implementation, which can help the users to make a decision based on the results. Hence, the two research objectives of the study are formulated as below

R1: Identifying the best model to predict SME performance after implementation of BCT technology.

R2: Predicting the impact of BCT technology on SME performance.

DOI: 10.6977/IJoSI.202409_8(3).0005
D. D, A. O N/Int. J. Systematic Innovation, 8(3), 56-62 (2024)

While finding solutions to these two research objectives, this paper makes a major contribution to both technologists and SME leaders. None of the existing research works predicts the impact of multiple input variables on multiple output variables. However, current research work aids the researchers in utilizing a multi-output regression model when there are multiple input and multiple output variables. Another contribution of this research work is for SME leaders. SME leaders are clueless about the benefits of implementing BCT in their company. Through a simulation study, this research work elucidates that SME improves efficiency in various factors through BCT implementation. This identification will help the SME leaders to make a decision on whether or not to go ahead with the implementation of BCT.

The remainder of the document is structured as follows: Section 2 details the background of the study. Section 3 includes the methodology used for the analysis as well as the results obtained from the analysis. Section 4 presents limitations of the study and Section 5 concludes the main findings obtained from this study.

## 2. Background

Even though blockchain technology (BCT) offers numerous benefits, like anonymity, immutability, transparency, and quick transactions (Abubakar & Alyou, 2021), it's critical to comprehend how the adoption of BCT will affect different businesses. A number of frameworks, including the Technology Acceptance Model (TAM) and the Technology-Organization-Environment (TOE) framework, were used to examine the effects of BCT adoption in SMEs (Kamble et al., 2022). The biggest driving variables for blockchain adoption are partner preparedness, perceived usefulness, perceived ease of use, and competitive pressure, which are important variables that must be considered. It is very important to analyze the individual's intent to adopt the blockchain rather than an organizational perspective (Kamble et al., 2022). Challenges of BCT adoption in the supply chain have been examined by Queiroz and Wamba (2019). The impact of Blockchain Technology (BCT) on variables such as supply chain stakeholders' trust, social influence, performance expectancy, effort expectancy, and facilitating conditions was examined. They employed partial least squares structural equation modeling (PLS-SEM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) to detect the impact of BCT. This study shows that adoption behavior varies in different countries Queiroz and Wamba (2019).

To test acceptance in terms of behavioral intention to use, two forms of information processing competency (private and public blockchain-based cooperation) are used (Giri & Manohar, 2021). Barriers to adopting blockchain technology in SMEs have been analyzed by Kaur et al. (2022). This study emphasizes the role of Government officials to concentrate on building infrastructure throughout the nation's various regions to support hosting blockchain systems. Such a homegrown blockchain platform can speed up and lower the cost of blockchain adoption for SMEs (Kaur et al., 2022). Another important factor that has to be considered for BCT adoption by SMEs is building the confidence of SMEs and making them aware of the positive impacts of BCT technology in their organization (Kaur et al., 2022).

Studies on the impact of BCT to improve the quality of accounting information, to help investors better understand the company was conducted by researchers such as Wu and Jin (2022). Several studies on the impact of BCT technology on various sectors, such as the maritime supply chain (Nasih et al., 2019), business models (Morkunas et al., 2019), and business interactions (de Oliveira et al., 2021).

Various methodologies such as additive symbiotic networks by Ferreira et al. (2023) are used to understand the impact of BCT technology. Content analysis was used by de Oliveira et al. (2021) and case study analysis was utilized by Stranieri et al. (2021). None of the methodologies in the literature make use of machine learning methods to identify the impact of BCT on various industries.

All these existing works in the literature do not utilize machine learning approaches to predict the impact of BCT in SMEs. This research utilizes a machine learning-based approach to identify the impact of BCT implementation on SMEs. The next section identifies the important variables to determine the impact of BCT implementation on SMEs and applies a machine learning model to predict the impact of BCT on SMEs' performance.

## 3. Methodology

In order to understand the impact of BCT on SME performance various factors that affect SME performance have to be analyzed first. These variables are identified by using various grounded theories such as the sense-making theory (Savolainen, 1993),

DOI: 10.6977/IJoSI.202409_8(3).0005
D. D, A. O N/Int. J. Systematic Innovation, 8(3), 56-62 (2024)

Organizational information processing theory, and Resource Based View (RBV) (Savolainen, 1993). Organizational information processing theory (Koh, 2016) believes that "the greater the uncertainty of the task, the greater the amount of information that must be processed between decision makers during the execution of the task to get a given level of performance". Diffusion of Innovation (Yu et al., 2021) helps to pinpoint the aspects of innovation that affect acceptance.

Based on the above-mentioned theories, this research identified the important variables for this study, which are listed in Table 1.

**Table 1.** Variables influence the BCT adoption in SMEs

| Sl. No. | Variables | Notation |
|---|---|---|
| 1 | Cost of BCT implementation | V1 |
| | Software Revision of BCT | V2 |
| | Maintenance cost of BCT | V3 |
| 2 | Level of sophistication of IT usage | V4 |
| | IT Project Management | V5 |
| 3 | Innovation Complexity | V6 |
| | Organization Compatibility | V7 |
| | Benefits among other technologies | V8 |
| | Observability | V9 |
| | Trialability | V10 |

Using these variables, a machine learning model has been developed in this research work to forecast the impact of BCT on SME's performance.

## 3.1 Multi-output regression model

Prediction of two or more numerical variables is the goal of multi-output regression. This model was developed to address the gap where the multivariate regression approach is preferred over separate univariate predictions (Rogers, 2010; Schmid, et al., 2023). The method used in this research includes a simulation study with different weights for the blockchain technology variables.

The model used for this model is given in **Figure 1.**

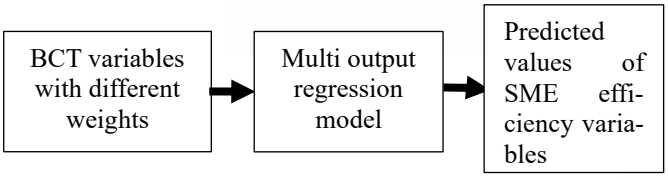



**Fig 1.** Structure of multi-output regression model

Three different types of algorithms are used to develop a multi-output regression model.

In this research, a multi-output regression model is developed, which takes different BCT variables as input and identifies the impact of those variables on SME's efficiency. The parameters to represent the SME's efficiency are selected as the output variables. Based on the literature we found that the Implementation cost of blockchain technology, the Maintenance cost of blockchain technology, IT Project Management, Compatibility with the organization, Benefit compared to other existing technological choices, and trialability are the input variables for the multi-output regression model. From the literature, we have identified the amount spent on Software Revision of BCT, the Level of sophistication of IT usage by the employees and customers, the Complexity of the innovation, and the Observability that determines the efficiency of SMEs after BCT implementation. Current research work uses a hypothetical situation of a company that implements BCT. In order to represent the hypothetical company, an input dataset is generated using simulation. A thousand samples are generated using simulation, which is utilized for the current analysis. In order to identify the best model that predicts the impact of BCT implementation on SME's efficiency, three different types of machine learning models are utilized. The first model presents a linear regression model with multiple outputs, the second one presents a KNN regressor model, and the last one presents a decision tree model. All these three models predict the impact of SME's efficiency after BCT implementation. A comparative analysis of the three models determines the one that has higher accuracy during prediction.

### 3.1.1 Simulation

**Case A: Multi-output Linear Regression Model**

Step 1: Create datasets

Step 2: X, y = make_regression (n_samples=1000, n_features=6, n_informative=5, n_targets=4, random_state=1, noise=0.5)

Step 3: Define model Linear Regression ()

Step 4: Fit the regression model using model.fit(X, y) function

Step 5: Make a prediction when all the selected variable inputs are zero

Step 6: Summarize prediction using regression

Step 7: Obtain the output for the selected output variables

Step 8: Evaluation of the model using cross-validation

Step 9: Generation of Mean absolute error to get the accuracy of the model

### Case B: Multi-output Knn Regression Model

Step 1: Create datasets

Step 2: X, y = make_regression (n_samples=1000, n_features=6, n_informative=5, n_targets=4, random_state=1, noise=0.5)

Step 3: Define model KnnRegression()

Step 4: Fit the regression model using model.fit(X, y) function

Step 5: Make a prediction when all the selected variable inputs are zero

Step 6: Summarize prediction using regression

Step 7: Obtain the output for the selected output variables

Step 8: Evaluation of the model using cross-validation

Step 9: Generation of Mean absolute error to get the accuracy of the model

### Case C: Multi-output Decision Tree Regression Model

Step 1: Create datasets

Step 2: X, y = make_regression(n_samples=1000, n_features=6, n_informative=5, n_targets=4, random_state=1, noise=0.5)

Step 3: Define model Decision_TreeRegression ()

Step 4: Fit the regression model using the model.fit(X, y) function

Step 5: Make a prediction when all the selected variable inputs are zero

Step 6: Summarize prediction using regression

Step 7: Obtain the output for the selected output variables

Step 8: Evaluation of the model using cross-validation

Step 9: Generation of Mean absolute error to get the accuracy of the model

After implementing three different types of multi-output regression models, cross-validation has been done to identify the best model for predicting SME performance. SME performance is measured by predicting the values of variables such as Software Revision of blockchain technology which represents how much is assistance obtained from BCT variables for software revision, how much the improvement in the Level of sophistication of IT usage, how better the complexity of the innovation factor improves with implementation of BCT and how much improvement BCT creates in the observability of SME.

## 3.2 Results and Discussion

Initially, all the input BCT variables are set to zero to understand how much the performance of SMEs without the implementation of BCT is given as Case A: Scenario 1.

### Case A: Multi-output linear regression output

***Scenario 1***: SME performance before implementation of BCT variables:
[ 0.00679014 0.00967136 -0.00181127 -0.00393957]

Later the same algorithm can be implemented by setting up the weights of BCT variables as one. This represents a scenario in which the BCT algorithm is implemented in the SME, and we can identify the SME output variable performance that can be calculated.

### Case A: Multi-output linear regression output

***Scenario 2:*** SME performance after implementation of BCT variables:
[188.8041768 269.25581906 196.06294426 340.31991746]

MAE: 0.395 (0.010)

Similarly, SME's performance is predicted using Multi-output Knn regression in Case B. Scenario 1 presents SME's efficiency before BCT implementation, and scenario 2 presents SME's efficiency after BCT implementation.

### Case B: Multi-output Knn regression output

***Scenario 1:*** SME performance before implementation of BCT variables:
[-21.65768054 -14.59638324 -5.78102635 -10.62669042]

***Scenario 2:*** SME performance after implementation of BCT variables:

[167.81018346 199.85623374 136.88660717 262.07654543]

MAE: 28.537 (2.179)

Same scenarios are replicated using Multi-output Decision Tree regression which is given in Case C.

### Case C: Multi-output Decision Tree regression output

***Scenario 1:*** SME performance before implementation of BCT variables:
[ 7.63187532 -4.92012703 -9.25101485 -9.2697882]

***Scenario 2:*** SME performance after implementation of BCT variables:

[75.28231803 128.24568503 107.82181375 206.19706852]

MAE: 53.369 (3.364)

With the aim of selecting the best methodology for prediction, a comparative study of outputs of multi-regression output models has been performed through a simulation study. Results obtained from the study revealed that multi-output linear regression has higher performance compared to the other models. This can be identified by checking the three models' Mean Absolute Error (MAE). From the outputs, it is clear that the multi-output regression model implemented using linear regression has a higher performance compared to Knn regressor or the decision tree model.

In scenario 1, where BCT is not implemented in the SME, output variables, which show the performance of SME, are very low compared to scenario 2, in which the output variables are measured after implementing the BCT. Predicting the output variables helps the SME to understand their performance improvement while implementing BCT. Thus, the hypothesis that implementation of BCT improves the performance of the SME is well established through this simulation study.

Modern technologies like machine learning, AI, IoT, and BCT improve the performance of various industries. Various studies exist about the barriers and enablers of all these technologies. Many researchers use models to understand how BCT affects different industries, including the Technology Acceptance Model (TAM), Technology-Organization-Environment (TOE), and Unified Theory of Acceptance and Use of Technology (UTAUT). However, a forecasting model is very important in predicting the impact of BCT implementation on SMEs. Therefore, studies related to predicting the performance of SMEs after and before BCT technology is very important. This research work tries to predict the impact of BCT in SMEs. Three different models have been utilized to predict SME's efficiency. A detailed explanation regarding the implementation of these three models given in section 3.2 helps future researchers to select the best model to predict the output variables in their study. SME leaders can benefit from this study by identifying the improvement in each output variable after BCT implementation.

## 4. Limitations of the study

A major limitation of this research work is the fact that this research doesn't collect real-life datasets from an SME. Instead, a simulation study based on the hypothetical situation is implemented in this research paper. Hence, in the future, the same set of variables can be collected from an SME through an empirical study, which brings more clarity to the procedure. Pre and Post analysis (before and after implementation of BCT) can be conducted in a real-life situation.

## 5. Conclusion

Due to the lack of technological infrastructure, many SMEs struggle to implement BCT technology. Many of the factors, like perceived ease of use and usefulness, came into the picture while analyzing the probability of implementing BCT. The implementation cost of BCT, the maintenance cost of BCT, IT project management, compatibility, the benefit compared among other technologies, and trialability with the organization are considered as the input variables for this study. Their impact on the dependent variables' software revision, level of sophistication, complexity of innovation, and observability are analyzed using a multi-output regression model. Results show that BCT implementation increases the efficiency of SMEs with respect to all these dependent variables. Therefore, implementing BCT in SMEs is recommended for better efficiency.

## References

Abubakar, A.M. & Al-zyoud, M.F. (2021). Problematic internet usage and safety behavior: does time autonomy matter, Telematics and Informatics, 56, 1-11.

Chowdhury, S., Rodriguez-Espindola, O., Dey, P. & Budhwar, P. (2022). Blockchain technology adoption for managing risks in operations and supply chain management: evidence from the UK. Annals of Operations Research. https://doi.org/10.1007/s10479-021-04487-1

de Oliveira F.C., Zanoni R., Dalla-Rosa R. & Verschoore J.R. (2021). Blockchain technology and relational gains in business interactions: a fuzzy set qualitative comparative analysis of supply chain specialists' perceptions, International Journal of Advanced Operations Management, 13 (4), 372 – 390.

Ferreira, A. I., Palazzo, G. & Carvalho, H. (2023). Implications of the blockchain technology adoption by additive symbiotic networks, Cleaner Logistics and Supply Chain, 6, 100095 https://doi.org/10.1016/j.clscn.2023.100095

Giri, G. & Manohar, H.L. (2021). Factors influencing the acceptance of private and public blockchain-based collaboration among supply chain practitioners: a parallel mediation model, Supply Chain Management: An International Journal. 28 (1), 1-24.

Kamble, S., Gunasekaran, A., Kumar, V., Belhadi, A. & Foropon, C. (2022). A machine learning based approach for predicting blockchain adoption in supply Chain, Technological Forecasting and Social Change, 163. 10.1016/j.techfore.2020.120465.

Kaur, J., Kumar, S., & Narkhede, B.E. (2022). Barriers to blockchain adoption for supply chain finance: the case of Indian SMEs, Electronic Commerce Research. https://doi.org/10.1007/s10660-022-09566-4

Koh, S. C. L., Morris, J., Ebrahimi, S. M., & Obayi, R. (2016). Integrated resource efficiency: Measurement and management, International Journal of Operations & Production Management, 36 (11), 1576–1600.

Morkunas V.J., Paschen J., & Boon E. (2019). How blockchain technologies impact your business model, Business Horizons, 62 (3), 295 – 306. 10.1016/j.bushor.2019.01.009

Nasih S., Arezki S. & Gadi T. (2019). Blockchain technology impact on the maritime supply chain. ACM International Conference Proceeding Series. DOI: 10.1145/3368756.3369104

Queiroz, M. M. & Wamba, S.F. (2019). Blockchain adoption challenges in supply chain: An empirical investigation of the main drivers in India and the USA, International Journal of Information Management, 46, 70-82. https://doi.org/10.1016/j.ijinfomgt.2018.11.021.

Rogers, E.M. (2010). Diffusion of Innovations. Simon and Schuster.

Salim, T. A., Barachi, M. E., Alfatih D. A. Mohamed, Halstead, S. & Babreak, N. (2022). The mediator and moderator roles of perceived cost on the relationship between organizational readiness and the intention to adopt blockchain technology, Technology in Society, 71. https://doi.org/10.1016/j.techsoc.2022.102108

Savolainen, R. (1993). The sense-making theory: Reviewing the interests of a user-centered approach to information seeking and use, Information

Processing & Management, 29 (1), 13-28, https://doi.org/10.1016/0306-4573(93)90020-E

Schmid, L., Gerharz, A., Groll, A. & Pauly, M. (2023). Tree-based ensembles for multi-output regression: Comparing multivariate approaches with separate univariate ones, Computational Statistics & Data Analysis, 179. https://doi.org/10.1016/j.csda.2022.107628

Seebacher, S. & Schuritz, R. (2017). Blockchain technology as an enabler of service systems: a structured literature review. Lecture Notes in Business Information Processing, Springer, Cham, 279.

Stranieri S.,Riccardi F., Meuwissen M.P.M. & Soregaroli C. (2021). Exploring the impact of blockchain on the performance of agri-food supply chains, Food Control, 119. DOI: 10.1016/j.foodcont.2020.107495

Wu C. & Jin S. (2022). Does Blockchain Technology Promote the Quality of Enterprise Accounting Information? Journal of Global Business and Trade, 18 (5). DOI: 10.20294/jgbt.2022.18.5.1

Yu, W., Zhao, G., Liu, Q. & Song, Y. (2021). Role of big data analytics capability in developing integrated hospital supply chains and operational flexibility: An organizational information processing theory perspective, Technological Forecasting and Social Change, 163. https://doi.org/10.1016/j.techfore.2020.120417

DOI: 10.6977/IJoSI.202409_8(3).0006
N. Qassrawi, M. Azzeh /Int. J. Systematic Innovation, 8(3), 63-83 (2024)

# Drug sales forecasting in the pharmaceutical market using deep neural network algorithms

Noura Qassrawi[1], Mohammad Azzeh[2*], Mohammad Hijjawi[1]
[1] Department of Computer Science, Applied Science Private University, Amman, Jordan
[2] Department of Data Science, Princess Sumaya University for Technology, Amman, Jordan
* Corresponding author E-mail: m.azzeh@psut.edu.jo
(Received 21 September 2023; Final version received 08 January 2024; Accepted 29 February 2024)

## Abstract

Drug sales and price forecasting have become an attractive investigation topic due to their important role in the pharmaceutical industry, A sales forecast helps every business to make better business decisions in overall business planning, budgeting, marketing, and risk management. The traditional forecasting method focuses on a conventional statistical model, which highly depends on the availability of historical sales data. However, for new drug entities, where not enough historical data is available, new methods of Machine Learning are applied. The aim of this paper is to identify an efficient Deep Neural Network algorithm suitable to forecast drug sales and pricing by applying Deep Neural Network Algorithms such as Multilayer Perceptron, Convolutional Neural Network, and Long Short-Term Memory, which are expected to perform well on this issue. The results are carried out to determine the efficiency of these algorithms by evaluating the performances of the models using MAE and RMSE performance metrics to identify the best algorithm for Drug Sales and Price Forecasting. The accepted accuracy should be more than 80\% of the actual value for quantity which is less than three thousand by unit and less than two dollars (USD) for price, Based on the results of the experiments Long Short Term Memory performed better than MLP and CNN for generating predictions with average Root Mean Square Error of for sales is 1.28(k) and Mean Absolute Error of about 0.85(k), and with average Root Mean Square Error for USD Prices is about 0.75, and Mean Absolute Error is about 0.44. The forecasts are then used to adjust stock levels according to the predictions.

*Keywords: Drug sales Forecasting, Multilayer Perceptron, Convolutional Neural Network, Long Short-Term Memory.*

## 1. Introduction

The pharmaceutical industry involves the discovery, research, development, manufacturing, distribution and marketing of pharmaceutical products, which are drugs or medications used for the treatment or prevention of diseases (Taylor, 2015). It is one of the most resilient, fastest growing, and research-intense industries (Lakner et al., 2019). The industry is dominated by Multinational companies mainly in the United States (US) and European Union (EU). These companies invest highly in the research and development of innovative drugs which are protected by IP law once approved by the relevant regulatory bodies (such as the USFDA). This patent protection (for about 20 years) enables these companies to set high prices for their products to recover the large costs incurred in the discovery phase. Upon the expiry of an innovative drug's patent, a reformulation of the same active ingredient can be developed by other companies. This is called a generic drug. The process of developing a generic drug is much less expensive and less time consuming. For these reasons, in addition to the presence of a number of competitors producing the same active ingredient, the prices of generic drugs are significantly lower than those of innovative (originator's) drugs.

Most local and regional pharmaceutical manufacturing companies based in the MENA region are generic companies that produce copies of innovative drugs after the expiry of their patents. Drug pricing policies are

country-dependent1. For example, in the Kingdom of Saudi Arabia and Jordan, drug pricing is regulated by the Saudi FDA and the Jordanian FDA, respectively. The price depends on many factors, including whether the product is originator or generic, the patent is valid or expired, and the number of generic products in the market. There are also continuous governmental initiatives to control drug prices. On the other hand, the Pharmaceutical market is a highly fragmented market that includes international, regional, and local companies. There are no regulations to control drug prices. For the same product, the price may differ between pharmacies, warehouses, or regions. The presence of illegal and parallel imports only complicates the issue.

Pharma manufacturers strive to enhance their efficiencies through accurate planning and sales forecasting. Sales forecasting is vital across various functions in the industry; it plays a major role in formulating effective business plans and gaining valuable and powerful insights that empower companies to make important decisions related to their costs, performance, and profitability (Tiriveedhi, 2018). Time is a very important variable in the analysis of sales data using the Time-series method of forecasting. Data is collected at equally spaced intervals of time. This data from current and past periods is used to predict the values in future periods as well as show trends and repeating patterns (seasonality). Sales forecasting is a complex process as it involves different internal (company-related) and external factors (political, economic…etc.). In this paper, we aim to identify the critical factors that influence drug products' sales, and then Artificial Neural Network algorithms are applied to forecast the sales of drugs, these algorithms will be studied and compared against each other to identify the most efficient algorithm suitable to forecast sales of drugs based on the results obtained for a given dataset to provide the most accurate predictions.

This paper aims to identify the critical features that affect drug sales and also to identify the machine learning algorithms suitable for resolving the drug sales forecasting problem and find the most efficient algorithm among the chosen algorithms based on the models' performances for the given dataset. Particularly, we studied three of Deep Artificial Neural Network algorithms for sales and pricing forecasting: Long Short-Term Memory (LSTM), Multi-Layer-Perceptron (MLP), and Convolutional Neural Network (CNN). The current research is driven by the following research questions:

- **RQ1**: What are the critical features that influence drug sales?

**Motivation**: This research question's motivation is to find the critical features that will influence the sales of drugs in the private pharmaceutical market. This will help us to improve the quality of the results.

- **RQ2**: How can the Deep Artificial Neural Network algorithms be chosen to resolve the sales forecasting problem?

**Motivation**: The motivation for this research question is to examine suitable Deep Artificial Neural Network models for sales forecasting.

- **RQ3**: Which Deep Neural Network model is efficient for forecasting the sales of drugs in the private pharmaceutical market?

**Motivation**: The motivation of this research question is to identify efficient Deep Neural Network algorithms among the selected algorithms for forecasting the sales of the drugs in the pharmaceutical market based on the obtained results.

The critical features identified from RQ1 are used to develop the Machine Learning model using different Deep Neural Network algorithms that were selected in RQ2. An efficient machine learning model is then identified by comparing the performances of these models using various metrics such as Mean Absolute Error and Root Mean Square Error Metrics for the dataset to address RQ3.

## 2. Background

Sales prediction is an important part of modern business intelligence (C. H. Wang & Yun, 2020). It can be a difficult problem, especially in the case of lack of data, missing data, and outliers' behavior. Sales can be analyzed as a time series. At present, different time series models have been developed, for example, by Holt-Winters, ARIMA, GARCH, SARIMAX, SARIMA.. ect. (Andrawis et al., 2011; Doganis et al., 2006; Lim & Zohren, 2021). Time-series approaches have some

[1] https://www.sfda.gov.sa/en

limitations in sales forecasting, below are some of them: 1) Historical data for a long time period is required to capture seasonality. However, we often do not have historical data for a target variable, for example, when a new product is launched. Simultaneously, we have a sales time series for similar products, so we could expect that our newly launched product will have the same sales pattern (Ensafi et al., 2022). 2) Drug sales data could have missing data and outliers, which should be cleaned before using a time series approach (Ensafi et al., 2022) , and 3) We should study all critical factors that affect drug sales to predict the more accrue value.

The drug sales forecast in the pharmaceutical market is a regression problem. Practice shows that by Machine-learning models, we can find patterns in the time series also complicated patterns in the dynamics of sales using supervised machine-learning methods. There are several methods to forecast the future sales for the products in different business areas. Forecasts are used for planning production and other business activities such as purchasing materials, inventory management and often more across most industries. Traditional forecasting approaches were primarily focused on experienced employee opinions or statistical analysis of previous data such as time series and linear regression, but in recent years Machine Learning and Artificial Neural Networks techniques have been implemented with great success in this field (Boyapati & Mummidi, 2020) . The next subsections describe the main ideas of the above forecasting methods.

## 2.1 Time-Series Forecasting

Time series is a time-dependent sequence of observations of a variable (Benidis et al., 2022). Based on the rate at which the data is collected, time series is categorized into two types: Continuous time series and Discrete-time series. A continuous-time series is a sequence of observations made continuously through time. A time series is a discrete time series when the observations are collected at fixed or equal intervals of time such as Daily closing price of Google stock, Monthly sales of cars, Yearly rate of change of global temperature and so on (Yeasmin et al., 2022). A continuous time series can be sampled at equal intervals of time to form a discrete time series (Lim & Zohren, 2021). Time series data can be analyzed for several purposes such as to describe the (seasonal or trend) variations of time series data, to use variations of one time series to gain insights into another time series (Lim & Zohren, 2021). Time series forecasting is useful to develop a model by analyzing the past observations of a time series to describe the relationships in the time series. This model is then used to predict future values of the time series (Ji et al., 2016).

## 2.2 Artificial Neural Networks Algorithms

Artificial Neural Networks (ANN) comprises of multiple nodes that initiate biological neurons of the human brain. The neuron is connected by links, interact with each other using these links. The node takes the input data through the input layer and operates on the hidden layer's data. These operations result is passed to the other neurons. After computation, the result is passed to the output layer (Hossen et al., 2017). when the hidden layers are more than two in any neural network, it is known as a deep neural network, which uses a cascade of multiple layers of non-linear processing units for feature extraction. The output of the current layer is fetched to the next layer as an input.

CNN is a feed-forward neural network. Although based on the traditional architecture of a neural network, CNN includes input layers, hidden layers, and output layers. CNN was designed to automatically and adaptively learn spatial hierarchies of features, from low- to high-level patterns. CNN is a mathematical construct typically composed of three types of layers (or building blocks): convolution, pooling, and fully connected layers. The first two, convolution and pooling layers, perform feature extraction, whereas the third, a fully connected layer, maps the extracted features into final output, such as classification. A convolution layer plays a key role in CNN, which is composed of a stack of mathematical operations, such as convolution, a specialized type of linear operation (Kim et al., 2016). CNN's convolutions are popularly known to work on spatial or 2D data. What is less popular is that there are also convolutions for 1D data. This allows CNN to be used in more general data type including texts and other time series data. Instead of extracting spatial information, you use 1D convolutions to extract information along the time dimension as shown in Figure 1.

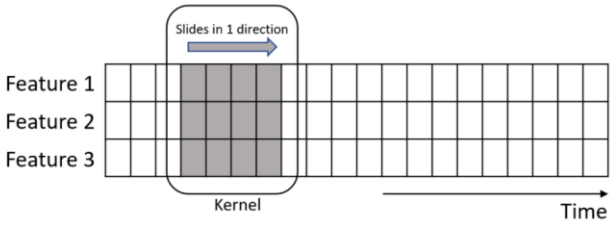



**Fig** 1. Conv1D Convolving on time dimension

### 2.3.2 Multi-layer Perceptron (MLP):

Multilayer Perceptron is a type of feed-forward neural network where the information flows from the input layer towards the output layer through the hidden layer as shown in Figure 2. Multilayer Perceptron, or MLPs for short, can be used to model univariate, multivariate, and multi-step time series forecasting problems (Dai et al., 2008). Rectified Linear Unit is used as activation function for Multilayer Perceptron algorithm. MLP makes use of a supervised learning algorithm called backpropagation for training the network (Lim & Zohren, 2021). In backpropagation, the error is propagated backward throughout the network. The error is calculated by taking the difference between the network output and the actual output. The network parameters called weights are modified to minimize this error based on this method. This process is repeated several times until a stopping condition is reached.

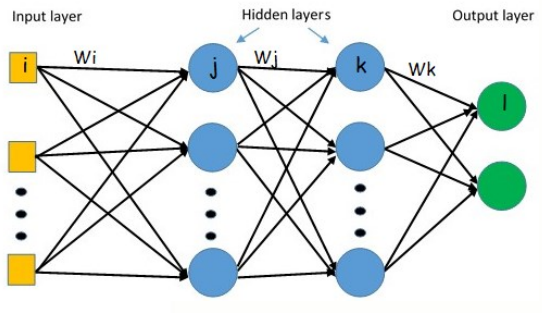



**Fig** 2. Multilayer Perceptron

$$y = f(x_1 w_1 + x_2 w_2 + \cdots + x_n w_n) \quad (1)$$

where '$x$' and '$y$' represents input and output of the network. '$W_i$' represents the connection weights between two input and hidden layers, $W_j$ represents the connection weights between 1st and 2nd hidden layer, $W_k$ denotes connection weights between hidden layer and output layer, '$f$' represents the activation function, '$i$','$j$','$k$','$l$' represents the nodes in the network (Yeasmin et al., 2022).

Long Short-Term Memory networks, or LSTMs for short, can be applied to time series forecasting. There are many types of LSTM models that can be used for each specific type of time series forecasting problem (Boyapati & Mummidi, 2020). Long Short-Term Memory (LSTM) is a type of recurrent neural network. It consists of regulators or LSTM units called gates. LSTM can learn long-term dependencies because of the usage of the gating mechanism and a memory cell. LSTM can overcome the vanishing gradient and exploding gradient problems faced by RNN (Dai et al., 2008; Yeasmin et al., 2022). A typical LSTM consists of a memory cell, input gate, output gate, and forget gate. The memory cell can remember information over arbitrary time periods. The gating mechanism regulates the flow of information to and from the cell. A typical LSTM consists of a memory cell, input gate, output gate, and forget gate. The memory cell can remember information over arbitrary time periods, and the gating mechanism regulates the flow of information to and from the cell (Z. Wang & Lou, 2019).

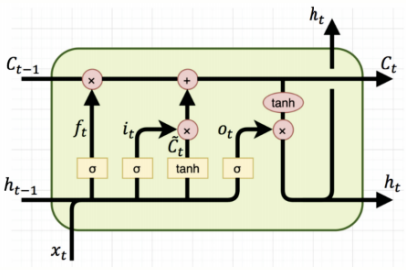



**Fig** 3. Long Short-Term Memory

In Figure 3, $c_t$ is the cell state which is used to carry information throughout the sequence chain which acts as memory. Forget gates represented by ft are used to determine which information should be eliminated from the cell state. For this purpose, they subject the input vector to a sigmoid function and then perform a pointwise multiplication operation with cell state $C_{t-1}$. The input gate it is used to determine the values (or input vector ($[h_{t-1},x_t]$) we are going to update. tanh function is used to generate new values from the input vector. The result from the input gate and tanh are combined by making use of a pointwise multiplication operation which is added to the cell state by making use of a pointwise addition operation. Finally, the output gate represented by $o_t$ determines which values of the input vector, we need output by applying a sigmoid function to the input vector. The value of hidden state $h_t$ is calculated by subjecting the cell state to a *tanh* function and then multiplying it with the output gate using pointwise multiplication. This information is then passed along the chain in a sequence and the above process is repeated (Z. Wang & Lou, 2019).

## 3. Related Work

Previously, several comparative studies between traditional models and neural networks have been carried out, researchers have extensively worked and examined alternative methods to find out the most efficient sales prediction methodology, and they have identified important features of time series data to enable the sales

prediction methodology to forecast sales efficiently. This has been done by using statistically-based methods like Linear Autoregressive models (AR) which are flexible to model many stationary processes (Hossen et al., 2017). The ARMA (Autoregressive Moving Average) model is used for short-term time series forecasting. The ARMA model failed because it only gave a linear relationship between features and could not accurately predict the evolution of non-linear and non-stationary data. Whenever there is highly fluctuating time series data due to seasonal factors or time trends it shows a degraded performance (Hossen et al., 2017). Therefore, most statistical methods are limited to non-linear and stationary time series forecasting assuming an AR type structure. To overcome the challenge of linear statistical time-series models, many non-linear machine-learning models like artificial neural networks (ANNs) have been proposed in the literature (Aras & Kocakoç, 2016; Kamruzzaman et al., 1 C.E.). ANNs belong to the data-driven approach, where training depends on the available data with little prior rationalization regarding relationships between variables (Aras & Kocakoç, 2016). ANNs do not make any assumptions about the statistical distributions of the underlying time series and they can naturally perform non-linear modeling (Aras & Kocakoç, 2016). As a result, ANNs are self-adaptive by nature. Recently we have seen the higher performance of Artificial Neural Networks (ANN) in classification and regression problems and have received focused attention in the time series forecasting methods. When we compare ANN with the normal statistical techniques, we find that ANN has many unique features such as: 1) non-linear and data-driven, 2) Not having a requirement for an explicit underlying model and 3) it is applicable to complicated models.

A literature review was performed in this study to identify a suitable neural network model for drug sales forecasting. To handle this work, some of the methods such as Machine Learning models, hybrid models, and statistical models will be helpful. (Dai et al., 2008) Have implemented an Artificial Neural Network for time series forecasting. They tried to study the use of artificial Neural Network in time series forecasting. They proposed that ARNN gives the best results to predict consumer goods' sales compared to SVM and Arima. ANNs are also known for their ability to map non-linear functions, making them suitable for sales forecasting. (Bing et al., 2014) proposed an algorithm to predict the stock price movement with an accuracy up to 76.12% by investigating public social media information represented in tweets data. Bing used a model to analyze public tweets and hourly stock price trends. NLP techniques have been used along with data mining techniques to identify correlation patterns between public sentiment and numeric stock prices. This study examines whether there is an internal association in the multilayer hierarchical structures, and found a relation between internal layers and the top layer of unstructured data. This study considers only daily closing values for historical stock prices. (Islek & Oguducu, 2015)studied with the use of bipartisan graphic clusters that clustered different warehouses according to sales behavior. They discussed the application by applying the Bayesian network algorithm in which they managed to produce the enhanced forecasting experience. (Siami-Namini & Namin, 2018) compared the accuracy of ARIMA and LSTM when it was forecasting time series data as representative techniques. These two techniques have been implemented and applied to a set of financial data, and the results have shown that LSTM is superior to ARIMA. (Kraus & Feuerriegel, 2017) used LSTM with transfer learning using text mining through financial news and the stock market data, Similarly, (Ding et al., n.d.) implemented Deep Stock Ranker, an LSTM based model for stock ranking using 11 technical indicators. (Mehdiyev et al., 2017) proposed a new multi-stage approach to deep learning for multivariate time series classification issues. They used the stacked LSTM autoencoders after extracting the features from the time series data in an unsupervised manner. The objective of the case study is to predict post-processing activities depending on the detected steel surface defects using the time series data obtained from the sensors installed in various positions of the steel casting process facility and the steel's chemical properties.

Meanwhile, in some of the papers, CNN models were preferred. (Kim et al., 2016) use CNNs in a bank telemarketing case study, whereby the aim is to predict whether a customer will take up a particular marketing campaign based on a number of numeric and nominal features per customer. The results for this study yield an impressive 76.70% accuracy, which yields the highest accuracy amongst 7 classifiers. In order to incorporate external features in the forecasting model. In (Sci & 2014, n.d.) the author uses CNNs to predict stock price changes based on the image of the time series plot. The author also attempts to color code the time series, however, the results of this approach were not positive. (Ding et al., n.d.) designed a model for stock market prediction driven by events. First, events are extracted from financial news and represented by word embedding as dense vectors. They trained a deep CNN to model on

stock price events both short term and long-term influences. Their proposed model in S&P 500 index prediction and individual stock prediction gave better performance than SVM. They also use a deep convolutional neural network to model short- and long-term influences of events of stock price movements. Results from this study show that CNNs can capture longer-term influence of news events than standard feed-forward networks. The authors of (Processing & 2017, 2017)used 250 features for the prediction of the private brokerage company's real data of risky transactions. They used CNN and LSTM for stock price forecasting. (Zhang et al., 2003) use convolutional neural networks for recognition of human activity (HAR). Their methodology capitalizes on the fact that a combination of unsupervised learning and supervised classification of features can increase the discriminative power of features. (Hernández et al., 2016) proposed an Auto-encoder and MLP based deep learning model to predict the daily accumulated rainfall. The authors used Auto-encoder to select features and MLP to predict. The results showed a better performance of their proposed model than other approaches. To make predictions, they used the deep feed-forward neural network. From the above literature review LSTM, CNN, and MLP are recommended as a deep learning models used for sales forecasting. Previously, most of the studies focused on considering the metrics as a mean absolute error, mean squared error, root mean squared error and k-fold cross-validation is used for training and testing data. Metrics like mean absolute error and root mean squared error are considered in this research. In this study, a stratified K-fold cross-validation technique is used for training and testing to increase the results' efficiency. In this study, LSTM, CNN, and MLP are chosen for sales forecasting.

## 4. Methodology

An experiment is chosen as a research method to answer the research questions because the experiment is considered a suitable research method for dealing with quantitative data as experiments would give greater control over variables. The experiments aim to evaluate the performance of Deep neural networks Multilayer Perceptron, Convolutional Neural Network, and Long Short-Term Memory on quarterly sales data extracted from the order management database of the private pharmaceutical sales in the Iraqi market. The efficient algorithm among the chosen algorithms is identified by analyzing and comparing the experiment results on the given dataset. We can describe the procedure followed in this experiment as follows: 1) Extracting the required data for the sales. 2) Applying Multilayer Perceptron, CNN, and LSTM algorithms. 3) The performance of the output can be enhanced by comparing metrics such as Mean Absolute Error and Root Mean Square Error, and 4) Based on assessment tests, the best suitable algorithm can be selected. We used Python, Pandas Sklearn and seaborn libraries to accomplish this work.

### 4.1 Study area and Dataset

The importance of data in the Pharmaceutical Industry is growing rapidly. Market data is one of the key factors that help companies monitor their performance and make strategic decisions to improve their efficiency. In this study, the private pharmaceutical market in Iraq (which is one of the most challenging markets in MENA) is selected to be the study area. Dataset is provided by Advanced Marketing Statistics company AMS. The data is collected quarterly at the SKU level (Stock Keeping Units) in volume and price (price is at pharmacy purchase level; i.e.: from drug store to pharmacy). The data covers the period from Q3-2010 To Q1-2021. AMS data reflects the quantitative sales of pharmaceutical products in the private sector of the Iraqi Market. It doesn't include tenders or governmental sales data. This dataset covered medicines that have had sales in the Iraqi market from at least 2013 which consists of about 80k records:

**Table 1**. Data set description

| No. | Features | Description |
|---|---|---|
| 1 | ATC2 | Anatomical Therapeutic Area at level 2 |
| 2 | Company Type | Gx / Rx |
| 3 | Product | Brand Name |
| 4 | Molecule | Active Ingredient of the product |
| 5 | Product Source | Legal/ Smuggled |
| 6 | Launch Date | The quarter at which a product has been first marketed |
| 7 | No. of Competitors | No. of other brands of the same molecule that are available in the market |
| 8 | Quantity | No. of standard units sold in a specific period of time |
| 9 | Price | Standard unit price in USD |
| 10 | Year | 2010 to 2021 |
| 11 | Quarter | Q1(Jan, Feb, March) Q2(April, May, June) Q3(July, Aug, Sep) Q4(Oct, Nov, Dec) |

The features in the previous dataset represent drug description, and sales are recorded on a quarterly basis. This data set might contain discrepancies in the names or codes, and also might contain missed data or outliers or errors; So, we should prepare data as in data preprocessing. The dataset is then transformed into a

supervised learning dataset using the sliding window approach.

## 4.2 Data Preprocessing

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. Data preprocessing is used in database-driven applications such as customer relationship management and rule-based applications (like neural networks). In Machine Learning (ML) processes, data preprocessing is critical to encoding the dataset in a form that could be interpreted and parsed by the algorithm (Zhang et al., 2003). Data goes through a series of steps during preprocessing: 1) Data Cleaning: Data is cleansed through processes such as filling in missing values or deleting rows with missing data, smoothing the noisy data, or resolving the inconsistencies in the data (García et al., 2016). 2) Data Integration: Data with different representations are put together and conflicts within the data are resolved (García et al., 2016).

## 4.3 Sliding Window

Time Series data can be transformed into a supervised learning problem by making use of Sliding Window Method. This transformation will enable us to use standard linear and nonlinear machine learning algorithms. A time series dataset can be transformed by making use of previous time steps as inputs and new time steps as output variables (Benidis et al., 2022).

## 4.4 Kwiatkowski, Phillips, Schmidt and Shin (KPSS) Test

Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) Test is a statistical test that is used to test whether a time series is stationary or not. It is used in this study to determine whether a time series is trend stationery or it consists of a unit root. The null hypothesis of the test is that the time series is stationary or trend stationary towards and the alternative hypothesis of a unit root series.

Null Hypothesis: If null hypothesis is accepted, the time series is considered to be trend stationary. Alternative Hypothesis: If the null hypothesis is rejected, the time series consists of unit root, meaning it is non-stationary. The results from the tests can be interpreted as follows: If p-value > 0.05 the null hypothesis is not rejected; the time series is trend stationary. Otherwise, the null hypothesis is rejected, and the time series consists of a unit root i.e., the series is non-stationary.

## 4.5 Feature Selection:

Feature selection refers to a class of methods for assigning values to input features to a predictive model which determines the relative significance of each factor while forecasting. Feature selection scores provide an overview of the model. Most significant scores are determined using a prediction approach that was fitted to the dataset. Inspecting the score of importance gives insight into that particular model and what features are the most essential and least important to the model while making a prediction. Feature Importance can be used to enhance a predictive model. This can be accomplished by selecting those features to remove (lowest scores) or those features to retain, using the importance scores. This is a type of selection of features that can simplify the modeling problem, accelerate the modeling process, and in certain cases, improve model performance.

## 4.6 Performance Metrics:

Performance Metrics should be selected depending on the regression problem and the dataset used for the experiment. Several metrics can be used while evaluating how well a model is performing. It is necessary to understand how each metric is measured to select the evaluation metric to better assess the model. This thesis's main objective is to compare the performance of neural network algorithms by evaluating all of these performance metrics such as Root Mean Square Error, and Mean Absolute Error. Root Mean Square Error as shown in equation 2 is the square root of the difference between the values predicted by the model and the real or observed values. The value of RMSE indicates the fit of the models on a particular dataset. Values close to zero implies a better fit thereby reducing the impact of outliers

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - yi\hat{})} \qquad (2)$$

Where '$n$' represents the feature variables, '$y_i$' represents actual values and '$y_i\hat{}$' represents the predicted or forecasted value.

Mean Absolute Error as shown in equation 3 is calculated by taking the average of absolute difference between values predicted by the model and the real or actual values. Similarly, the accuracy of the model is higher when MAE values are close to zero.

$$MEA(x,y) = \frac{1}{\text{Nsamples}} \sum_{i=0}^{\text{Nsamples}-1} |yi - xi| \quad (3)$$

## 4.7 Walk Forward Validation:

In time series modeling, the predictions over time become less and less accurate and hence it is a more realistic approach to re-train the model with actual data as it becomes available for further predictions. Since training of statistical models are not time consuming, walk-forward validation with sliding windows is the most preferred solution to get the most accurate results as shown in Figure 4. Walk Forward Validation by sliding windows is a re-sampling technique used to evaluate the machine learning model because it keeps the temporal order in the dataset while splitting time series data. The data is divided into training and test splits as fixed window size (*n*) training set (*1-n-1*) Test set (*n*) for each time then window slide one step or time stamp and predict (*n*) value. each window evaluated by RMSE & MAE, these results are stored to get the average of our performance metrics (*RMSE & MAE*).

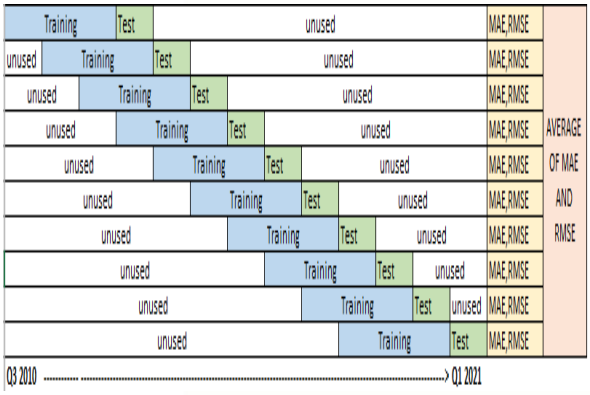



**Fig 4**. Walk Forward Validation

## 5. Results

## 5.1 Stationarity Test

Kwiatkowski, Phillips, Schmidt, and Shin Test was used to test the stationarity of the dataset and the results obtained are shown in Figure 5 and Figure 6.

```
KPSS Test Results (Quantity):
Results of KPSS Test:
Test Statistic            0.09931
p-value                   0.10000
Lags Used                10.00000
Critical Value (10%)      0.34700
Critical Value (5%)       0.46300
Critical Value (2.5%)     0.57400
Critical Value (1%)       0.73900
dtype: float64
```

**Fig** 5. Kwiatkowski, Phillips, Schmidt and Shin Test (Quantity)

Based on Figure 5, the p-value is greater than 0.05 So, the null hypothesis cannot be rejected. This indicates the data is stationary which can further be utilized for performing time series analysis.

```
KPSS Test Results (PriceUSD):
Results of KPSS Test:
Test Statistic            0.495799
p-value                   0.042613
Lags Used                10.000000
Critical Value (10%)      0.347000
Critical Value (5%)       0.463000
Critical Value (2.5%)     0.574000
Critical Value (1%)       0.739000
dtype: float64
```

**Fig 5**. Kwiatkowski, Phillips, Schmidt, and Shin Test (USD PRICE)

Based on Figure 6, the p-value is less than 0.05 So, the null hypothesis is rejected. This indicates that the data is non-stationary. However, we want to test Deep learning models to perform the analysis of this type of time series. LSTM, MLP, and CNN have been used for forecasting sales and prices of medicines in the Iraqi pharmaceutical market, and the following results depict how LSTM, MLP, and CNN have performed based on various performance metrics used in this thesis.

## 5.2 Learning curve

Learning curves are used to evaluate the performance of a model to diagnose whether the model underfits, overfits, or is a good fit on the chosen dataset. They can also be used to determine if the statistical properties of the training dataset are relative to the properties of the validation dataset. Learning curves are used to identify learning performance changes by taking a plot of loss or error over time. It can also be used to identify how well a model generalizes to unseen data by using a validation dataset. Based on the structure and dynamics of the learning curve, the configuration of the model can be changed to enhance the learning and performance of the model.

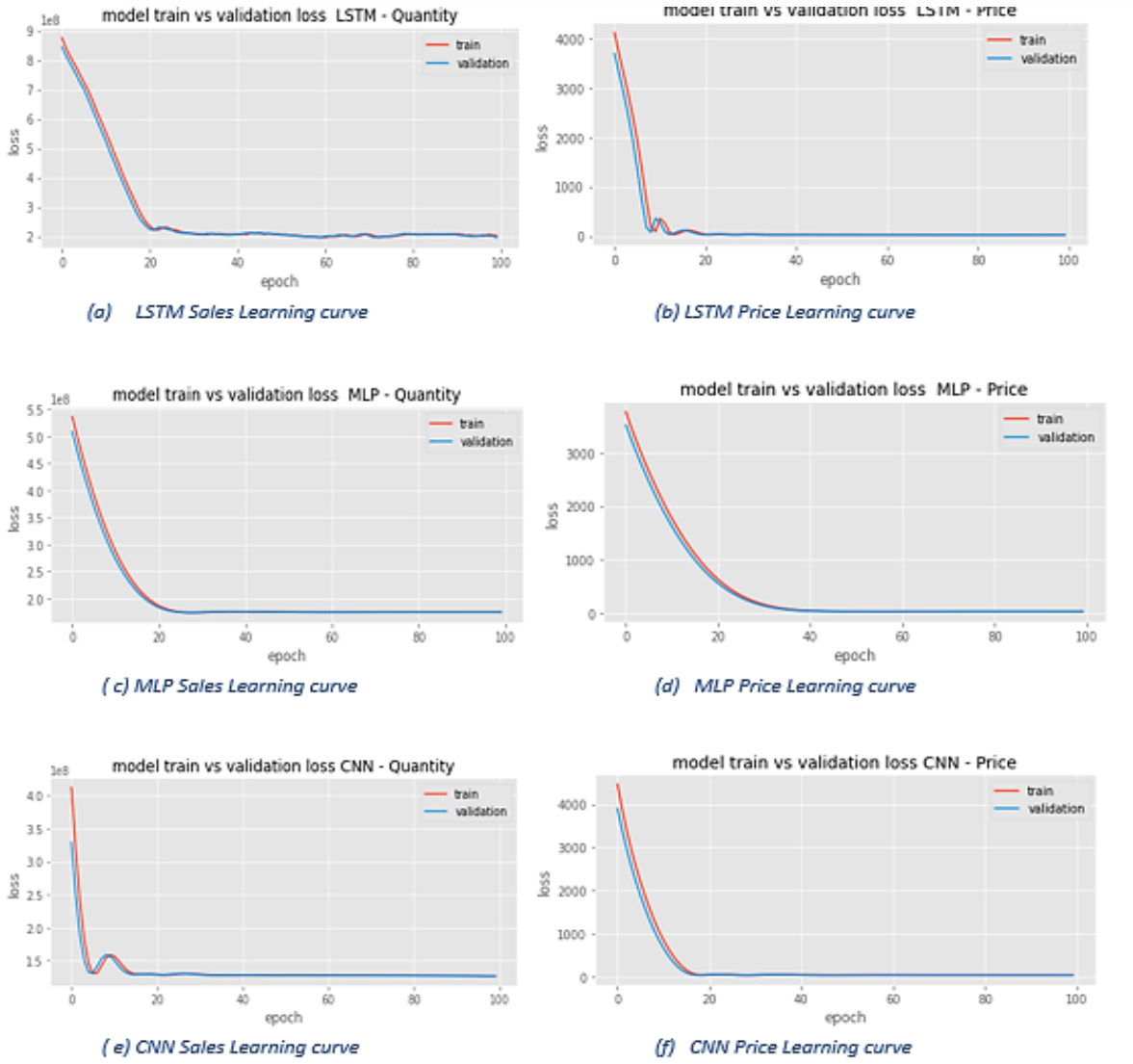



**Fig** 6. (LSTM, MLP, CNN) Learning curves

**Figure 7** (a - f) represents the Learning curves of LSTM, MLP, and CNN for Sales and Prices. We can see that the training and validation errors decrease to an optimal point with a minimal difference in loss values between them, indicating that the used ANN models are a good fit.

## 5.3 Forecast results

LSTM, MLP & CNN are trained using walk forward validation approach and the data is split in 80% to 20% ratio where 80% was used for the training set and 20% was used as the test set. The training set is further split into training and validation sets for in 5 steps as the used approach. The performance of these models is estimated using the performance metrics RMSE and MAE and the results obtained are shown below for each type of data sets, the used dataset in this thesis was classified based on the source of medicines in the market as below:

1- Actual (To Market) dataset. (40% OF DATASET)
2 -Projected medicines (In Market). (40% OF DATASET)
3- Smuggled medicines. (20% OF DATASET)

DOI: 10.6977/IJoSI.202409_8(3).0006
N. Qassrawi, M. Azzeh /Int. J. Systematic Innovation, 8(3), 63-83 (2024)

**Table 2**. Performance metrics for Actual medicines (To Market) dataset.

| ID | Medicine | LSTM | | | | MLP | | | | CNN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Sales (k) | | Price USD | | Sales (k) | | Price USD | | Sales (k) | | Price USD | |
| | | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| 1 | ADOL-TAB 500MG 96`S | 1.9 | 2.5 | 0.07 | 0.14 | 55.2 | 86.0 | 0.12 | 0.22 | 18.4 | 29.1 | 0.06 | 0.13 |
| 2 | AMARYL-TAB 2MG 30'S | 0.30 | 0.38 | 0.38 | 0.66 | 17.6 | 21.8 | 0.70 | 0.97 | 1.63 | 2.83 | 0.41 | 0.66 |
| 3 | SUPRAX-CAP 200MG 8`S | 0.33 | 0.61 | 0.35 | 0.47 | 1.80 | 2.95 | 0.50 | 0.66 | 0.39 | 0.61 | 0.31 | 0.43 |
| 4 | CONCOR-TAB 10MG 30`S | 0.005 | 0.006 | 0.21 | 0.48 | 10.3 | 12.6 | 0.52 | 0.87 | 2.46 | 3.69 | 0.47 | 0.79 |
| 5 | NOVATEN-TAB 100MG 28`S | 1.23 | 1.54 | 0.005 | 0.009 | 34.2 | 49.0 | 0.006 | 0.01 | 16.8 | 20.9 | 0.006 | 0.009 |
| 6 | CRESTOR-TAB 10 MG 28'S | 0.63 | 0.99 | 0.85 | 1.95 | 5.85 | 7.28 | 2.52 | 4.35 | 1.96 | 3.19 | 1.39 | 2.47 |
| 7 | ZESTRIL-TAB 5MG 28'S | 1.81 | 2.97 | 0.26 | 0.56 | 3.88 | 4.90 | 0.81 | 1.11 | 1.08 | 1.93 | 0.22 | 0.51 |
| 8 | APROVEL-TAB 150MG 28`S | 0.12 | 0.22 | 1.60 | 2.34 | 1.22 | 1.51 | 2.68 | 3.74 | 0.17 | 0.72 | 1.65 | 2.79 |
| 9 | PLAGIN-TAB 75MG 30'S | 0.49 | 0.60 | 0.12 | 0.14 | 1.57 | 2.11 | 0.12 | 0.16 | 0.60 | 0.87 | 0.13 | 0.14 |
| 10 | ATACAND-TAB 8MG 28`S | 1.29 | 2.09 | 0.83 | 1.62 | 10.99 | 13.94 | 1.46 | 2.42 | 4.41 | 6.05 | 1.00 | 1.82 |
| 11 | SUPRAX-CAP 200MG 8`S | 0.16 | 0.26 | 0.35 | 0.47 | 1.17 | 1.61 | 0.51 | 0.66 | 0.29 | 0.54 | 0.31 | 0.43 |
| 12 | PLATIL-TAB 75MG 30'S | 0.024 | 0.034 | 0.73 | 1.05 | 14.35 | 17.91 | 0.70 | 1.31 | 4.53 | 6.02 | 0.52 | 1.02 |
| 13 | MYOGESIC-TAB 450MG 20`S | 0.26 | 0.34 | 0.09 | 0.17 | 62.02 | 80.85 | 0.19 | 0.27 | 27.93 | 37.96 | 0.10 | 0.17 |
| 14 | JOINTACE-CAP 30'S | 0.075 | 0.124 | 0.20 | 0.29 | 3.96 | 5.42 | 0.26 | 0.42 | 1.99 | 3.22 | 0.194 | 0.289 |
| 15 | NEXIUM-TAB 40MG 14`S | 2.41 | 3.07 | 0.31 | 0.89 | 34.39 | 42.02 | 1.85 | 2.74 | 11.36 | 18.27 | 0.47 | 1.02 |
| Average Performance Metrics | | 0.74 | 1.05 | 0.42 | 0.75 | 17.23 | 23.33 | 0.86 | 1.33 | 6.27 | 9.06 | 0.48 | 0.85 |

This table obtained the MAE and RMSE for Sales & USD Prices for 15 products Selected randomly from the Actual dataset. Outliers in Actual (To-Market) Data have been overcome by dividing the sales data by 2. Suppliers normally ship large quantities of drugs to secure enough stock for at least six months. **Figure 8** below represents the actual sales and the forecasted sales of NEXIUM-TAB 40MG 14`S using LSTM, MLP, and CNN algorithms, where the green line indicates the predicted value of the target variant and the blue line represents the actual values.

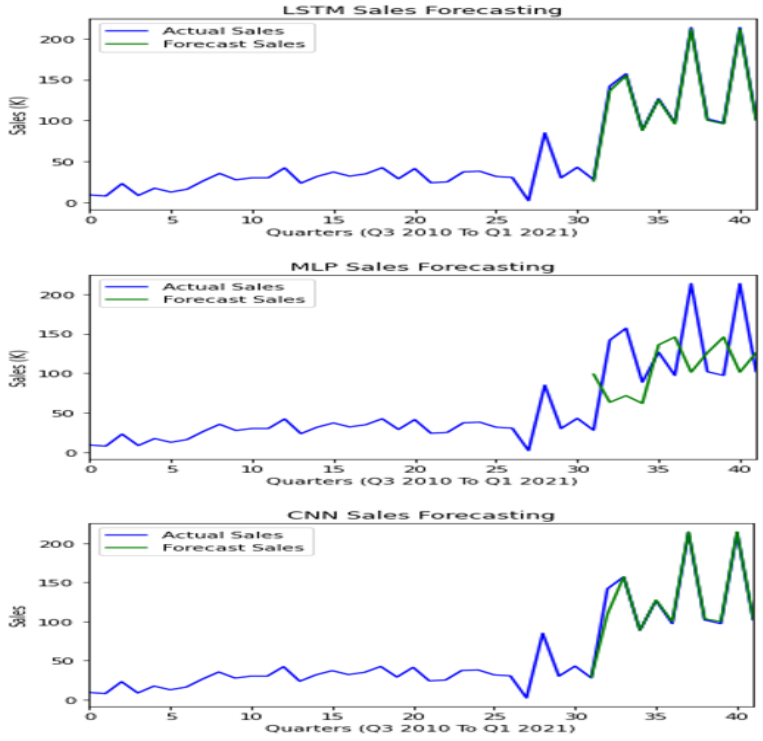



**Fig 7.** NEXIUM-TAB 40MG 14`S (Actual Sales (k))

**Figure 8** represents the actual prices and the forecasted prices of **NEXIUM-TAB 40MG 14`S** using LSTM, MLP, and CNN algorithms, where the green line indicates the predicted price of the target variant and the blue line represents the actual price.

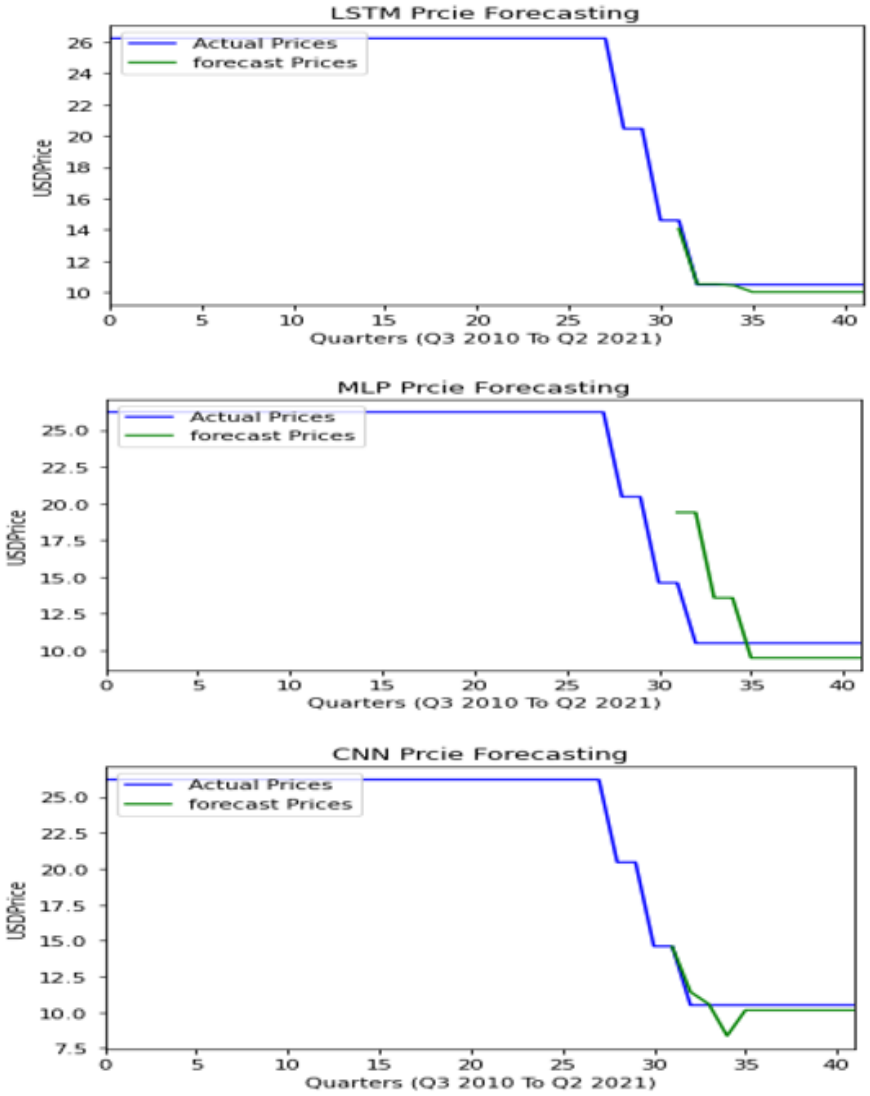



**Fig** 8. NEXIUM-TAB 40MG 14`S (Actual USD Price)

**Table 3** obtained the MAE and RMSE for Sales & USD Prices for 15 products Selected randomly from the Projected dataset. Outliers replaced by "Mean" value.

**Table 3.** Performance metrics for Projected medicines

| ID | Medicine | LSTM | | | | MLP | | | | CNN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Sales (k) | | Price USD | | Sales (k) | | Price USD | | Sales (k) | | Price USD | |
| | | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| 1 | ACTIFED-TAB 30`S | 0.037 | 0.045 | 0.036 | 0.044 | 3.641 | 4.830 | 0.06 | 0.14 | 0.196 | 0.442 | 0.038 | 0.044 |
| 2 | ALDACTON-TAB 25MG 20 'S | 0.027 | 0.036 | 0.20 | 0.23 | 3.66 | 4.55 | 0.32 | 0.53 | 0.420 | 0.727 | 0.223 | 0.335 |
| 3 | MOTILIOSYR-TAB 10MG 30`S | 1.483 | 3.250 | 0.047 | 0.100 | 16.31 | 20.32 | 0.079 | 0.158 | 5.015 | 7.64 | 0.047 | 0.100 |
| 4 | TOFRANIL-TAB 10 MG 50`S | 0.015 | 0.036 | 0.315 | 0.530 | 2.68 | 3.86 | 0.59 | 0.80 | 0.317 | 0.484 | 0.312 | 0.529 |
| 5 | LIBROXIDE-TAB 10MG 10`S | 6.19 | 8.78 | 0.007 | 0.008 | 34.25 | 42.66 | 0.007 | 0.009 | 8.62 | 13.89 | 0.007 | 0.008 |
| 6 | XENICAL-CAP 120MG 84'S | 0.011 | 0.022 | 1.439 | 2.965 | 1.238 | 1.666 | 12.80 | 18.66 | 0.093 | 0.195 | 4.769 | 6.866 |
| 7 | CO-AMOXI-1000 TAB 12`S | 1.62 | 4.64 | 0.240 | 0.378 | 25.64 | 32.74 | 0.355 | 0.593 | 13.59 | 16.92 | 0.285 | 0.439 |
| 8 | LOFRAL-5 OP 30'S | 0.050 | 0.078 | 0.120 | 0.159 | 15.03 | 18.35 | 0.145 | 0.208 | 6.975 | 9.375 | 0.138 | 0.169 |
| 9 | BUTADIN-TAB 2MG 10`S | 5.209 | 8.49 | 0.002 | 0.003 | 103.7 | 145.4 | 0.002 | 0.003 | 52.29 | 80.26 | 0.001 | 0.002 |
| 10 | IMURAN-TAB 50MG 100`S | 0.062 | 0.094 | 1.579 | 2.439 | 2.16 | 2.83 | 2.38 | 3.43 | 0.605 | 0.654 | 1.69 | 2.53 |
| 11 | DAKTACORT-CREAM 15G | 0.282 | 0.360 | 0.346 | 0.739 | 10.07 | 11.56 | 0.639 | 0.99 | 4.35 | 6.72 | 0.434 | 0.758 |
| 12 | CATAFLAM-TAB 50MG 20`S | 0.261 | 0.665 | 0.260 | 0.410 | 13.72 | 19.08 | 0.319 | 0.527 | 5.76 | 9.43 | 0.146 | 0.349 |
| 13 | ADVIL SINUS-CAP 20`S | 0.71 | 0.826 | 0.242 | 0.372 | 12.40 | 14.54 | 0.46 | 0.635 | 2.298 | 3.748 | 0.239 | 0.365 |
| 14 | RABEZOLE-TAB 20MG 28`S | 0.106 | 0.129 | 0.269 | 0.455 | 6.29 | 11.76 | 0.357 | 0.718 | 2.83 | 6.173 | 0.236 | 0.453 |
| 15 | DILATREND-TAB 25MG 30`S | 0.510 | 0.810 | 1.16 | 1.95 | 2.18 | 2.74 | 1.73 | 2.55 | 0.605 | 0.905 | 1.255 | 2.065 |
| Average Performance Metrics | | 1.10 | 1.88 | 0.42 | 0.72 | 16.86 | 22.46 | 1.35 | 2.00 | 6.93 | 10.5 | 0.65 | 1.00 |

**Figure 9** represents the actual sales and the forecasted sales of CO-AMOXI-1000 TAB 12`S using LSTM, MLP, and CNN algorithms, where the green line indicates the predicted value of the target variant and the blue line represents the actual values.

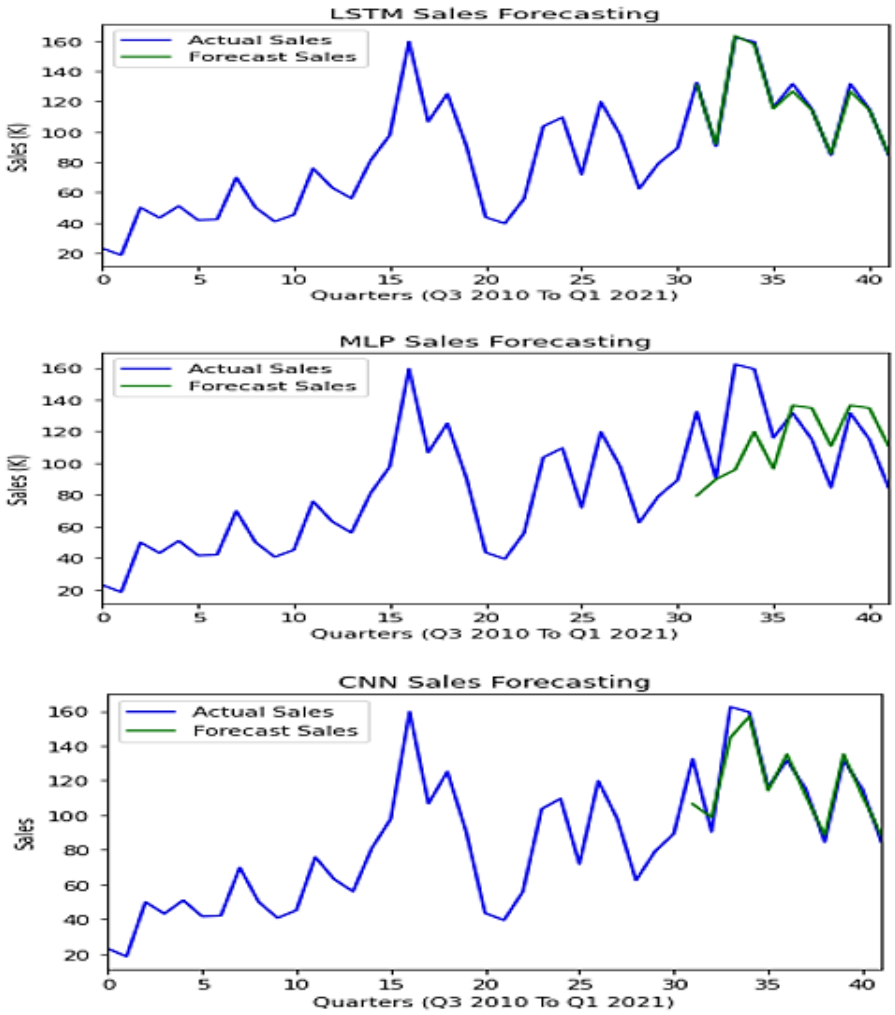



**Fig** 9. CO-AMOXI-1000 TAB 12`S (Projected Sales(k))

**Figure 10** represents the actual prices and the forecasted prices of CO-AMOXI-1000 TAB 12`S using LSTM, MLP, and CNN algorithms, where the green line indicates the predicted price of the target variant and the blue line represents the actual price.

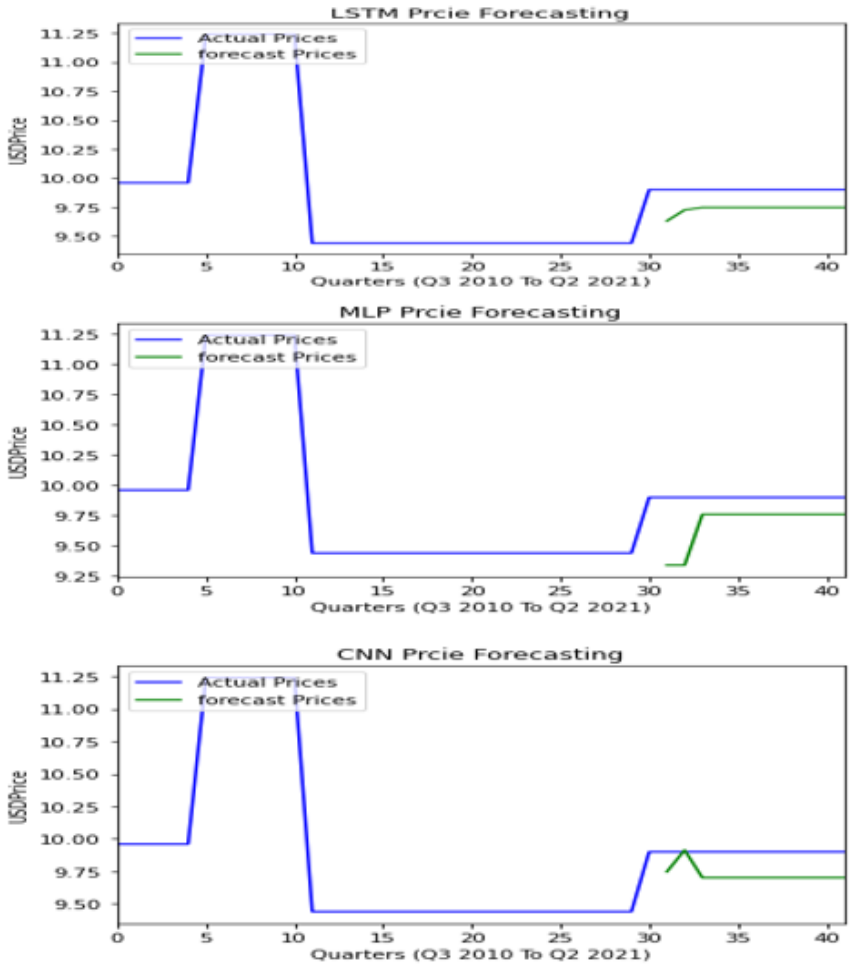



**Fig** 10. CO-AMOXI-1000 TAB 12`S (USD Price)

**Table 4** obtained the MAE and RMSE for Sales & USD Prices for 15 products Selected randomly from the Smuggled dataset. Figure 11 represents the actual sales and the forecasted sales of NEXIUM-TAB 40MG 14`S using LSTM, MLP, and CNN algorithms, where the green line indicates the predicted value of the target variant and the blue line represents the actual values.

**Table 4**. Performance metrics for Smuggled medicines

| ID | Medicine | LSTM | | | | MLP | | | | CNN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Sales (k) | | Price USD | | Sales (k) | | Price USD | | Sales (k) | | Price USD | |
| | | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| 1 | TOFRANIL-TAB 25MG 50'S | 0.571 | 0.842 | 0.222 | 0.459 | 9.05 | 12.70 | 0.470 | 0.684 | 1.514 | 2.195 | 0.227 | 0.460 |
| 2 | ALDOMET-TAB 250MG 30`S | 0.550 | 0.757 | 0.201 | 0.363 | 5.68 | 6.689 | 0.257 | 0.570 | 2.46 | 3.76 | 0.198 | 0.372 |
| 3 | AMARYL-TAB 4MG 30'S | 1.324 | 1.461 | 0.516 | 0.794 | 12.15 | 17.54 | 0.926 | 1.176 | 2.449 | 3.307 | 0.535 | 0.800 |
| 4 | ACTIFED-TAB 12`S | 1.934 | 2.422 | 0.066 | 0.086 | 32.68 | 42.68 | 0.067 | 0.092 | 12.48 | 20.6 | 0.055 | 0.075 |
| 5 | ISOPTIN-TAB 40MG 30'S | 0.155 | 0.345 | 0.261 | 0.523 | 1.60 | 2.10 | 0.74 | 0.89 | 0.169 | 0.308 | 0.263 | 0.530 |
| 6 | NEXIUM-TAB 40MG 14`S | 0.022 | 0.031 | 0.703 | 1.13 | 3.46 | 4.01 | 1.3 | 1.95 | 0.850 | 1.42 | 0.77 | 1.19 |
| 7 | PLAVIX-TAB 75MG 28`S | 0.44 | 0.52 | 1.317 | 2.179 | 3.75 | 5.064 | 3.574 | 4.356 | 0.28 | 0.52 | 1.90 | 2.87 |
| Average Performance Metrics | | 0.71 | 0.91 | 0.47 | 0.79 | 9.77 | 12.97 | 1.05 | 1.39 | 2.89 | 4.59 | 0.59 | 0.90 |

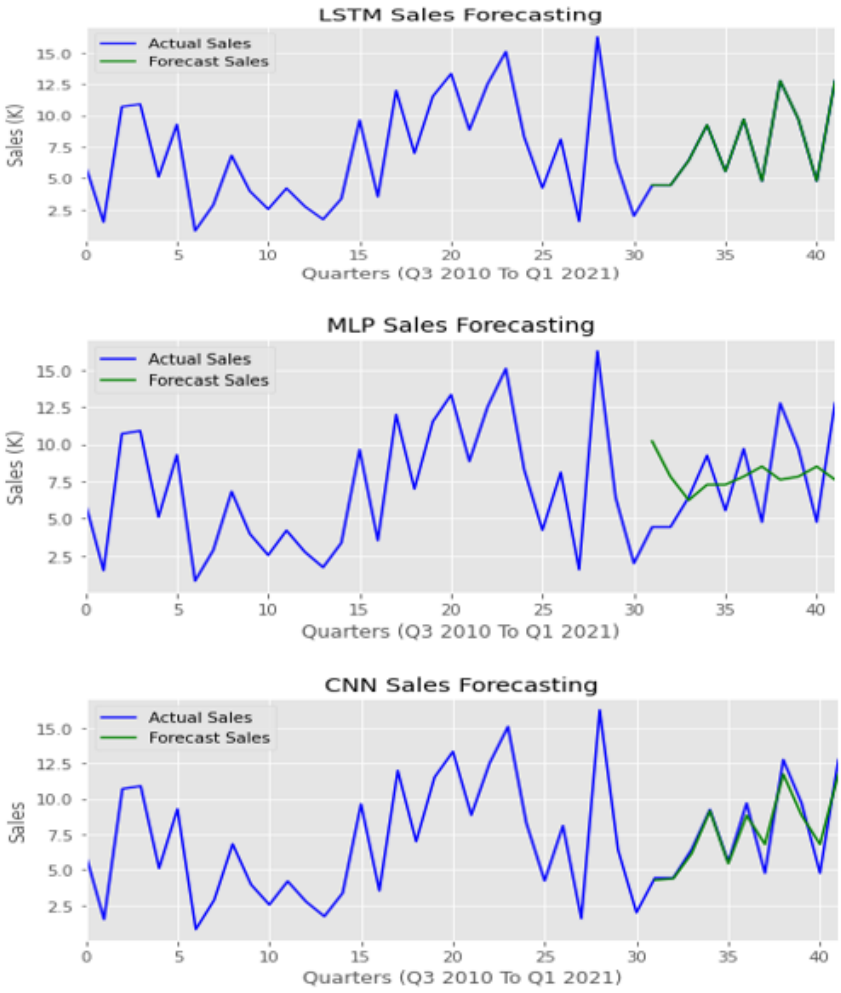



**Fig** 11. NEXIUM-TAB 40MG 14`S (Smuggled Sales forecasting)

**Figure 12** represents the actual prices and the forecasted prices of NEXIUM-TAB 40MG 14`S using LSTM, MLP, and CNN algorithms, where the green line indicates the predicted price of the target variant and the blue line represents the actual price.

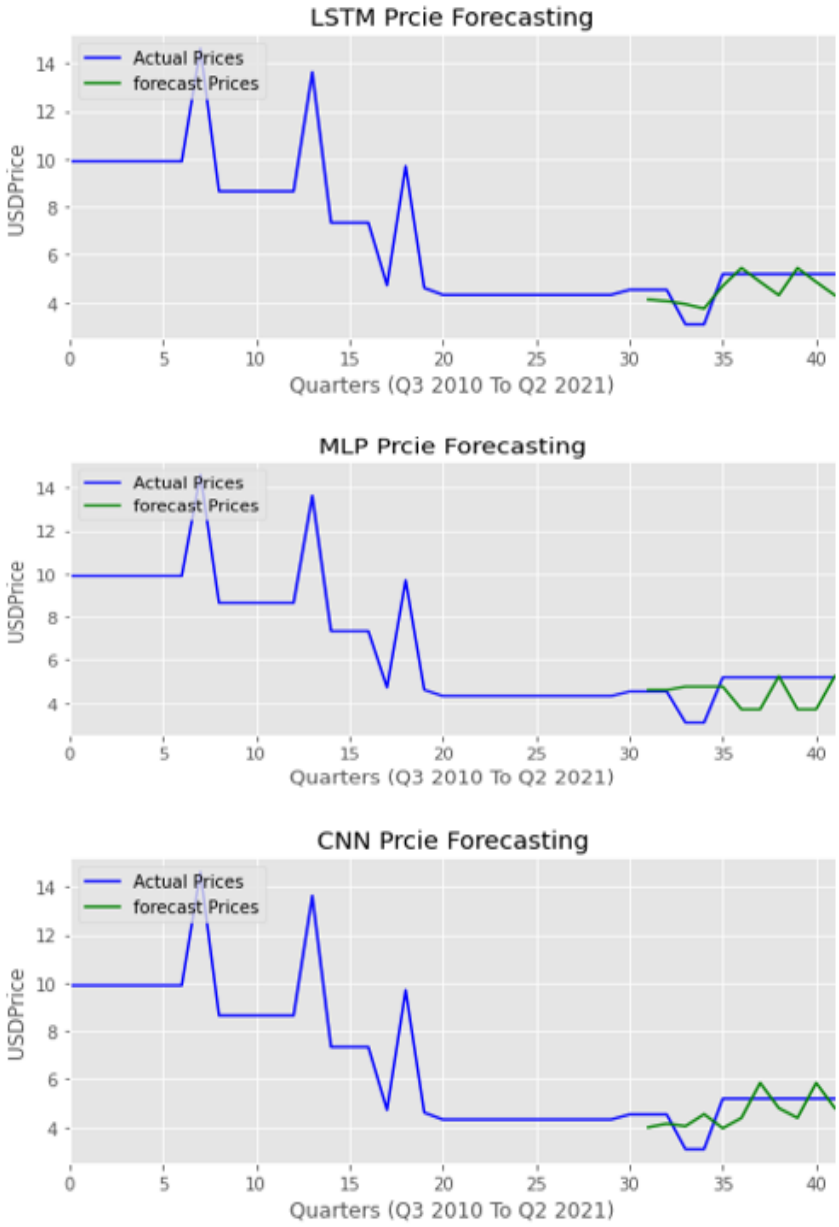



**Fig** 12. NEXIUM-TAB 40MG 14`S (Smuggled USD Price)

## 5.3 Analysis of experiment results

**Figure 14** represents the average mean absolute error of Long Short-Term Memory, Multilayer Perceptron, and Convolutional Neural Network algorithms for Sales forecasting for the three types of data. Long Short-Term memory performed better in the three types of data when compared to Multilayer Perceptron and Convolutional Neural networks. Multilayer Perceptron has the highest average error across the three types of data.

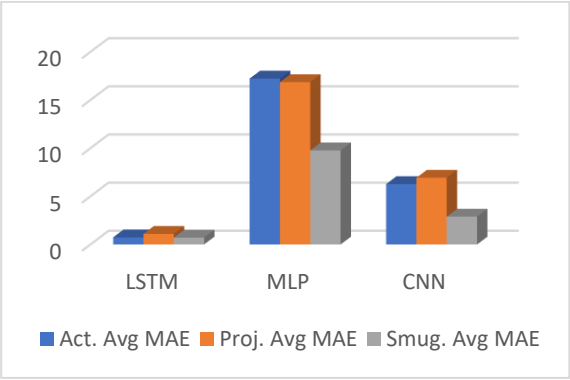



**Fig** 13. Mean Absolute Error for Sales Forecasting

**Figure 15** represents the average mean absolute error of Long Short-Term Memory, Multilayer Perceptron, and Convolutional Neural Network algorithms for USD Price forecasting for the three types of data. Also, we can see that Long Short-Term memory performed better in the three types of data when compared to Multilayer Perceptron and Convolutional Neural networks. Multilayer Perceptron has the highest average error across the three types of data.

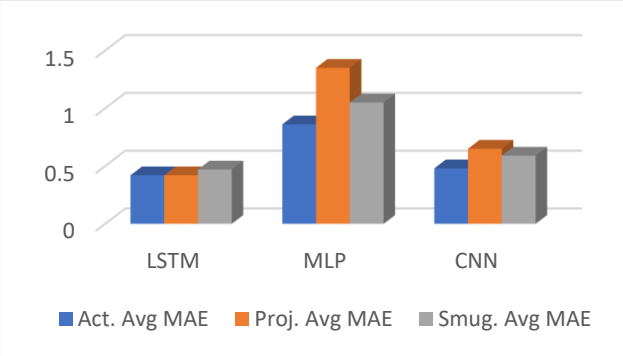



**Fig 14**. Mean Absolute Error for USD Price Forecasting

**Figure 16** represents the Root Mean Square error of Long Short-Term Memory, Multilayer Perceptron, and Convolutional Neural Network algorithms for Sales forecasting for the three types of data. Long Short-Term memory performed better in the three types of data compared to Multilayer Perceptron and Convolutional Neural networks. However, Multilayer Perceptron has the highest average error across the three types of data.

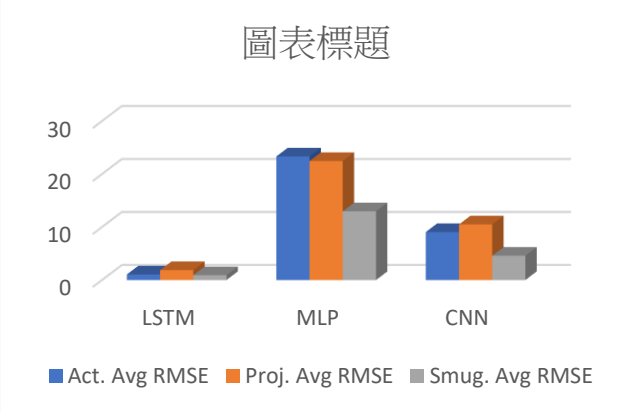



**Fig** 15. Root Mean Square Error for Sales Forecasting

**Figure 17** represents the Root Mean Square error of Long Short-Term Memory, Multilayer Perceptron, and Convolutional Neural Network algorithms for Price (USD) forecasting for the three types of data. Long Short-Term memory performed better in the three types of data compared to Multilayer Perceptron and Convolutional Neural networks. However, Multilayer Perceptron has the highest average error across the three types of data.

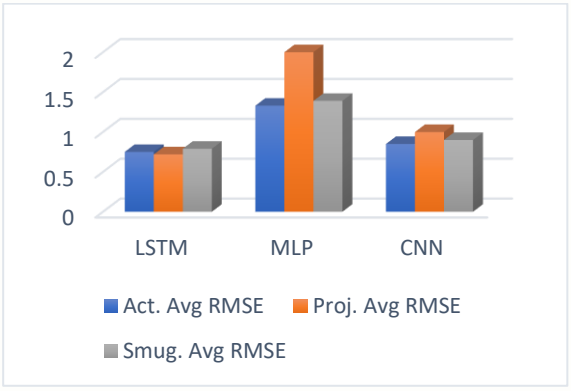



**Fig** 16. Root Mean Square Error for Price (USD) Forecasting

## 5.4 Performance Evaluation

The main objective of this thesis is to compare MLP, CNN, and LSTM forecasting accuracy in medicine sales and Prices in different data sources.

**Table 5**. Comparison of performance evaluation results for sales forecasting

| Algorithm | MAE (k) | RMSE (k) |
|---|---|---|
| LSTM | 0.85 | 1.28 |
| MLP | 14.62 | 19.58 |
| CNN | 5.363 | 8.05 |

The average error as shown in Table 5 indicates that LSTM based forecasting performed better than the other two algorithms with a Root Mean Square Error of about 1.28(k) and a Mean Absolute Error of about 0.85(k).

**Table 6.** Comparison of performance evaluation results for Price (USD) forecasting

| Algorithm | MAE | RMSE |
|---|---|---|
| LSTM | 0.44 | 0.75 |
| MLP | 1.09 | 1.57 |
| CNN | 0.57 | 0.92 |

The average error as shown in Table 6 indicates that LSTM based forecasting performed better than the other two algorithms with a Root Mean Square Error of about 0.75 and a Mean Absolute Error of about 0.44.

## 6. Discussion

**RQ1**. What are the critical features that influence drug sales?

Our objective was to forecast future sales and prices. Year & Quarter are the main features required to build date for the time series, Sales (Quantity), and USD Prices are our targets. Data sources also affect sales volumes and the change of the prices and each source has different behavior in the market. based on that we split our dataset into three other datasets to study them separately.

**RQ2**. How can the Deep Artificial Neural Network algorithms be chosen to resolve the sales forecasting problem?

Deep neural network algorithms had been selected based on the literature review and what is recommended for time series, we have chosen LSTM, MLP, and CNN and checked if these models have a good fit for the data by reviewing the learning curve.

**RQ3**. Which Deep Neural Network model is efficient for forecasting the sales of drugs in the private pharmaceutical market?

Long Short-Term Memory proved to be an efficient Deep learning algorithm. From the experiment, Long Short-Term memory performed better when compared to Multilayer Perceptron and Convolutional Neural networks. This is because LSTM makes use of temporal information from the data. The average Root Mean Square Error of Long Short-Term Memory for sales is 1.28(k) and Mean Absolute Error of about 0.85(k). The average Root Mean Square Error of Long Short-Term Memory for USD Prices is about 0.75, and the Mean Absolute Error is about 0.44. Which is less when compared to that of Multilayer Perceptron and Convolutional Neural networks.

Final remarks Different sources of data, moving stock between agents or pharmacies, the change of sample size & universe size in the different periods between (2010 to 2021) and the smuggled sales for agents and pharmacies coming from neighboring countries are the primary reasons that make data tracking very complex. The proposed model was implemented using Deep Artificial Neural Network algorithms, and the main objective was to obtain a suitable algorithm for sales and price forecasting from selected algorithms (LSTM, MLP, CNN).

## 7. Conclusion

Sales forecasting plays an important role in the business sector in every field. With the sales forecasts improved, sales revenue analysis will help decision-makers get the required details to estimate both the revenue and the income. Artificial Neural Networks have been selected among several machine learning models because they can effectively handle non-linear data. In this thesis, three Artificial Neural Network algorithms: Multilayer perceptron, Convolutional Neural Network, and Long Short-Term Memory are identified as fitting machine learning algorithms for drug sales and price predicting. The three algorithms are evaluated using RMSE and MAE performance metrics. The model with the lowest value is considered to be an efficient model for generating forecasts. Based on the results from the experiment, Long Short-Term Memory performed better than MLP and CNN for generating predictions with an average Root Mean Square Error of for sales is 1.28(k) and a Mean Absolute Error of about 0.85(k), and average Root Mean Square Error for USD Prices is about 0.75, and Mean Absolute Error is about 0.44. The forecasts are then used to adjust stock levels according to the predictions. As part of future work, the predictions can be enhanced by including a set of influential factors as feature variables in the dataset, such as the effect of smuggled sales on legal sales and pricing in the market, bonuses, and discounts offered on specific variants. These are a sample of the factors that could be used to understand the variations in time series data, allowing us to further improve the model performance to generate reliable forecasts. The potential future work is to build a recommended system by finding a relation between sales and the price change from historical data to recommend the best price for a new product in a specific ATC to make the best sales.

## Acknowledgment

The authors thank Applied Science Private University and Princess Sumaya University for Technology for their support to carry out this research project.

## References

Andrawis, R. R., Atiya, A. F., & El-Shishiny, H. (2011). Forecast combinations of computational intelligence and linear models for the NN5 time series forecasting competition. International Journal of Forecasting,

27(3), 672–688. https://doi.org/10.1016/J.IJFORECAST.2010.09.005

Aras, S., & Kocakoç, T. D. (2016). A new model selection strategy in time series forecasting with artificial neural networks: IHTS. Neurocomputing, 174, 974–987. https://doi.org/10.1016/J.NEUCOM.2015.10.036

Benidis, K., Sundar Rangapuram, S., Flunkert, V., Research, A., Yuyang Wang, G., Maddix, D., Caner Turkmen, U., Gasthaus, J., Bohlke-schneider, M., Salinas, D., Stella, L., Aubet, F., Callot, L., Wang, Y., Turkmen, C., Bohlke-Schneider, M., Aubet, F.-X., Turkmen, C., Gasthaus, J., … Maddix, D. (2022). Deep Learning for Time Series Forecasting: Tutorial and Literature Survey. ACM Computing Surveys, 55(6). https://doi.org/10.1145/3533382

Bing, L., Chan, K. C. C., & Ou, C. (2014). Public sentiment analysis in twitter data for prediction of a company's stock price movements. Proceedings - 11th IEEE International Conference on E-Business Engineering, ICEBE 2014 - Including 10th Workshop on Service-Oriented Applications, Integration and Collaboration, SOAIC 2014 and 1st Workshop on E-Commerce Engineering, ECE 2014, 232–239. https://doi.org/10.1109/ICEBE.2014.47

Boyapati, S. N., & Mummidi, R. (2020). Predicting sales using Machine Learning Techniques. https://www.diva-portal.org/smash/get/diva2:1455353/FULLTEXT02

Dai, J., Zhang, P., Mazumdar, J., Harley, R. G., & Venayagamoorthy, G. K. (2008). A comparison of MLP, RNN and ESN in determining harmonic contributions from nonlinear loads. IECON Proceedings (Industrial Electronics Conference), 3025–3032. https://doi.org/10.1109/IECON.2008.4758443

Ding, X., Zhang, Y., Liu, T., & Duan, J. (n.d.). Deep Learning for Event-Driven Stock Prediction.

Doganis, P., Alexandridis, A., Patrinos, P., & Sarimveis, H. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. Journal of Food Engineering, 75(2), 196–204. https://doi.org/10.1016/J.JFOODENG.2005.03.056

Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. International Journal of Information Management Data Insights, 2(1), 100058. https://doi.org/10.1016/J.JJIMEI.2022.100058

García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: methods and prospects. Big Data Analytics 2016 1:1, 1(1), 1–22. https://doi.org/10.1186/S41044-016-0014-0

Hernández, E., Sanchez-Anguix, V., Julian, V., Palanca, J., & Duque, N. (2016). Rainfall prediction: A deep learning approach. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9648, 151–162. https://doi.org/10.1007/978-3-319-32034-2_13/COVER

Hossen, T., Plathottam, S. J., Angamuthu, R. K., Ranganathan, P., & Salehfar, H. (2017). Short-term load forecasting using deep neural networks (DNN). 2017 North American Power Symposium, NAPS 2017. https://doi.org/10.1109/NAPS.2017.8107271

Islek, I., & Oguducu, S. G. (2015). A retail demand forecasting model based on data mining techniques. IEEE International Symposium on Industrial Electronics, 2015-September, 55–60. https://doi.org/10.1109/ISIE.2015.7281443

Ji, S., Yu, H., Guo, Y., & Zhang, Z. (2016). Research on sales forecasting based on ARIMA and BP neural network combined model. ACM International Conference Proceeding Series. https://doi.org/10.1145/3028842.3028883

Kamruzzaman, J., Sarker, R. A., & Begg, R. (1 C.E.). Artificial Neural Networks: Applications in Finance and Manufacturing. Https://Services.Igi-Global.Com/Resolvedoi/Resolve.Aspx?Doi=10.4018/978-1-59140-670-9.Ch001, 1–27. https://doi.org/10.4018/978-1-59140-670-9.CH001

Kim, K. H., Lee, C. S., Jo, S. M., & Cho, S. B. (2016). Predicting the success of bank telemarketing using deep convolutional neural network. Proceedings of the 2015 7th International Conference of Soft Computing and Pattern Recognition, SoCPaR 2015, 314–

317. https://doi.org/10.1109/SOCPAR.2015.7492828

Kraus, M., & Feuerriegel, S. (2017). Decision support from financial disclosures with deep neural networks and transfer learning. Decision Support Systems, 104, 38–48. https://doi.org/10.1016/J.DSS.2017.10.001

Lakner, Z., Kiss, A., Popp, J., Zéman, Z., Máté, D., & Oláh, J. (2019). From Basic Research to Competitiveness: An Econometric Analysis of the Global Pharmaceutical Sector. Sustainability 2019, Vol. 11, Page 3125, 11(11), 3125. https://doi.org/10.3390/SU11113125

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: a survey. Philosophical Transactions of the Royal Society A, 379(2194). https://doi.org/10.1098/RSTA.2020.0209

Mehdiyev, N., Lahann, J., Emrich, A., Enke, D., Fettke, P., & Loos, P. (2017). Time Series Classification using Deep Learning for Process Planning: A Case from the Process Industry. Procedia Computer Science, 114, 242–249. https://doi.org/10.1016/J.PROCS.2017.09.066

Processing, N. A.-F. in S., & 2017, undefined. (2017). Neural networks for financial market risk classification. Isaac-Scientific.ComN AbroyanFrontiers in Signal Processing, 2017•isaac-Scientific.Com, 1(2). https://doi.org/10.22606/fsp.2017.12002

Sci, A. S.-S. U. D. C., & 2014, undefined. (n.d.). Convolutional networks for stock trading. Cs231n.Stanford.EduA SiripurapuStanford Univ Dep Comput Sci, 2014•cs231n.Stanford.Edu. Retrieved September 21, 2023, from http://cs231n.stanford.edu/reports/2015/pdfs/ashwin_final_paper.pdf

Siami-Namini, S., & Namin, A. S. (2018). Forecasting Economics and Financial Time Series: ARIMA vs. LSTM. https://arxiv.org/abs/1803.06386v1

Taylor, D. (2015). The pharmaceutical industry and the future of drug development. Pharmaceutical in the Environment. https://pubs.rsc.org/en/content/chapterhtml/2015/9781782622345-00001?isbn=978-1-

Tiriveedhi, V. (2018). Impact of Precision Medicine on Drug Repositioning and Pricing: A Too Small to Thrive Crisis. Journal of Personalized Medicine 2018, Vol. 8, Page 36, 8(4), 36. https://doi.org/10.3390/JPM8040036

Wang, C. H., & Yun, Y. (2020). Demand planning and sales forecasting for motherboard manufacturers considering dynamic interactions of computer products. Computers & Industrial Engineering, 149, 106788. https://doi.org/10.1016/J.CIE.2020.106788

Wang, Z., & Lou, Y. (2019). Hydrological time series forecast model based on wavelet de-noising and ARIMA-LSTM. Proceedings of 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2019, 1697–1701. https://doi.org/10.1109/ITNEC.2019.8729441

Yeasmin, N., Amin, S. H., & Tosarkani, B. M. (2022). Machine Learning Techniques for Grocery Sales Forecasting by Analyzing Historical Data. 21–36. https://doi.org/10.1007/978-3-030-85383-9_2

Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. Applied Artificial Intelligence, 17(5–6), 375–381. https://doi.org/10.1080/713827180

DOI: 10.6977/IJoSI.202409_8(3).0006
N. Qassrawi, M. Azzeh /Int. J. Systematic Innovation, 8(3), 63-83 (2024)

## AUTHOR BIOGRAPHIES

**Noura Qassrawi** is Master student pursuing MSc in computer Science at Applied Science Private University. His research interests include Machine learning and data mining.

**Mohammad Azzeh** is a full professor at the Department of Data Science at Princess Sumaya University for Technology. He earned his PhD in Computing from University of Bradford in 2010, Bradford, UK. M.S.C in Software Engineering from University of the West of England, Bristol, UK. Dr. Mohammad has published over 50 research articles in reputable journals and conferences such as IET Software, Software: Evolution & Process, Empirical Software Engineering and Systems & Software. His research interests focus on Software Cost Estimation, Empirical Software Engineering, Data Science, Mining Software Repositories, Machine Learning for Software Engineering Problems. Dr. Mohammad was Conference chair of CSIT2016 and CSIT2018, and he is co-chair of many IT-related workshops.

**Mohammad Hijjawi** is an associate professor at the Department of Computer Science in the Faculty of Information Technology at Applied Science University. He has published many research articles in reputable journals and conferences. His research interests focus on Natural Language Processing and Health informatics.

DOI: 10.6977/IJoSI.202409_8(3).0007
R. S. Bindu, S. Shalgar, etc/Int. J. Systematic Innovation, 8(3), 84-98 (2024)

# Design, development & performance evaluation of sustainable, hybrid air-conditioning system for automobiles

Rupa Sunil Bindu*[1], Sandeep Shalgar[2], Avinash Salunke*[1], Ankur Salunkhe*[1]
[1] D Y Patil Institute of Technology, Pimpri, Pune
[2] Tata Technologies Limited, Pune
[1] roopabindoo@gmail.com
(Received 05 December 2023; Final version received 08 May 2024; Accepted 21 June 2024)

## Abstract

Performance of the vapor adsorption refrigeration system is based on controlling the pressure of refrigerant, multi fluid temperatures, ambient and cabin air temperatures, and humidity to maintain thermal equilibrium. A direct evaporative cooling (DEC) system combined with adsorption refrigeration technique is sustainable as it is driven by vehicle exhaust. The heat potential in the exhaust gases of a vehicle which otherwise is going to be wasted to the atmosphere, can be (HAC). The designed and fabricated HAC model is fitted in a small vehicle and vehicle exhaust is provided to this model for finding its COP and cooling potential with the help of a customized data acquisition and controlling system. This research describes the design development and testing of an innovative HAC system for small cabin volume cars. The theoretical and actual COP of HAC system is derived at maximum blower speed and two DEC fans working conditions. The theoretical COP of HAC system varies from 0.57 to 0.66 and the actual COP varies from 0.57 to 0.47 with vehicle speed. As velocity increases COP decreases. Increased COP of the HAC system promises its easy adaptability for automotives at reduced cost.

*Keywords: sustainable cooling technologies, direct evaporative cooling, specific cooling power, cellulose grid, vehicle compartment's air cooling, vapor adsorption refrigeration*, activated carbon granules.

## 1. Introduction

Since the previous four decades, human pain has been growing due to significant changes in the environment and atmospheric influences. As a result, the elective aspect of incorporating air conditioning in automobiles is today becoming a significant system of a vehicle and has become a need. These higher cooling requirements necessitate more air conditioning power and the consumption of additional hydrocarbon fuel. Vapor compression refrigeration is extensively employed in current vehicles propelled by internal combustion engines, with highly optimized and efficient components. To control engine power needs due to air conditioning system power, extra fuel is burned, increasing engine speed by 200 to 300 rpm (Johnson 2002). Hence, an engine exhibiting a thermal efficiency ranging from 25% to 30% incurs an extra 10% fuel consumption in both operational and non-operational states, increasing more combustion gases like carbon dioxide (CO2), carbon monoxide (CO), sulfur dioxide (SO2), nitrogen dioxide (NOx) emissions (Atan, 1998).

Regulations limiting the use of chlorofluorocarbon (CFC) and hydrofluorocarbons (HFC) to reduce ozone depletion and greenhouse gas emissions must also be severely enforced. As shown in **Figure 1**. below, a belt-driven mechanical nature of the compressor is utilized, which utilizes a designated portion of the power supplied by the engine.

Over the previous decade, efficiency enhancement and thermal load optimization have evolved, resulting in smaller total component sizes and more system flexibility in automobiles. These advancements facilitated the design of the complex space within the vehicle. The control of the compressor and crank pulley engagement is governed by an electronic control unit, which maps the temperature set requirement for the cabin's environment.

## 2. Literature Review

### 2.1 Development of adsorption air conditioning system

Engine operating at 30% thermal efficiency releases 70% of the waste heat through exhaust gases, coolant, and engine compartments by radiation (Johnson 2002). Atan et al. devoted their efforts to employing the

waste energy in radiator water and exhaust gases for air conditioning purposes (Atan 1998). Johnson et al. illustrated two varieties of refrigeration systems: vapor absorption and other is solid adsorption, explaining the adsorption of refrigerant by porous media. Zeolite, gel of silica, activated carbon, and composites have porous structures that absorb or desorb the refrigerant gases on cooling or heating. The solid material is called an adsorbent while refrigerant gases are called adsorbate. The effectiveness of the adsorption system majorly depends on adsorbate uptake at working conditions (Johnson 2002).

## 2.2 Adsorbate and adsorbent working pairsItre

The different working pairs of adsorbate-adsorbent have their unique adsorption equilibrium characteristics. Many researchers have found mathematical correlations to express various pair parameters for the formation of steady state equilibrium working conditions. **Figure 2**. shows the development of different adsorbent-adsorbate pairs till today. L. Wang et al. summarized the benefits and drawbacks of various working pairs. along with mathematical models (L. Wang, R. Wang, and Oliveira 2009). Many researchers put their efforts into adopting a variety of cooling techniques for different cooling applications nevertheless cooling performance improvement is the key focus area.

Extensive review work on air-conditioning of commercial vehicles focused on absorption refrigeration using engine exhaust and fuel cell exhaust. The yearly refrigerant leakage rate from presently used vapor compression system can be as high as 25% moreover, these refrigerants exhibit noteworthy Global Warming Potential (GWP). The review also found very scare contribution of absorption refrigeration to automobile air conditioning along with vehicle level experimentation (Venkataraman et al. 2020). Tiawri et al. developed the exhaust gas operated refrigeration system for trucks with 1kW capacity. The experimentation is done at a laboratory fitted engine with an Activated Carbon-ammonia working pair. It was possible to get a cooling power of 1 to 1.2 kW. The system's coefficient of performance (COP) falls within the spectrum of 0.40 to 0.45. the cooling effect requires around 10 minutes of heating time. The overall weight for a cooling capacity of 1kW is 3000g of Activated Carbon-ammonia (Tiwari and Parishwad 2012). Li Gang et al. experimentally analyzed the performance of the secondary loop refrigerant system for automobiles using R152a and HC-290. The experimental outputs are compared with those of single-loop R134a traditional system. The results have revealed an improvement in cooling performance by 15% with R152a while 60% with HC-290.The systematic uncertainty for the working range of instruments are also presented. The lower pressure of R134a is 3 bar while higher side pressure is about 14 bar for an ambient heat load of 35°C (G. Li et al. 2014).

## 2.3 Activated carbon (AC) as adsorbent

Activated carbon has graphite lattice that is very porous and amorphous. It's often made up of little pellets or powder or granules. Carbonization and activation or oxidation are the two steps in the manufacturing process. The carbonization process entails drying and heating in the range of 600-900°C in a nitrogen environment in order to remove by-products, such as tars and other hydrocarbons, from the source material, as well as expel any generated gases. In the activation or oxidation, the carbonized product is exposed to oxygen above 300°C. Further to improve porosity by opening its microstructure, chemical treatment is done with different hydroxides and chlorides.

Yeo et al., reviewed the adsorption refrigeration system using improved porosity of the adsorbent. The utilization of readily accessible commercial activated carbon without antecedent treatment has led to a diminished performance outcome. Oxidized activated carbon has a better adsorption property on hexavalent chromium (VI) as compared to the untreated activated

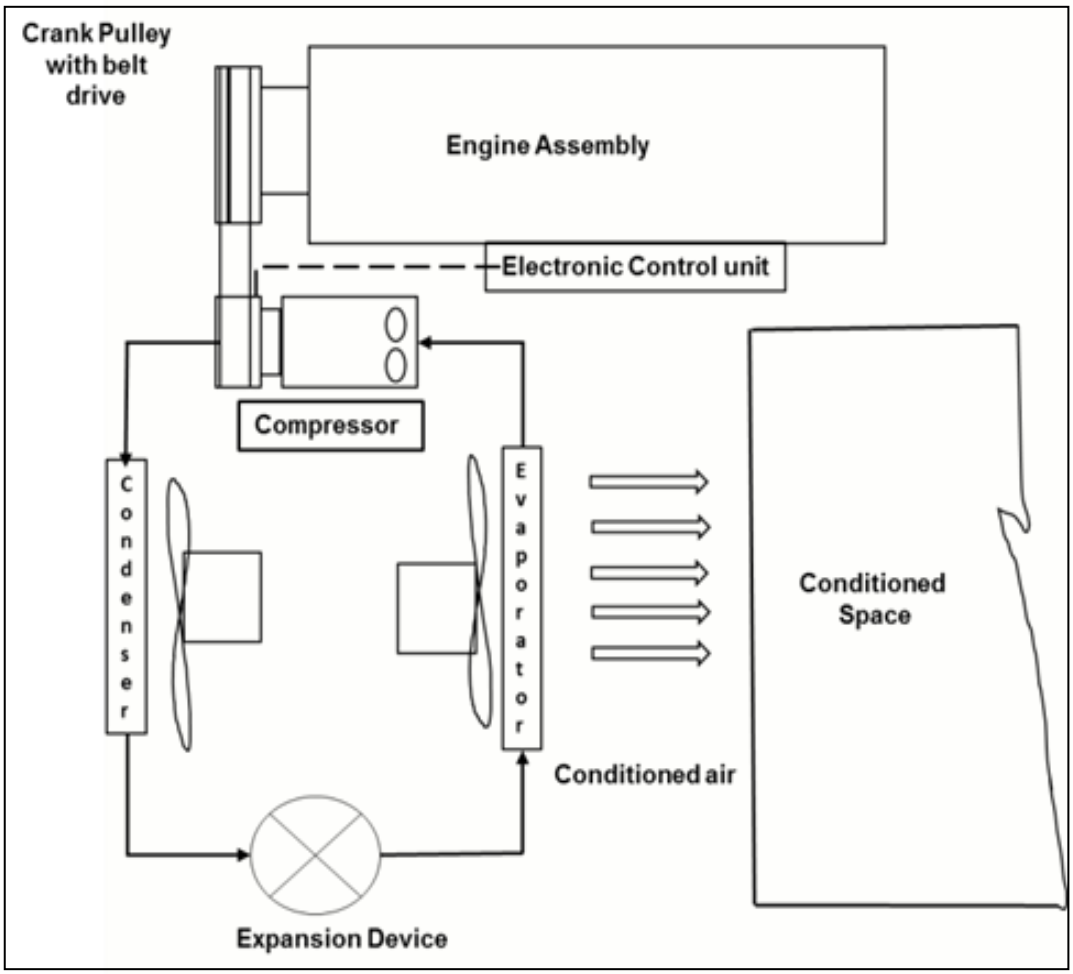



**Fig 1**. Conventional compression refrigeration

carbon. Oxidative treatment of activated carbon was highly effective in increasing metal ion uptakes. Generally, the thermal treatment of activated carbon proved effective in enhancing the adsorption capacity of organic components (Yeo, Tan, and Abdullah 2012). Ruzhu Wang et al. explained the microstructure of different types of activated carbon. Other adsorbents and adsorbates are also explained with their properties. Different working pairs with their performance parameters and thermodynamic properties of physical adsorption are also explained. Fig 3. (a) and Fig 3 (b) below show the pictorial and microscopic view of granular activated carbon (GAC) made of purified coconut shells (Ruzhu Wang, Liwei, and Jingy 2014).

Zeng et al., Investigated the impact of particle diameter on thermal conductivity. The thermal conductivity of the adsorbent is also affected by the refrigerant's thermal conduction coefficient and the number of voids present in the adsorbent layer. When the pressure drop in the adsorber is reduced. The mass transfer performance improves. As permeability increases, pressure drop reduces. The permeability experiment was carried out on adsorbent with diameters of 305µm, 390µm, 513µm, and 605µm, and it was discovered that the permeability increases as the adsorbent diameter and void ratio increase (Zeng et al. 2017).

The researcher designed a refrigeration system based on adsorption utilizing granular activated carbon as the adsorbent and refrigerant R134a as the adsorbate. The system uses solar thermal energy or waste heat in exhaust gases. The maximum temperature used for exhaust gases is 100°C. The calculated maximum value of the coefficient of performance is 0.36. The maximum specific cooling energy (SCE) of the system obtained from the test is 71 kJ/kg compared with the theoretical highest SCE of 82 kJ/kg (Askalany et al. 2013). The author analyzed the physical properties of granular activated carbon and R-134a refrigerant pair at temperatures from 21 ℃ to 61℃ and pressures up to 11 kgf /cm2. After 20 minutes of operation at 21°C, the highest rate of

adsorption was found equal to 1.90kg of R134a per kg of granular activated carbon. These outcomes as well

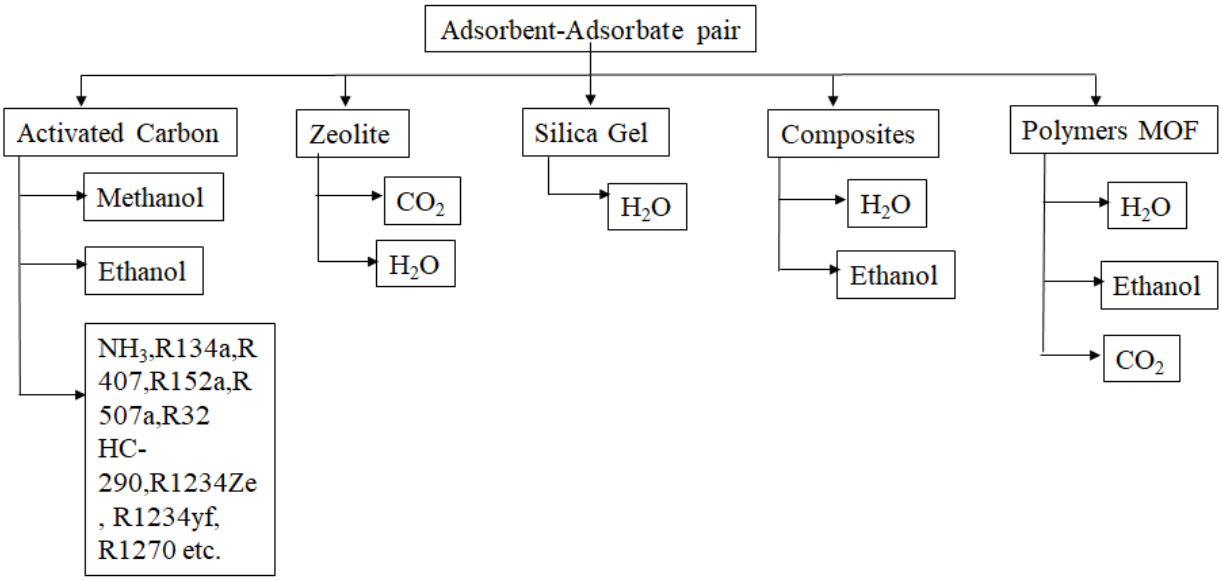



**Fig 2**. Development of different adsorbent-adsorbate pairs (Shabir et al. 2020).

depict "The adsorption capacity per kilogram of adsorber rises with an increase in heat transfer area.". This has given rise to a tube heat exchanger with fins which can be used as an adsorber. For capacities ranging from 0.21 to 1.82 kg of adsorbate/kg of adsorbent, The isosteric heat of adsorption spans from 120 to 340 kJ/kg. The Author analyzed the refrigeration system with R134a. Isotherms of the adsorption process are obtained for Maxsorb III in the temperature spans of 6°C – 70 °C and pressures reaching 12 bars during desorption. In current testing data, the isosteric adsorption heat is predicted for the working pair (Askalany et al. 2013). Shmroukh et al., designed adsorption chillers of 5 kW capacity. A tube and fin exchanger were used at the adsorber center. The surface was sealed with adhesive material. Powder activated carbon/R-134a, Powder activated carbon /R-407c, Powder activated carbon /R-507A, Granular activated carbon/R-507A, Granular activated carbon /R-407c, and Granular activated carbon /R-134a were all tested at varied adsorption temperatures ranging from 26°C to 52°C. At 25°C, PAC/R-134a had a highest adsorption capacity of 0.8352 kg/kg, whereas at 50°C, it had a peak adsorption capacity of 0.3207 kg/kg. The Powder activated carbon /R-134a pair is strongly suggested as a working pair due to its extraordinary adsorption capability in the tested working pairs (Shmroukh et al., 2015). Habib et al., depicted the kinematic characteristics of R134a and R507A for adsorption on pitch type Activated Carbon for various temperatures ranging from 21°C to 61°C using a variable pressure at constant volume. The adsorbent can adsorb the refrigerant R134a 1.6 kg/kg in 1200s at 25°C. As the ad-sorption temperature rises to 60°C, equilibrium absorption decreases to 1.0 kg/kg.

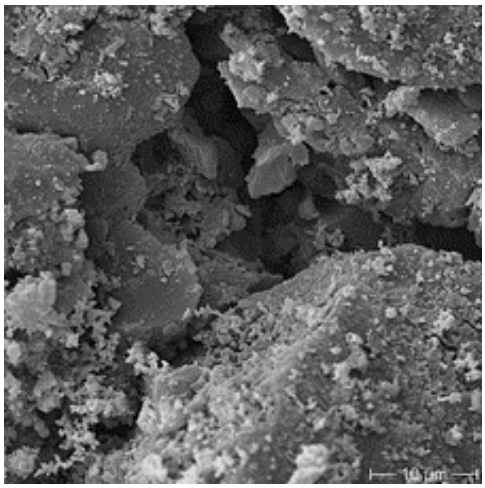
The equilibrium position, on the other hand, takes only 600 seconds to attain (Habib et al., 2010). The effectiveness of an adsorption chiller which has our adsorbent beds using Maxsorb III activated carbon was examined by Jribi et al., with 81 kg of Maxsorb and produced from solar energy at 85°C or waste heat. The system is able to generate 2 kW cooling power. The adsorption cooling performance of R1234ze is found to be virtually equal to that of R134a. Both R134a and R1234ze refrigerants have well-formulated mathematical models for adsorption cooling systems (Jribi et al., 2013).

Astina et al. in his experimentation used AC-R32 working pair. The adsorption capacity is increased using nitric acid and sulfuric acid and compared with pure activated carbon. The obtained COPs are in the range of 0.14 to 0.22 and SCP ranged between 4.0 and 6.3 W/kg of Activated Carbon. The OMB-DAQ-2461

data logger is used to map the temperature after 5 seconds interval and pressure is measured manually. The Clausius-Clapeyron equation stating the linkage between the evaporator pressure, the condenser pressure,

and system temperatures is described. The evaporator cooling capacity is dependent on a specific cooling power, cycle time, and mass of activated carbon in the bed. Thus, an increase in the number of adsorber bed than one and an optimum mass-accommodated design with a higher heat transfer coefficient enhances the cooling capacity. The adsorber bed leakage testing and evacuation is also mentioned before charging refrigerants into the adsorption system (Astina et al., 2018).

(a)



(b)

**Fig 3**. (a) Pictorial view of GAC; (b) Microscopic view of GAC

## 2.4 Activated Carbon-R134a working pair

Ahmed et al., explained the physical, chemical and composite adsorption future working pairs. In his work a review of many new and current work is covered: out of that AC/R134a pair is found much effective implementable working pair for automobiles applications. An adsorption isotherm at a pressure of 0.8 MPa, 30°C temperature, R134a -AC had a maximum capacity of adsorption uptake as 2 g/g. The adsorption time at 25°C was calculated to be 1200 seconds (Ahmed and Shehata 2017). Ahmed N. Shmroukh et al. reviewed the current advancements in the utilization of adsorption working pairs and revealed that the Maxsorb III/R134 combination stands out with the peak adsorption capacity. Activated carbon /Methanol has max. the adsorption capacity of 0.259kg/kg and Activated carbon/R134a has a maximum adsorption capacity of 2kg/kg at 30°C of bed temperature (Shmroukh et al., 2015).

Ojha et al., in their review, described different working pairs and recent progress in adsorption refrigeration along with different working pairs and applications. In his review, the maximum operating temperature is 250°C. It is observed that the use of activated charcoal is very frequent due to its highly porous volume. This sets activated carbon apart from typical adsorbents. In adsorption refrigeration systems, to maintain the continuity of the cooling effect produced, a minimum of two adsorber beds is a requisite (Ojha et al., 2021).

The system parameters of the GAC/R134a pair were reviewed by Askalany et al., The studies were carried out using a laboratory-scale test rig that was conceived and manufactured. At various temperatures, the GAC's adsorption ability was investigated. During the experiments, pressure and time were monitored. After 1000 seconds, The maximum capacity measured was to be 1.68 kg of refrigerant /kg of adsorbent at a temperature of 25°C (Askalany et al., 2013). Shmroukh et al., tested experimentally the adsorption properties of the granular activated carbon/R-134a pair in the temperatures 20°C to 60°C and at maximum pressures of 10 bars. At the end of 1200 seconds at 20°C, the highest adsorping capacity was obtained equal to 1.92 kg/kg. The test results show that as the thermal exchange surface area escalates the adsorping capacity /kg of adsorber also increases, based on which an adsorption heat exchanger with fins is developed. For adsorbing powers between 0.2 to 1.8 kg/kg, the isosteric heat of adsorption was estimated to be 120 to 340 kJ/kg (Shmroukh et al., 2015). S.A.Wani et al. reviewed experimental evaluation of adsorption refrigeration unit driven by exhaust gasses heat of automobiles. With a source temperature of 350°C to 450°C, the cooling power achieved was 5KW, COP was 0.25 and SCP of 165 W/kg to 202 W/kg of activated carbon (S.A.Wani et al., 2015).

Shabir et al., in his review work mentioned isotherm model for different adsorbate and adsorbent working pairs. For activated carbon/R134a working pair Dubinine Astakhov (D-A) model is used which gives

fundamental equations for finding adsorption uptake isotherm at different pressures. Granular Activated carbon (GAC) with different heat treatment and oxidation processes improves the refrigerant uptake capacity. The granular activated carbon has a total pore volume of 0.878 $cm^3/g$ and a surface area of 899 $m^2/g$. At 32°C saturation conditions, the equilibrium adsorption uptake is 1.639 kg/kg. It is evident that extensive research is underway to pioneer energy-efficient adsorption systems as a substitute for traditional air conditioning technology (Shabir et al., 2020).

Banker et al., in their performance analysis, used the AC-R134a pair along with thermal compression addition to existing conventional air conditioning systems. It showed an energy saving of 40% in hybrid compression over conventional compressors. Such a type of system is suitable for large cooling capacity requirements and where heat source at low temp. is available in abundant form like in industrial applications (Banker et al., 2008).

Kılıç et al., in his work, uses the 4 adsorber beds filled with activated carbon and R134a working pair. D-A equation model is used and parameters for creating the isotherm model are given. The mathematical equations for adsorber bed's parameters like length, diameter and inner diameter, coolant pipe diameters, mass of adsorbent, and adsorbate based on driving temperature are presented. The Saturation pressure at evaporator is 4.23 bar while and condenser is 8.68 bar. For evaporator, the temperature achieved is 10.6°C, and the cycle time considered is 960 seconds. The experimentation compared activated carbon granules-R134a or R404a paired with activated carbon pallets and isotherms are plotted at different pressures and temperatures with governing equation models for these pairs. The process of plotting the adsorption isotherm is mentioned with the procedure of vacuuming at 0.05kPa and desorbing residual gases. To prevent condensation within the adsorption tank during refrigerant charging, measures were taken to ensure that the pressure during the filling process was retained below the saturation pressure of the adsorbate fluid. Computed adsorption capacity and isosteric heat are presented compared with experimental data (Kılıç and Gönül 2018).

Pinto et al., reviewed details of Activated Carbon/R134a pair with four beds system had compared simulation and experimental results. Heated water at 84.9°C while cold water at 29.9°C is flown in the circuit. The ID of bed is 36.4 mm, 4.499 kg of R134a is filled. Optimized cycle time is found as 40 mins (10 mins to each process). At no load on evaporator condition, the lowest cooling of 14.5°C with a refrigeration capacity of 430±13W with COP 0.5 (Pinto et al. 2019). Dakkama et al. explore the use of Simulink/MATLAB software for testing the performance of several adsorbent/refrigerant working pairs in cascaded adsorption refrigeration. Activated carbon with an alkali activation process (Maxsorb) is used with R134a, R152a, Propane and Ethanol. Amongst all refrigerants R134a and propane paired with Maxsorb found the highest cooling capacity at 350s of cycle time with a COP of 0.088 (Dakkama et al., 2015).

X.H. Li et al. studied and reviewed the various methods for maximizing the Performance of an adsorption refrigeration system which depends basically on the size and shape of the adsorber bed. One approach involves reducing resistance to the heat transfer of the bed, achievable by increasing the heat transfer area. Options to achieve this include utilizing a plate-finned bed, a spiral plate bed, or a pin-fin bed, all of which effectively enhance the transfer area. The cycling time is reduced and the COP is improved when the heat pipe technology is introduced into the bed. It has been discovered that lowering the thermal contact resistance between the wall and the adsorbent and also reducing the heat resistance of the adsorbent itself can improve heat transfer performance. By adding new advances in the heat transfer process, adsorption cycles can be reduced (X. H. Li et al. 2015). Sharafian et al. studied different adsorber bed designs. Tubular adsorber beds with fins and extensions are developed as implementing practical solutions to augment heat and mass transfer rates within the bed for the required SCP of 350W/kg of dry adsorbent after optimizing fin parameters like fin pitch and fin length and improved adsorbent material's thermal conductivity. The metal wool is added to activated carbon to improve thermal conductivity for enhanced heat transfer rate (Sharafian and Bahrami 2014).

## 3. Procedure for design of Hybrid Air Conditioning System (HAC)

Hybrid air conditioning systems incorporate various cooling methods or cooling machines which have an energy-saving potential. Kojok et al. extensively reviewed a variety of hybrid systems. The evaporative cooling when compared with vapor compression cycle showed an increase in saving of energy from 20% to 49%. The use of cooling pads for liquid cooled vapor compression condenser improves system performance. The reduction in temperature of condenser even by 1°C leads to 2-4% COP improvement. Hybrid systems

utilizing adsorption cooling offer advantages primarily in scenarios with modest air conditioning loads and where discontinuous cooling is acceptable. This is particularly applicable in applications such as cold storage, air conditioning for vehicles, and similar use cases (Kojok et al., 2016). Kilic et al., in his experimentation, combine vapor adsorption and vapor compression refrigeration cycle. The refrigeration system uses silica gel- water pair. The flow of the mathematical model is given. Adsorption and desorption time is considered as 500s while cooling and heating time is 50s. The vapor compression cycle tested with different refrigerants found with highest COP and energy-saving ratio of 8.8 (Kilic and Anjrini 2020). Sultan et al., focus on desiccant dehumidification for air conditioning applications. Desiccant refrigeration combines desiccant dehumidification and evaporative cooling to manage humidity and temperature. The Desiccant air-conditioning (DAC) is appealing since it is devoid of harmful refrigerants, and regeneration is feasible using lower-grade heat. Silica gel, synthetic zeolite, activated alumina, polymer desiccant are the materials used for DAC system whose COP values found from 0.35 to 0.44. The desiccant efficiently bears the latent load of AC, lowering energy consumption and increasing system COP. The higher COP requires higher regenerated temperatures and humidity, so the system is mostly useful in highly humid and hot ambient regions (Sultan et al. 2015).

## 3.1 Heat load in modelled vehicle

The necessary cooling load for cooling of the vehicle compartments is determined through the conventional way of calculating various heat load scenarios. This includes considerations for solar radiation on vehicle rooftops, vehicle walls, the number of individuals within the cabin and windows, air heat load, utility equipment loads, as well as sensible and latent heat loads. For load assessment, a model vehicle with cabin dimensions of (2.23 x 1 x 1.4) m3 is used. According to environmental statistics from West India's local region, the maximum atmospheric temperature is 40°C (Verde et al., 2016). The DEC system is developed with a critical condition of 40°C and 40% RH. Considering the fluctuating load throughout the year, a specified comfort target temperature which is 25°C desired.

The cumulative heat load required for space cooling in the model car is assessed to be 1.7 kW, hence a refrigeration system of 0.5 TR is considered adequate to keep the vehicle compartment temperature at 25°C (Waghchore, Jumde and Somwanshi . 2013).

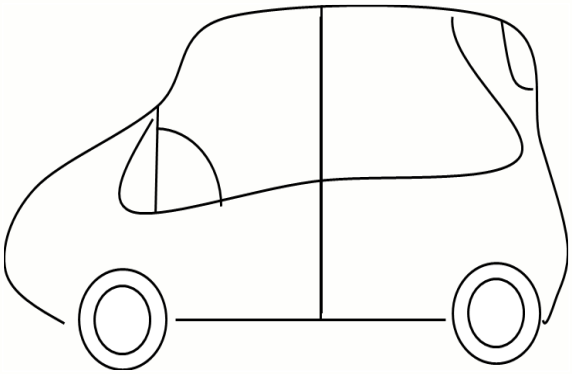


**Fig 4**. Outline of model vehicle for evaluation

## 3.2 Working of hybrid air conditioning system

**Figure 5** above illustrates a schematic representation of the recently devised hybrid air conditioning system (HAC). using adsorption refrigeration combined with evaporative cooling.

The Direct Evaporative Cooling system (DEC) makes use of pads of cellulose material and takes latent enthalpy of vaporization of moisture, from air, lowering the temperature of the air. This cooled air from DEC system is then passed over the surface of finned evaporator tubes which is also a part adsorption refrigeration circuit. So, the air is further cooled over the evaporator tubes and then passed to the passenger compartment space for cooling. Thus HAC System consists of different circuits viz Flow of refrigerant through a circuit, Flow of exhaust gas circuit, and Flow of coolant circuit. As shown in Fig 5 above, the refrigerant circuit consists of two adsorber beds (B1 and B2) which work alternately in adsorption and desorption modes. In bed B1 when the adsorption of refrigerant is going on, bed B2 is involved in the desorption of the refrigerant process. The heat required for pressurization and desorption is taken from the waste heat of exhaust gases. The cooling of adsorber beds is done by passing coolant from the coolant circuit. Bed B1 is connected to the condenser through valve V3 through high-pressure line of refrigerant. Condenser outlet is connected by high-pressure line to the throttling valve and evaporator. The evaporator outlet is connected to bed B1 through valve V5 by low-pressure lines. Similarly, adsorber bed B2 is connected to the condenser through valve V4 by a high-pressure line. Condenser outlet is connected by a high-pressure line to the throttling valve and evaporator. Outlet of evaporator is connected to bed B2 through valve V6 by low-pressure lines.

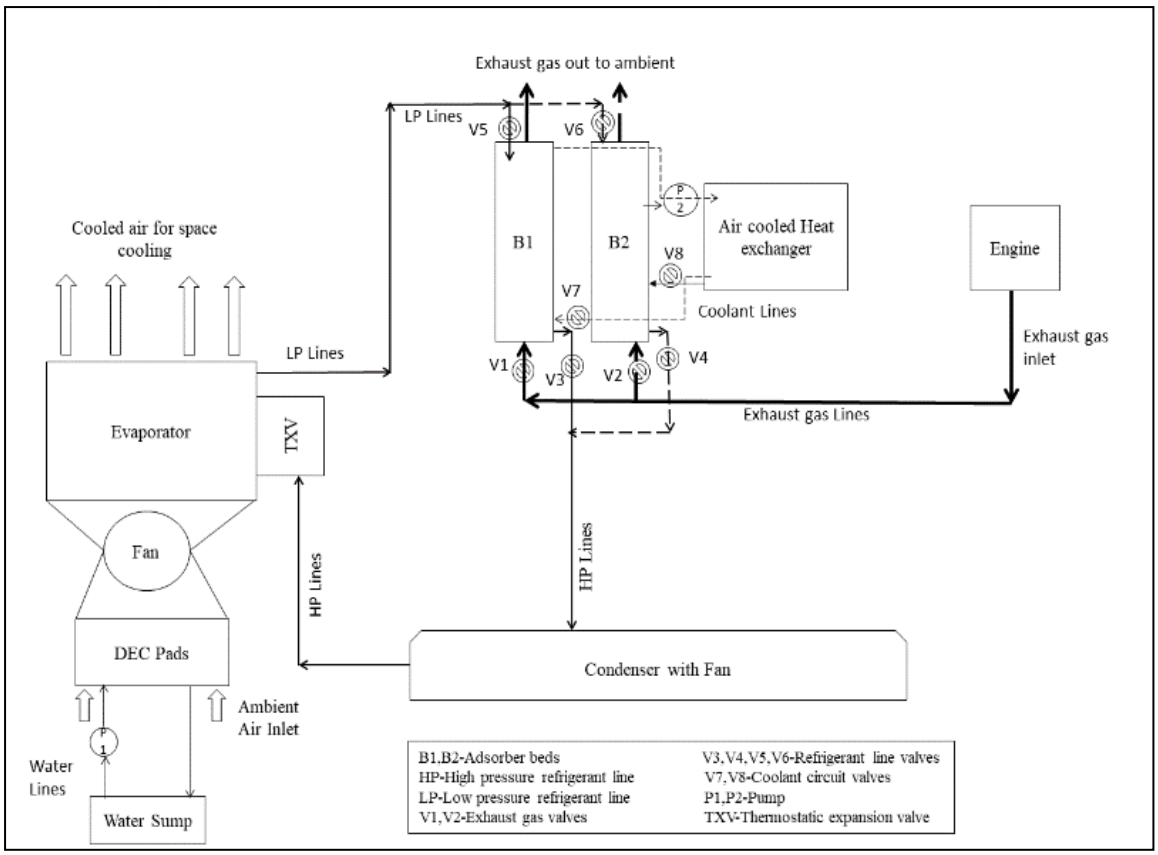



**Fig 5**. Hybrid Air Conditioning System

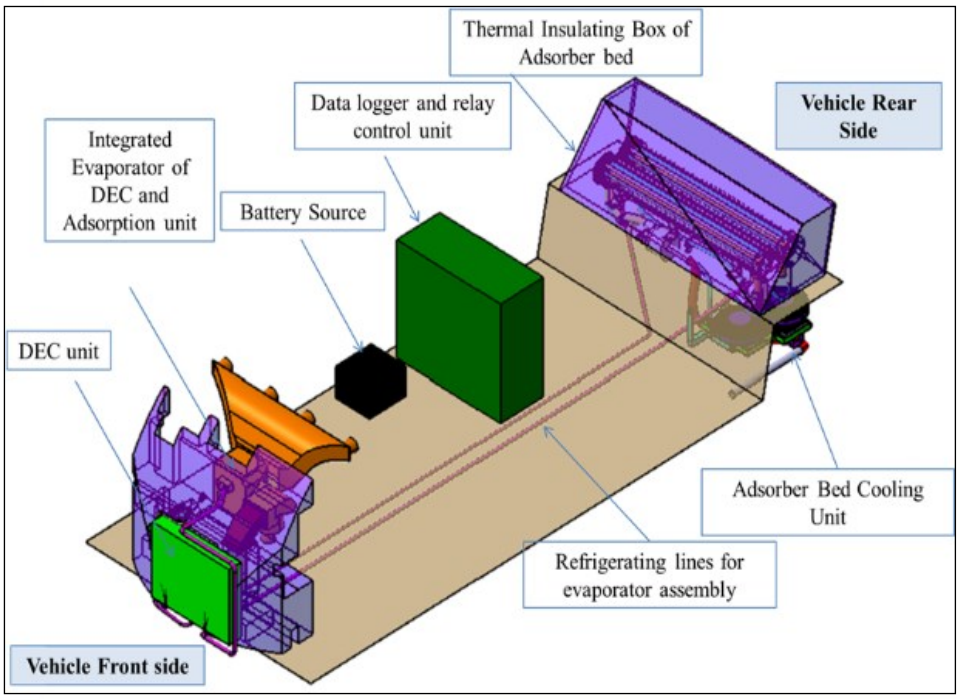
The exhaust gases from the engine are supplied alternately to bed B1 and B2 through valves V1 and V2 respectively. A water-glycol (50:50) mixture is used as a coolant to cool the adsorbed bed. The coolant absorbs heat from bed B1 and B2 and rejects it to air-cooled heat exchangers through alternate operation of valves V7 and V8 respectively. Pump P1 is used in DEC system to take water from the sump, pressurize it and spray over to DEC pads. Pump P2 sucks the heated coolant from bed B1 and B2 and forces it through air cooled heat exchangers.

## 3.3 Experimental setup in modelled vehicle

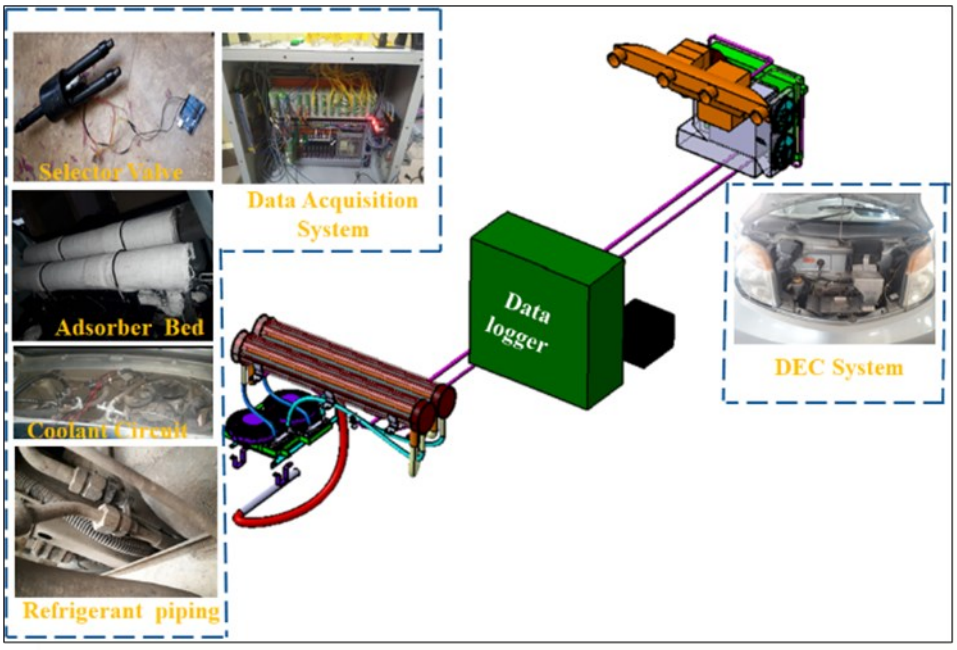
Refer to **Figure 6** (a) below which shows the proposed and developed HAC system integrated in selected passenger vehicles. The components are installed as proposed layout. DEC system is at the vehicle front side which transfers cooled air to cabin through a common evaporator assembly of the Vapor Adsorption System (VAS). The refrigerating line for evaporator cooling in VAS is routed form bottom of cabin floor from which refrigerant R134a is charged by vacuuming the circuit. Adsorber beds of VAS are installed at the vehicle rear side and are placed inside an insulated laminated box to avoid heat ingress to the cabin. The Data Acquistion System (DAS) is installed at the rear side of the vehicle and the battery at the center.

## 3.3 Vacuuming and determination of capacity of adsorption uptake

The capacity of adsorption uptake of an adsorbing material is determined by plotting an isotherm for the selected working pair. An isotherm is a graph that illustrates the correlation among pressure and adsorption uptake capacity for an activated carbon-R134a pair at a constant temperature. The concentration ratio for Vapor Adsorption Refrigeration System (VAS) is expressed as the ratio of adsorbed refrigerant mass to the activated carbon mass. The higher the concentration ratio, the higher will be the cooling effect. The pressure rise during the adsorption process is dependent on the adsorber bed design geometry hence this graph must be analyzed before experimentation of the HAC system. Before charging of gas, vacuuming is an important step where any minor leakage from any joints or component gets caught which may help in achieving stable performance system.



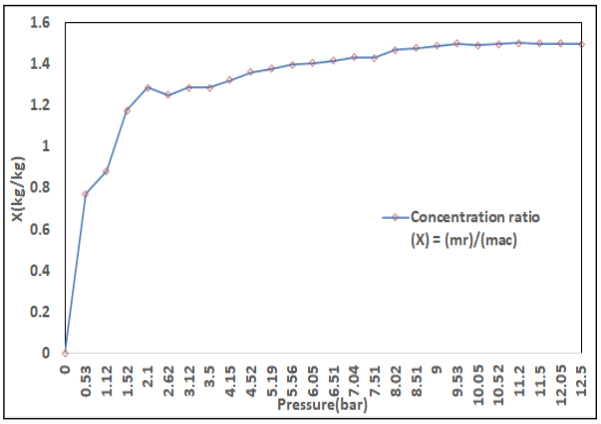

(a)



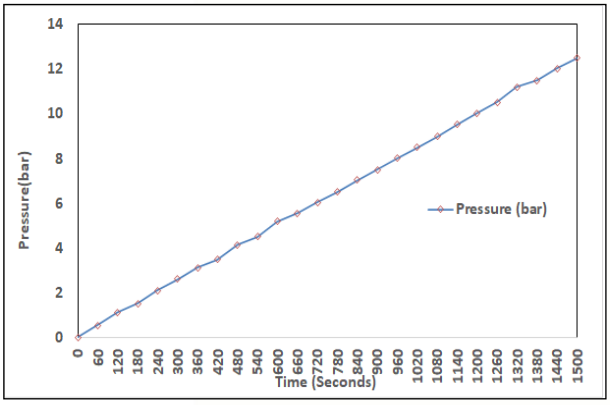

(b)

**Fig 6** (a) Proposed integration of HAC system inside vehicle; (b) Proposed and developed HAC system components.

The vacuum pump is employed to establish a vacuum within the adsorber bed, removing all gases and moisture until reaching the vacuum pump's limit, approximately 0.2 bar (absolute) (Jribi et al., 2013). In this process, any minor leakage in the refrigerating circuit leads to an increase in the vacuuming time and is critically overlooked to avoid gas leakage. In the process of vacuuming, the dust particles from the adsorber bed will dilute the vacuum pump oil and require to be changed after the vacuuming process.




(a)




(b)

**Fig 7**. (a) Adsorption bed uptake during R134a charging; (b) Adsorption bed R134a pressure rises with time.

The purity of refrigerant is important in achieving better performance. The adsorption bed can be heated during the vacuuming operation by keeping the exhaust gas input port open, allowing trapped gases from activated carbon to be released and increased adsorption uptake to be attained. This creates pressure difference in vacuuming leading to refrigerant flow from the refrigerating cylinder to the refrigerant circuit. Value make VES50B model of automatic charging machine is used to charge the refrigerant. The charging machine possesses two ports inlet and outlet ports. The Inlet port is connected to the refrigerant gas cylinder and the outlet port is connected to the vacuumed refrigerant circuit at the inlet refrigerant port of the adsorber bed. The outlet refrigerant ports of adsorbed beds are closed manually by turning off the solenoid valves by giving an electrical power supply. An adsorber bed coolant circuit is kept working to cool the adsorber. The gas cylinder is placed on the weighing platform of the automatic charging machine and the initial weight of the gas cylinder is noted. Now the outlet port of the gas charging machine is slowly opened. After 60s of refrigerant flow hold and wait for equilibrium to be achieved. Take the reading of the weight of the cylinder. The reduction in the weight of the cylinder is equal to the mass of the refrigerant transferred to the adsorber bed refrigerant circuit. As the refrigerant enters into the refrigerating circuit and adsorber beds its pressure increases. This pressure is also measured. Accordingly, the charging time, mass of R134a adsorbed, and pressure is tabulated. Based on this concentration equilibrium is calculated and the graph is plotted for adsorption concentration and pressure built as mentioned in Fig 7 (a) below concentration equilibrium is achieved after 1080 seconds and the total adsorbed refrigerant is 1.49kg. We decided the adsorption time is 600 seconds, hence at 600 seconds refrigerant uptake was 1.38kg/kg of activated carbon.

The adsorption and desorption process is pressure-dependent in the system, Fig 7. (b) shows the graph of pressure attained by refrigerant during charging with respect to time. At 600 seconds the adsorber bed pressure was 5.19 bar during the adsorption process while during the desorption process at 1200 seconds, it was 10.05 bar. So, our system will operate in this two-pressure range. There will be changes in the enthalpy of the refrigerant at these pressures which will give the desired cooling effect in the evaporator. These are the condenser and evaporator working pressures, Pc = 10.1 bar and Pe = 5.2 bar.

## 4. Dynamic testing for evaluating the performance of HAC: procedure, results, and discussion

**Figure 8**. (a) below represents the integration of the adsorption bed connected with the selector valve, solenoid valves, and DAS with a laptop. It has been discussed about the integration of DECS in **Figure 6**. (a) of the above section. The required thermosensors are instrumented as per **Figure 8.** (a) for measurement of temperatures of the ambient air, cabin air, refrigerant temperatures at adsorber bed inlet and outlet, exhaust gas inlets and

**Fig 8**. (Instrumentation of HAC System in Automobiles

in refrigeration circuits are placed for pressure measurement in adsorber bed inlet and outlets. Humidity sensors are located inside and outside the cabin for humidity measurements.

The vehicle along with HAC system fitted inside was run for a long distance in scarce area on dates 30th and 31st January 2023 on National Highway for dynamic testing. The testing of the vehicle HAC system started when thermal equilibrium temperature was achieved. The atmospheric conditions changed from 29ºC DBT, 54% RH to 31.5 ºC, 58% RH during testing.

**Figure 8.** (b) shows the time needed for heating the bed to raise the refrigerant pressure to the required working pressure, keeping the volume constant. With the escalation of vehicle speed, the triggering pressure also rises and it is achieved after 781 seconds at 72 kmph speed. The engine exhaust waste heat of 2.52kW to 3.14 kW is utilized for HAC system.

**Figure 9**. (a) below illustrates the theoretical and actual COP of the HAC system at maximum blower speed with two DEC fans in operation. The theoretical COP ranges from 0.57 to 0.66, while the actual COP is recorded as 0.57 and 0.47, exhibiting a decrease with increasing vehicle speed. This decline is attributed to the uncontrolled variations in gear change ratios, ambient air temperature, fluctuations in relative humidity, and variations in engine exhaust gas temperatures, which deviate from the ideal conditions.

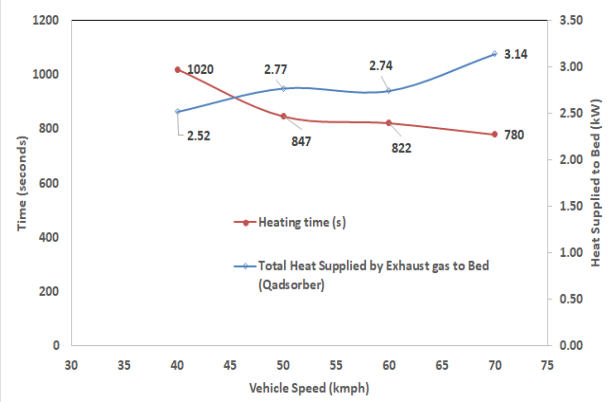
From experimentation, it is observed that the minimum heat utilized at 40 kmph vehicle speed is 2.52kW and it is 3.14kW at a vehicle speed of 70kmph. The percentage utilization at lower vehicle speed is more 76.21% and reduces to 37.16% as vehicle speed increases. This is due to the reduction in heating time to attain the required refrigerant pressures of triggering for achieving cooling effects when the vehicle is running at higher speeds.

## 5. Payback period of HAC system in passenger and commercial vehicles

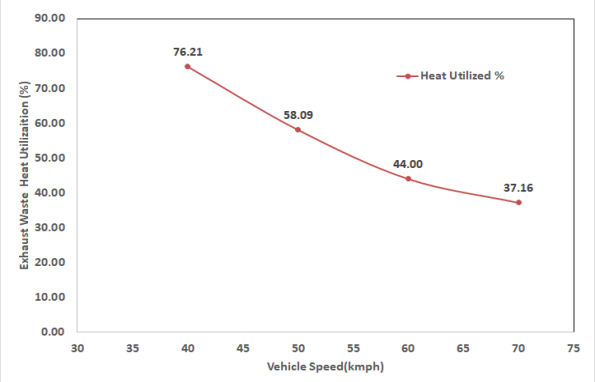
The implementation of any new technology products requires the overall system's cost consideration and its payback period. The payback period of HAC system

in passenger and commercial vehicles as compared to presently used Vapor Compression Refrigeration system is explained below based on following assumptions:

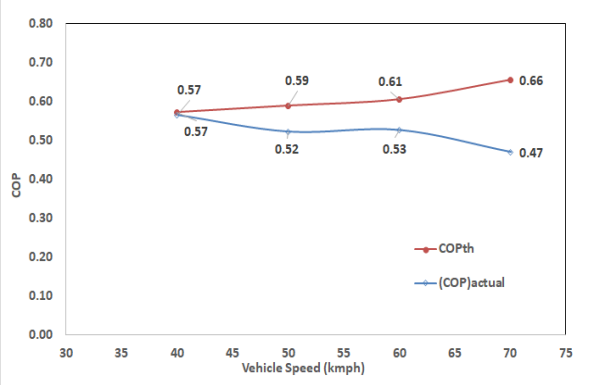
- Daily driving of passenger and commercial vehicle is considered as 250km, 30 days and 12 months of vehicle running.




(a)




(b)




(c)

**Fig 9**. (a) Heating time and heat supplied; (b) Theoretical and experimental COP of HAC; (c) Heat utilized by HAC system.

- Operating cost is 12 % of diesel consumption cost considered in vapour compression refrigeration system.
- Servicing cost is 10% of diesel consumption cost is considered for air filters change, cleaning of ducts, refrigerant charging, component failures if any for vapour compression refrigeration system and HAC system.

The AC operating cost of vapour compression refrigeration system in commercial trucks prominently affect the fuel consumptions over passenger vehicles due to very less fuel mileage. For passenger car's vapour compression refrigeration system the payback period worked out is 35.8 days and for commercial truck payback period is just 10days. But the air conditioning operating cost in commercial truck is 340% higher than that for passenger vehicles. Even though the initial cost of HAC system in passenger and commercial vehicles is 2.86 times higher than vapour compression refrigeration system but the yearly air conditioning operating cost for HAC system is 1.2 times lower than that of vapour compression refrigeration system.

Payback period for HAC system is 226 days for passenger vehicles and 87days for commercial vehicle which is significantly lower as compared to the human thermal comfort. The Indian scenarios of buying the commercial vehicles fitted with air conditioning system are significantly very less to 15% only due to higher AC operating cost. Hence the developed HAC system significantly plays very important role of highly comfort to the cost. Also, as per required cooling load of various Indian climates the HAC system provides flexibility to use either DEC system or HAC system whichever is beneficial.

## 6. Conclusion

DEC system making use of cellulose material pads was innovated for cooling the occupant's compartment of a small vehicle and was evaluated for its performance. The critical design parameters considered are the prevention of water leakage and noise. The electric power input to the blower can be reduced by 7.99% if the pads are just wetted by spraying water for only 60 seconds rather than keeping the pump on continuously. When the car is exposed to sun radiation in static conditions, the mean compartment air temperature rises by 17.5°C above atmospheric temperature, and inside air relative humidity is reduced by 27% at the 40$^{th}$ minute pertaining to the static test. The experimental outcomes became stable after 38$^{th}$ minutes from the beginning of the experiment. The temperatures at various locations in the system are measured after keeping the DEC system on for 10minutes time at three different fan speeds. At the lowest fan speed, the compartment front and rear temperatures got reduced from 41.2˚C to 35.9˚C and 36.3˚C

respectively. After this at the medium and high fan speeds, after 30th and 40th minutes from the start of the test, compartment air temperature is dropped by 6.3˚C and 9.5˚C respectively. The refrigerating capacity ascends with the augmentation of the fan speed. The highest refrigeration is obtained at the highest speed of the fan and is 1043W with ambient air temperature drop from 41.8˚C to 32.3˚C. The DEC system is effective for reducing air temperature by approximately 9.55˚C providing enhanced comfort for users. The coefficient of performance of DEC system obtained is 1.97 at blower speed 1 and 4.38 with blower speed 3. The maximum electric power consumed by electrical components is 238W. This increase in coefficient of performance of the system renders ease in adoptability at reduced cost and allows usage of smaller size automotive components. The theoretical and actual COP of HAC system was derived at maximum blower speed and two DEC fans working conditions. The theoretical COP of HAC system varies from 0.57 to 0.66 while actual COP is found 0.57 and 0.47 and COP decreases with a rise in vehicle speed. This is due to uncontrolled variation of gear change ratios, ambient air temperature, relative humidity change, and engine exhaust gas temperature variations that are not idealistic to those derived in numerical analysis. The HAC System model can be scaled up based on the space available in the compartment of the vehicle.

## • Uncertainty Analysis

The error of sensors, instruments, and calibration needs to be mapped at the primary stage of experimentation to ensure the acceptable variation range of measured output. Also, as required accuracy increases, the cost associated with the sensors and instrument also increases. Hence a careful selection of the instruments needs to be done as per the required accuracy only. Uncertainty can be computed using the equation:

$$\% \text{ uncertainty} = \frac{\text{Least count}}{\text{Minium value of output measured}}$$

Uncertainties in experiments can occur from selected instruments, their conditions and calibration, as well as the environment and test sampling. The temperature, pressure, humidity, mass flow, weight, and equipment current consumption are measured using the Data Acquisition System (DAS).

DAS comprises of 32 K-type thermocouples. These sensors have an operating range of about 0-600˚C having an accuracy of $\pm$1.5˚C. Temperature mapping is done using DVP-04TC modules. During the experiment, the lowest temperature observed was 22°C. Temperature uncertainty is derived as per the equation above is 6.81%. DAS consists of 2 humidity transducers of 947series and six Baumer make CTX pressure sensors. The pressure and humidity are mapped with DVP-04AD modules. The CPU used here is a Delta DVP-12SA2, which uses DVP16 SP I/O connections to handle acquired data. Uncertainties in airflow, weight, and current measurement are calculated using equation (a) and are mentioned in Table 1. given above which gives detailed insight into instruments, their make, operating range along with their experimental uncertainties

| Instrument Name | Operating Range | Accuracy | Measurement Minimum value | %Uncertainty |
|---|---|---|---|---|
| K-type Thermosensors | 0- 600°C | ± 1.5 ℃ | 22 ℃ | 6.81% |
| Humidity sensors | 0-100% RH | ± 3% RH | 32%RH | 3% |
| Pressure sensors | 1-40 bar | ± 0.5% bar | 0.53bar | 0.5% |
| Anemometer Mextech AM-4208 | 0-9999 ($m^3$ per minute) | ±(2%+1$m^3$ /min) | 0.833 $m^3$/min | 1.22% |
| Weight and charging machines VES 50B | 0.002 to 50 kg | ±0.05% | 5.383 | 0.05% |
| Clamp Meter RX266 | 20A-1000A | ± (2.0%) | 0.44A | 2% |

**Table 1**. Uncertainties in Experimentation

## Notes

- https://www.datasheetarchive.com/DVP-04AD-datasheet.html
- https://www.datasheetarchive.com/DVP-04TC-datasheet.html
- https://www.meanwell.com/Upload/PDF/LRS-100/LRS-100-SPEC.PDF
- https://www.datasheetarchive.com/DELTA+dvp+16sp-datasheet.html
- http://www.weatheronline.in/weather/maps/city

## References

Askalany, A. A., Saha, B. B., Ahmed, M. S., & Ismail, I. M. (2013). Adsorption cooling system employing granular activated carbon–R134a pair for renewable energy applications. International Journal of Refrigeration, 36(3), 1037–1044. https://doi.org/10.1016/j.ijrefrig.2012.11.009

Astina, I. M., Zidni, M. I., Hasugian, H. R., & Darmanto, P. S. (2018). Experiment of adsorption refrigeration system with working pairs of difluoromethane and activated carbon modified by sulfuric and nitric acids. 20015. https://doi.org/10.1063/1.5046599

Atan, R. (1998, February). Heat Recovery Equipment (Generator) in an Automobile for an Absorption Air-Conditioning System. https://doi.org/10.4271/980062

Banker, N. D., Dutta, P., Prasad, M., & Srinivasan, K. (2008). Performance studies on mechanical+adsorption hybrid compression refrigeration cycles with HFC 134a. International Journal of Refrigeration, 31(8), 1398–1406. https://doi.org/10.1016/j.ijrefrig.2008.03.009

Dakkama, H. J., Elsayed, A., Al-Dadah, R. K., Mahmoud, S. M., & Youssef, P. (2015). Investigation of Cascading Adsorption Refrigeration System with Integrated Evaporator-Condenser Heat Exchanger Using Different Working Pairs. Energy Procedia, 75, 1496–1501. https://doi.org/10.1016/j.egypro.2015.07.285

Habib, K., Saha, B. B., Rahman, K. A., Chakraborty, A., Koyama, S., & Ng, K. C. (2010). Experimental study on adsorption kinetics of activated carbon/R134a and activated carbon/R507A pairs. International Journal of Refrigeration, 33(4), 706–713. https://doi.org/10.1016/j.ijrefrig.2010.01.006

Johnson, V. H. (2002, June). Heat-Generated Cooling Opportunities in Vehicles. https://doi.org/10.4271/2002-01-1969

Jribi, S., Saha, B. B., Koyama, S., Chakraborty, A., & Ng, K. C. (2013). Study on activated carbon/HFO-1234ze(E) based adsorption cooling cycle. Applied Thermal Engineering, 50(2), 1570–1575. https://doi.org/10.1016/j.applthermaleng.2011.11.066

Kilic, M., & Anjrini, M. (2020). Comparative performance analysis of a combined cooling system with mechanical and adsorption cycles. Energy Conversion and Management, 221, 113208. https://doi.org/10.1016/j.enconman.2020.113208

Kılıç, M., & Gönül, E. (2018). An Experimental Study on Adsorption Characteristics of R134a and R404a Onto Granular and Pellet-Type Activated Carbon. In Exergetic, Energetic and Environmental Dimensions (pp. 751–761). Elsevier. https://doi.org/10.1016/B978-0-12-813734-5.00042-1

Kojok, F., Fardoun, F., Younes, R., & Outbib, R. (2016). Hybrid cooling systems: A review and an optimized selection scheme. Renewable and Sustainable Energy Reviews, 65, 57–80. https://doi.org/10.1016/j.rser.2016.06.092

Li, G., Eisele, M., Lee, H., Hwang, Y., & Radermacher, R. (2014). Experimental investigation of energy and exergy performance of secondary loop automotive air-conditioning systems using low-GWP (global warming potential) refrigerants. Energy, 68, 819–831. https://doi.org/10.1016/j.energy.2014.01.018

Li, X. H., Hou, X. H., Zhang, X., & Yuan, Z. X. (2015). A review on development of adsorption cooling - Novel beds and advanced cycles. Energy Conversion and Management, 94, 221–232. https://doi.org/10.1016/j.enconman.2015.01.076

Ojha, M. K., Shukla, A. K., Verma, P., & Kannojiya, R. (2021). Recent progress and outlook of solar adsorption refrigeration systems. Materials Today:

Proceedings, 46, 5639–5646. https://doi.org/10.1016/j.matpr.2020.09.593

Pinto, S. P., Karanam, P., Raghavendra, B. G., Boopathi, V., Mathad, P., Behera, U., & Kasthurirengan, S. (2019). Development and studies of low capacity adsorption refrigeration systems based on silica gel-water and activated carbon-R134a pairs. Heat and Mass Transfer, 55(2), 513–531. https://doi.org/10.1007/s00231-018-2441-0

Shabir, F., Sultan, M., Miyazaki, T., Saha, B. B., Askalany, A., Ali, I., Zhou, Y., Ahmad, R., & Shamshiri, R. R. (2020). Recent updates on the adsorption capacities of adsorbent-adsorbate pairs for heat transformation applications. Renewable and Sustainable Energy Reviews, 119, 109630. https://doi.org/10.1016/j.rser.2019.109630

Sharafian, A., & Bahrami, M. (2014). Assessment of adsorber bed designs in waste-heat driven adsorption cooling systems for vehicle air conditioning and refrigeration. Renewable and Sustainable Energy Reviews, 30, 440–451. https://doi.org/10.1016/j.rser.2013.10.031

Shmroukh, A. N., Hamza, A., Ali, H., Abel-Rahman, A. K., & Ookwara, S. (2015). Experimental Investigation on Adsorption Capacity of a Variety of Activated Carbon/Refrigerant Pairs. In Journal of Engineering Research and Applications www.ijera.com (Vol. 5, pp. 66–76). www.ijera.com

Sultan, M., El-Sharkawy, I. I., Miyazaki, T., Saha, B. B., & Koyama, S. (2015). An overview of solid desiccant dehumidification and air conditioning systems. Renewable and Sustainable Energy Reviews, 46, 16–29. https://doi.org/10.1016/j.rser.2015.02.038

Tiwari, H., & Parishwad, G. V. (2012). Adsorption Refrigeration System for Cabin Cooling of Trucks (Vol. 2, Issue 10). www.ijetae.com

Venkataraman, V., El-Kharouf, A., Pandya, B., Amakiri, E., & Steinberger-Wilckens, R. (2020). Coupling of engine exhaust and fuel cell exhaust with vapour absorption refrigeration/air conditioning systems for transport applications: A review. Thermal Science and Engineering Progress, 18, 100550. https://doi.org/10.1016/j.tsep.2020.100550

Verde, M., Harby, K., de Boer, R., & Corberán, J. M. (2016). Performance evaluation of a waste-heat driven adsorption system for automotive air-conditioning: Part I – Modeling and experimental validation. Energy, 116, 526–538. https://doi.org/10.1016/j.energy.2016.09.113

Waghchore, R. K., Jumde, M. R., & Somwanshi, A. H. (2013). INDIA. b,c Final year Mechanical Engineering Department. International Journal of Engineering Trends and Technology, 445001. http://www.ijettjournal.org

Wang, L. W., Wang, R. Z., & Oliveira, R. G. (2009). A review on adsorption working pairs for refrigeration. Renewable and Sustainable Energy Reviews, 13(3), 518–534. https://doi.org/10.1016/j.rser.2007.12.002

Wang, R., Wang, L., & Wu, J. (2014). Adsorption Refrigeration Technology: Theory and Application. Wiley Publication.

Yeo, T. H. C., Tan, I. A. W., & Abdullah, M. O. (2012). Development of adsorption air-conditioning technology using modified activated carbon - A review. Renewable and Sustainable Energy Reviews, 16(5), 3355–3363. https://doi.org/10.1016/j.rser.2012.02.073

Zeng, T., Huang, H., Kobayashi, N., & Li, J. (2017). Performance of an Activated Carbon-Ammonia Adsorption Refrigeration System. Natural Resources, 08(10), 611–631. https://doi.org/10.4236/nr.2017.810039 R.E. & Zadeh, L.A. (1970). Decision-making in a fuzzy environment, Management Science, 17(4), 141-164.

DOI: 10.6977/IJoSI.202409_8(3).0008
M. Kumar, U. Pilania /Int. J. Systematic Innovation, 8(3), 99-114 (2024)

# A framework for detection of drone using yolov5x for security surveillance system

Manoj Kumar[1], Urmila Pilania[2]*
[1,2]*Department of Computer Science, Manav Rachna University, Faridabad, India
urmilapilania@gmail.com
(Received 27 March 2024; Final version received 29 June 2024; Accepted 29 June 2024)

## Abstract

In response to the increasing use of drones, there is an emerging need for dependable security surveillance systems capable of detecting them. This article presents a theoretical framework for drone detection using the YOLOv5x (You Only Look Once) deep learning algorithm, which aims to improve security surveillance systems. The framework is composed of various hardware and software components, including drones, cameras, and computer systems, and utilizes YOLOv5x to accurately detect drones. To carry out the research, the Roboflow Drone 1 dataset is used, which includes a variety of images captured under various experimental conditions. The YOLOv5x algorithm is selected due to its high accuracy and low computational cost, which makes it an ideal solution for security surveillance systems. To measure the performance of the system, metrics such as F1-Score, Recall, Precision, Area under the Curve (AUC), and Mean Average Precision (mAP) are employed. The mAP is calculated at 0.95 at a learning rate of 0.5, precision is 0.901, recall is 0.97, AUC is 95.3, and F-1 score is 0.94. Overall, this framework provides a dependable and efficient approach for detecting drones, resulting in improved security of critical infrastructure, public events, and other sensitive locations. Furthermore, the proposed research is compared with existing state-of-the-artwork, and the experimental results verified that the proposed research outperformed the state-of-the-art.

*Keywords: Drone Detection*, YOLOv5x, Security, Surveillance, Unmanned Aerial Vehicle.

## 1. Introduction

A drone, named an Unmanned Aerial Vehicle (UAV), is a jet which is generally operated remotely by a human as per Kumar, M, et al. (2024). Drones have an extensive variety of applications, containing military operations, videography, surveying, inspection, and aerial photography. Modern drones are prepared having advanced features such as GPS, cameras, and autonomous flight capabilities. Drones can be categorized based on their cost, size, weight, capabilities, and intended use according to Taha, B., & Shoufan, A. (2019). This information is used to determine the appropriate level of regulations, safety measures, and limitations for the use of drones in various contexts. Some common categories of drones include as represented in **Table 1**:

**Table 1.** Types of Drones Zhu, X., et al. (2021)

| Characteristic | Consumer Drone | Commercial Drone | Military Drone | Agriculture Drone | Emergency Drone | Education Drone |
|---|---|---|---|---|---|---|
| Purpose | These types of devices are lightweight, small in size, and designed for leisure use, typically weighing less than 55 pounds | Larger, more capable drones designed for commercial use like midair photography, survey, package distribution, inspection of pipelines and power lines. | These types of drones are developed for military use like investigation, reconnaissance, and beset killings. | Drones specifically designed for agricultural applications, such as crop monitoring, mapping, and precision agriculture. | Drones are used for emergency response and disaster relief operations, such as search, rescue, and firefighting. | These are used for educational uses like teaching learners about drone machinery and flight. |
| Size | Small | Varies with application | Medium | Large depending upon application | Varies | Design as per use |

| | | | | | | |
|---|---|---|---|---|---|---|
| Range | Medium range | Short to long range | Long range | Medium depends on application | Varies depending upon application | Range varies with use |
| Sensor | Cameras and gyroscopes | Sensors for data collection | Cameras, radar, infrared | Multispectral and hyperspectral sensors | Advanced sensors for detecting heat and survivors. | Basic Sensors |
| Cost | Affordable | Cost is high due to more features | Very costly due to military grade material | Very costly due to specific instrument and sensors | Depending on features cost varies. | More affordable |

**Yolov5x-based drone classification:** It is a version of the YOLO object recognition method which could be utilized for drone classification. It is a deep learning technique which could automatically identify the type and size of a drone. Its intended use and capabilities can be analyzed using aerial images captured by the drone. The YOLOv5x works by separating an image into a network of cells by using a deep neural network to foresee the presence and class of objects within each cell. The model is skilled on a large dataset of pictures of different kinds of drones, so it can learn to identify specific features and patterns associated with each class of drone Pilania, U., et al. (2023).
After training the model, it could be applied to classify new images in real-time. By using YOLOv5x for drone classification, it is possible to quickly and accurately determine the type along with the abilities of a drone. It could be valuable for numerous applications, like security, surveillance, law enforcement, emergency response, and border protection. The proposed research works by training the model utilizing the roboflow drone 1 dataset of images to accurately detect drones. The YOLOv5x algorithm is used because of its good accuracy, very lower computational cost, allowing for efficient and effective surveillance. The method could be combined into existing surveillance systems to provide real-time monitoring, and alerts, enabling security personnel to take appropriate action when unauthorized drones are detected. The basic difference between YOLOv5 and YOLOv5x is given as in **Table 2**:

**Table 2.** Features of YOLOv5x

| Feature | YOLOv5 | YOLOv5x |
|---|---|---|
| Architecture | YOLOv5 comes in varying number of layers as per the requirement of the application. | It is designed with deeper and wider architecture. |
| Networks | It use CSPDarknet53 (Cross Stage Partial) network for extraction of features. | It uses ResNet with Cross Stage Partial connections Network (CSPNet) as backbone for extraction of more complex features. |
| Size | It uses small size variant resulting into low accuracy. | It uses large model size resulting into better accuracy. |
| Hardware Requirement | It can work with CPU and uses GPU for better performance. | It may use additional hardware for real-time applications because of high computational demands. |
| Resolution | It may work with varying resolution depending upon application. | It operates on high resolution resulting into good feature extraction and high accuracy. |
| Speed | Its speed is high due to due to smaller size of model. | Its speed is low due to large model size and complex structure. |

The main objective of the proposed work is to design a framework for drone detection using the YOLOv5x, precisely for security surveillance systems. The objective is to improve security and safety in domains like airports, public events, critical infrastructure, and private properties, by effectively identifying and monitoring the presence of drones in real-time.

## Research Contributions:

- The framework focuses on achieving real-time drone detection, which is crucial for security surveillance applications.
- To enhance the practicality and reliability of the system, the work contributes to reducing false alarms.

- It is applicable to different real-time security systems due to scalability and versatile nature.
- Proposed approach could be adapted to various types of cameras, sensors, and deployment scenarios.
- The research contributes to the growth of a user-friendly interface for security personnel. This includes a dashboard or software that provides a clear and intuitive display of detected drones and relevant information.
- Comprehensive performance evaluation and validation are conducted to measure the accuracy, precision, recall, and real-world applicability of the developed framework. This includes testing the model under numerous environmental conditions and drone types.

In summary, the research "A Framework for Detection of Drones using YOLOv5x for Security Surveillance System" aims to provide a robust, real-time solution for drone detection in security surveillance applications, with a focus on accuracy, speed, and practicality. The contributions lie in the customization of the YOLOv5x model, its integration, and the reduction of false alarms, all while addressing the security needs of various sectors.

## 2. Literature Review

The drones have been used gradually in recent years, as they have been established to be useful in numerous fields such as agriculture, construction, and aerial photography. However, the rise in unauthorized drone usage has led to significant security concerns. This literature review focuses on the framework for detecting drones using YOLOv5x for security surveillance systems.

These days' drones have been used widely due to their affordability and commercial availability. However, the use of drones also has increased the risk of illegal activities such as drug smuggling and terrorism. To prevent such actions, it is crucial to monitor and detect drones in restricted or special zones. One of the biggest challenges in drone detection is the resemblance between drones and birds, particularly when viewed against multifarious backgrounds in surveillance videos. To address this, a new image-based drone-detection model has been designed using YOLOv5 in Al-Qubaydhi, et al. (2022). Due to the limited availability of information, transfer learning was used to enhance the performance of the system. The results were outstanding, with an average precision rate of 94.7%. This system could be useful for safeguarding restricted areas and preventing illegal drone interventions.

The increasing quantity of drones has been used for both marketable and recreational purposes raised concerns about their potential misuse, including privacy violations and drug smuggling. However, spotting drones could be challenging due to other items in the sky, like aircraft, birds, and computerized systems requiring a huge amount of information and highly configured devices for real-time detection. To overcome the mentioned issues, the study Kumar, M., et al. (2024) proposed the use of a one-shot sensor called YOLOv5, which can be trained with pre-existing weights and data augmentation. The system was assessed using mAP, and recall, and attained a mAP of 90.40%, which is a 21.57% enhancement over the earlier model which applied YOLOv4 and was tried on the same dataset. This technology could suggest a solution to the problem of drone abuse and help protect people's privacy.

The advancements in drone technology have led to the emergence of object identification technologies which could be useful to various scenarios, including detecting illegal immigrants, natural disasters, missing people or objects, and industrial accidents. The authors De Galiza Barbosa, et al. (2023) aimed to discover ways to improve object recognition recitals under challenging conditions. Experimental data were gathered through photography in a confusing environment with varying environmental conditions. The F11 4K PRO drone and VisDrone dataset were used for the experiment. The study proposed a better form of the original YOLOv5 model and compared its performance with the original model. Key performance indicators, including precision, recall, F-1 score, and mAP (0.5), were calculated for both models. The improved YOLOv5_Ours model validated enhanced performance as mAP (0.5) and function loss than the original YOLOv5 model. Constructed on the statistical analysis, the study concludes by identifying a good object identification model for various challenging conditions. Detecting objects from drone-captured scenarios is a popular and recent task, but it comes with challenges. Drones navigate at different altitudes, resulting in varying object scales, which makes optimizing networks difficult. Additionally, motion blur from high-speed, low-altitude flights causes objects to become indistinguishable, adding to the challenge. To solve these challenges, the authors in Zhu, X., et al. (2021) proposed TPH-YOLOv5, a modified version of YOLOv5 that includes further prediction head to detect objects of diverse scales. TPH-YOLOv5 improved by about 7% compared to the baseline YOLOv5 model, which was promising and competitive.

Human action detection from drones has developed a significant challenge in current years, with potential applications in environmental monitoring, search, and rescue operations. Though, this challenge was complicated by the variability of human subjects' scales, orientations, and occlusions in drone-captured images. Authors Ahmad, T., et al. (2022) proposed low-resource machine learning approaches for action detection using the "Okutama-Action" dataset, which includes images with controlled image acquisition parameters. The proposed approach combines object recognition with a gradient-boosting classifier to identify actions in single images. The authors integrated YoloV5 with the classifier to achieve both scalability and efficiency in the object recognition system while accommodating the variable difficulty of samples. The proposed method beat erstwhile architectures used on the same dataset, which we attribute to the performance of Yolov5 and the adequacy of our pipeline to the Okutama dataset's specificities in terms of bias-variance tradeoff. The goal of drone detection is to locate the drone(s) within a video by identifying the smallest rectangle that encloses it. To resolve this issue, authors Aker, C., & Kalkan, S. (2017) presented an end-to-end object recognition method using Convolutional Neural Networks (CNN). To overcome the lack of information for training the network, the author developed a system that combines background-subtracted real images to create an extensive artificial dataset. This approach allowed us to achieve high precision and recall values simultaneously. After the literature survey, it has been concluded that there are several challenges associated with drone detection as represented by the **Table 3**:

**Table 3.** Mapping of issues with proposed work

| Challenge | Existing Work | Proposed Work |
|---|---|---|
| Size and Speed | Drones can vary in size and speed, making them difficult to detect. Small drones can be particularly challenging to detect, as they may not produce enough heat or sound to be detected by traditional sensors as per Singha, S., & Aydin, B. (2021) | Diverse training of the model on different datasets helps in handling varying size and shape of the drone. The proposed model is able to detect drones of varying size and shape. |
| Flight Patterns | Drones can fly in unpredictable patterns, making it difficult to predict where they will be at any given time. They can also fly at low altitudes, making them difficult to detect with radar as per paper Akyon, F. C., et al. (2021) | ResNet with CSPNet efficiently work on features such as shape, texture, color, and motion patterns. The model efficiently detects drones at very low height as well as at very high altitudes. |
| Background Noise | In noisy environments, it can be challenging to detect the sound of a drone, particularly if it is small and has a low acoustic signature as in paper Anwar, M. Z., et al. (2019) | As in feature extractions ResNet is used which effectively filter out the background noise. |
| Signal Interference | Drones can interfere with GPS and other communication signals, which can make it difficult to track their location Wu, M., et al. (2018) | Pre-processing and data augmentation techniques are used in YOLOv5x for signal inference. These techniques help in tracking the location of the signal accurately. |

## 3. Proposed Work

Traditional methods of detecting drone activity such as human surveillance, and radar systems, can be time-inefficient, resource-demanding, and could not always be effective. So, there is a requirement of more efficient and reliable method of detecting drones. To acknowledge the above challenges, the work proposes a framework for the identification of drones using the YOLOv5x deep learning algorithm. The framework provides a comprehensive solution for detecting drones in real-time, enhancing the security of critical infrastructure, public events, and other sensitive locations. The proposed work aims to develop an efficient and accurate drone identification system using modified YOLOv5x. For modification of YOLOv5x the backbone network, CSPDarknet53 is replaced with Residual Network (ResNet) and head network is updated to include extra layers for capturing the features more accurately. Additionally, the system will utilize advanced data augmentation techniques to improve its performance and robustness. The proposed model is trained using a large dataset of drone images under many environmental and lighting conditions. The proposed system's performance will be evaluated based on many constraints such as detection accuracy, recall, precision, and F1-score.

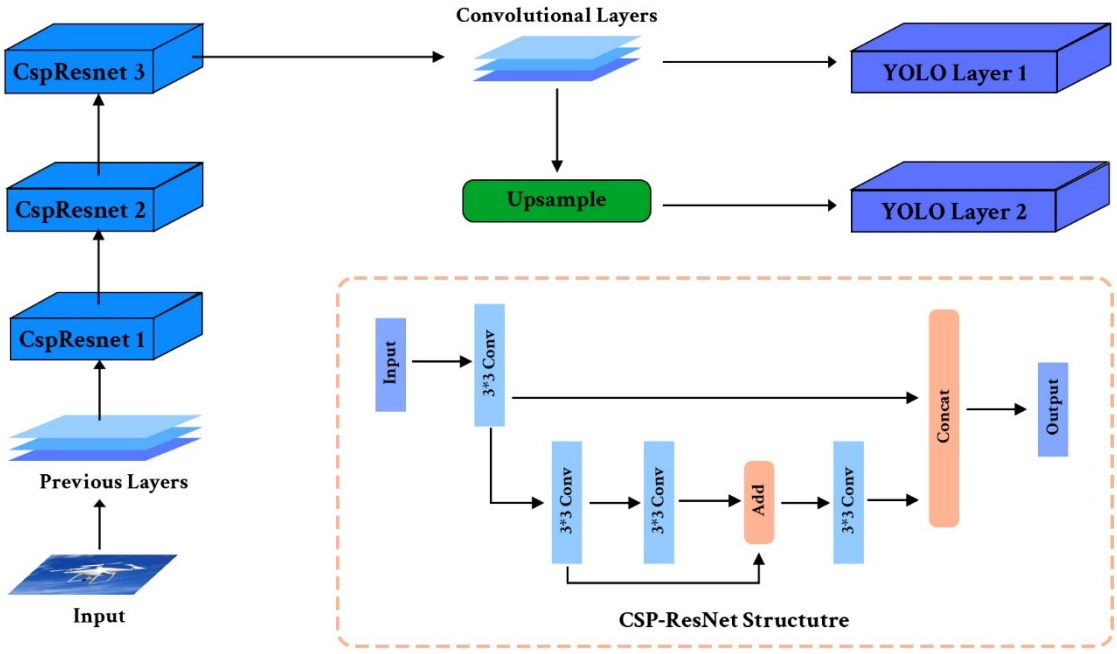



**Fig 1.** Layered Architecture of YOLOv5x

In the proposed work, Roboflow Drone 1 dataset is used which is a publicly available dataset containing images and annotations for drone detection. The dataset contains images of drones in various orientations, sizes, and backgrounds. The annotations include the bounding boxes and labels for the drones in the images. The dataset contains over 7,500 images with a resolution of 1080 x 720 pixels. The dataset includes a diversity of objects, such as buildings, trees, vehicles, birds and to simulate real-world scenarios. The Roboflow Drone 1 dataset is used for developing and testing deep-learning models for drone detection. It could be used to train and evaluate models based on deep learning algorithms, like YOLO, Faster R-CNN, and SSD. The dataset is available for download on the Roboflow website and can be accessed for free.

YOLOv5x is a larger and more accurate version of YOLOv5, designed to run on high-performance GPUs. It has a larger number of layers and parameters, allowing it to perform more complex computations and achieve higher accuracy in object detection tasks. In summary, YOLOv5s is optimized for speed and efficiency, while YOLOv5x is optimized for accuracy. The choice of a particular model depends on the specific use case and the available computational resources.

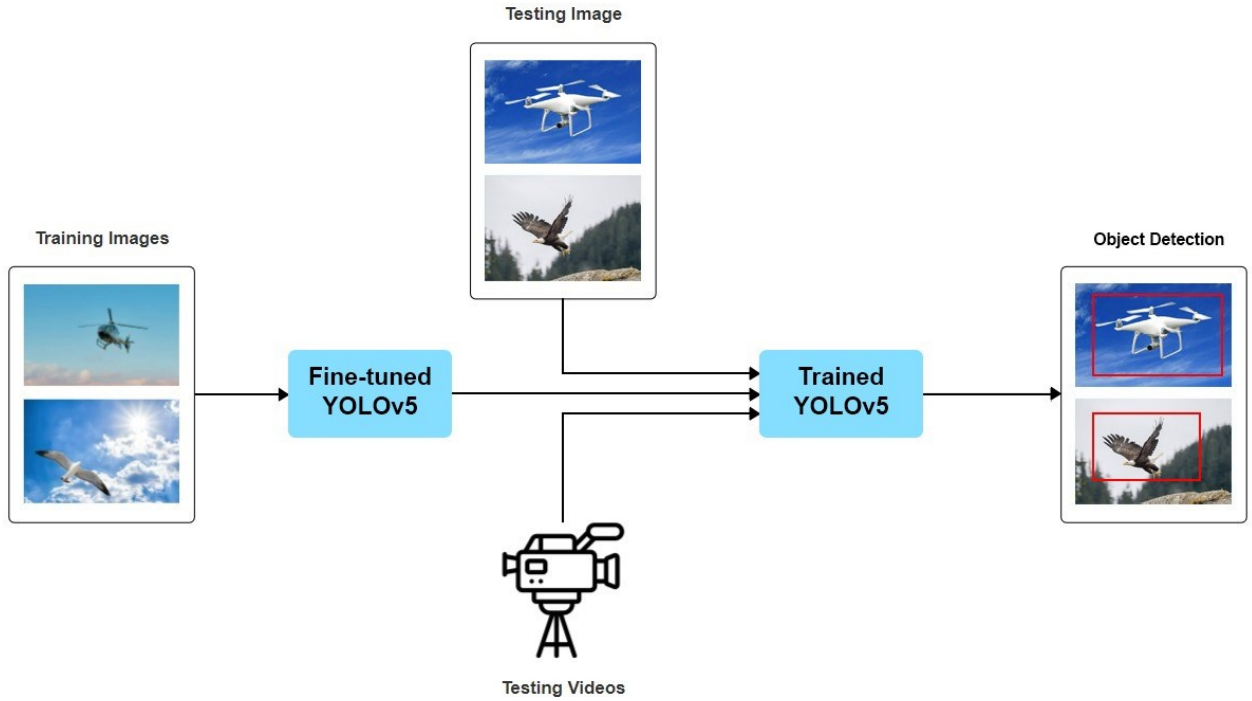



**Fig 2.**Proposed Flowchart

YOLOv5x is a standard object recognition model which uses a layered architecture to identify and categorize objects in images as depicted in **Figure 1**. **Figure 2** represents the proposed flowchart for the research work. **Figure 3** represents the scattering of the sizes and positions of drones in the dataset. The architecture of YOLOv5x is divided into several key components.

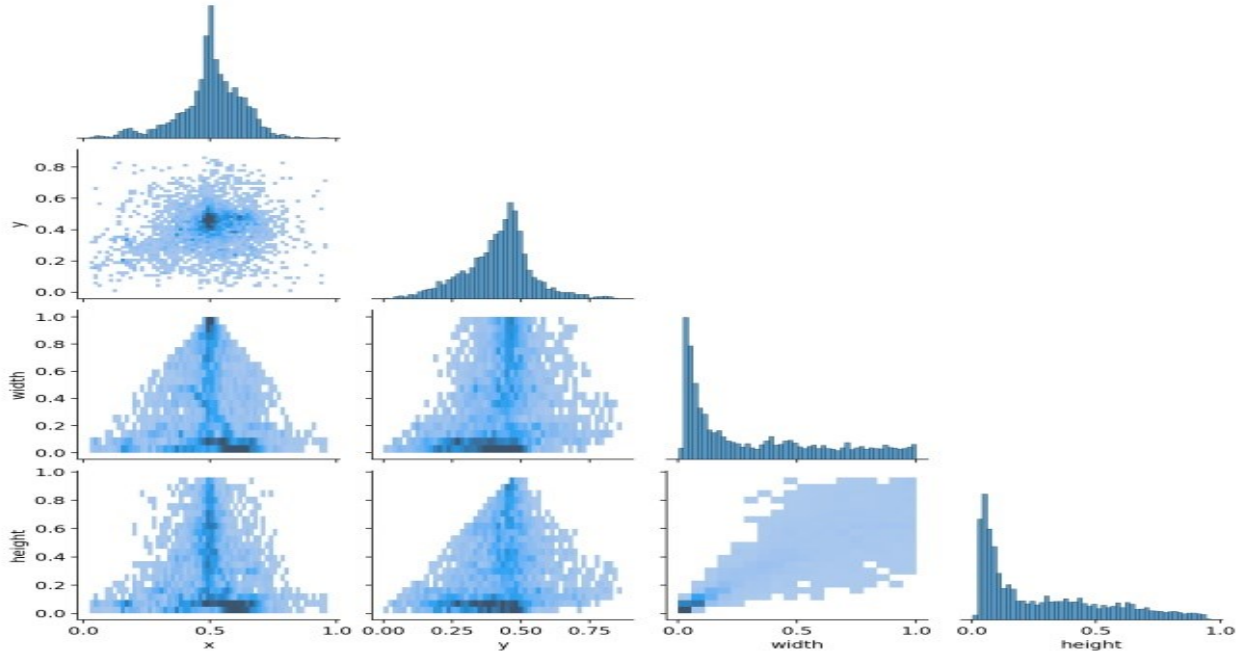



**Fig 3.**Distribution of the Sizes and Positions of drone in dataset

The backbone network of YOLOv5x consists of a deep CNN which retrieves features from the input image. It is built on the ResNet design, which improves the efficiency of feature extraction. CSPDarknet53 in the basic architecture of YOLOv5x has been replaced by the ResNet. ResNet has the ability to extract robust features from the input images. The required changes have been done to adjust ResNet with the neck and head network of the YOLOv5x. The feature map generated by the ResNet is compatible with different layers of YOLOv5x. ResNet has a sufficient number of resources for training the dataset and interfaces. The integrity of the ResNet is monitored using performance metrics. The purpose of changing the CSPDarknet53 to ResNet help in down-sampling operations, and skip connections which help in feature extraction at different spatial resolutions. ResNet uses CSP connections at different levels in the network which enable reuse and reduced computational efficiency. The CSP connections enhance the way of representation of the features and retrieve contextual data from the images. This network can handle images with different backgrounds, shapes, poses, and sizes. It makes a foundation for the detection of drones in the input images.

It is responsible for aggregating and refining the features retrieved by the backbone network. To achieve this, YOLOv5x employs a Spatial Pyramid Pooling (SPP) module in its neck network, which enables the model to capture features at various scales in an efficient manner. It incorporates many stages after the backbone network for the refinement of features extracted by the backbone network. All these stages focus on enhancing the presentation of extracted features, context capturing, and information aggregation. All these steps lead to more accurate detection of drones in minimum time.

The head network of YOLOv5x is in charge of forecasting bounding boxes and class likelihoods for objects in the image. Contrasting classical object recognition approaches that use anchor boxes, YOLOv5x uses an anchor-free detection approach that eliminates the need for anchor boxes, resulting in more precise object detection. It is trained by the backbone network and intermediate stages to detect drones in input images. Due to the process of training, it is able to put the boundary box on the detected drones.

The last stage is post-processing. The predictions made by the head network of YOLOv5x are post-processed to enhance the final results by eliminating redundant detections. This is accomplished using a Non-Maximum Suppression (NMS) algorithm, which removes overlapping detections and retains the most confident detections for each object. YOLOv5x employs a layered architecture that empowers it to achieve top-notch performance in object detection tasks, exhibiting high precision and efficiency.

## 4. Result Analysis

The evaluation of performance metrics for drone detection may differ based on the technology employed and the intended use case. YOLOv5x algorithm's drone detection performance may be influenced by several factors, like drone size, speed, lighting conditions, and weather. Hence, it's necessary to assess the performance metrics in various scenarios to ensure that the algorithm can function efficiently in diverse environments. YOLOv5x is a more advanced version of the Yolov5 object detection algorithm, which employs a larger and more intricate model architecture. The training and validation procedures for Yolov5x are similar to those for the standard YOLOv5 algorithm. To train a Yolov5x model, the dataset is first split into training and validation sets using the --split flag when running the train.py script. During the training phase, YOLOv5x generates several metrics like training loss, validation loss, average precision, and recall. These metrics are used to access the model's performance during training and make necessary adjustments to hyperparameters.

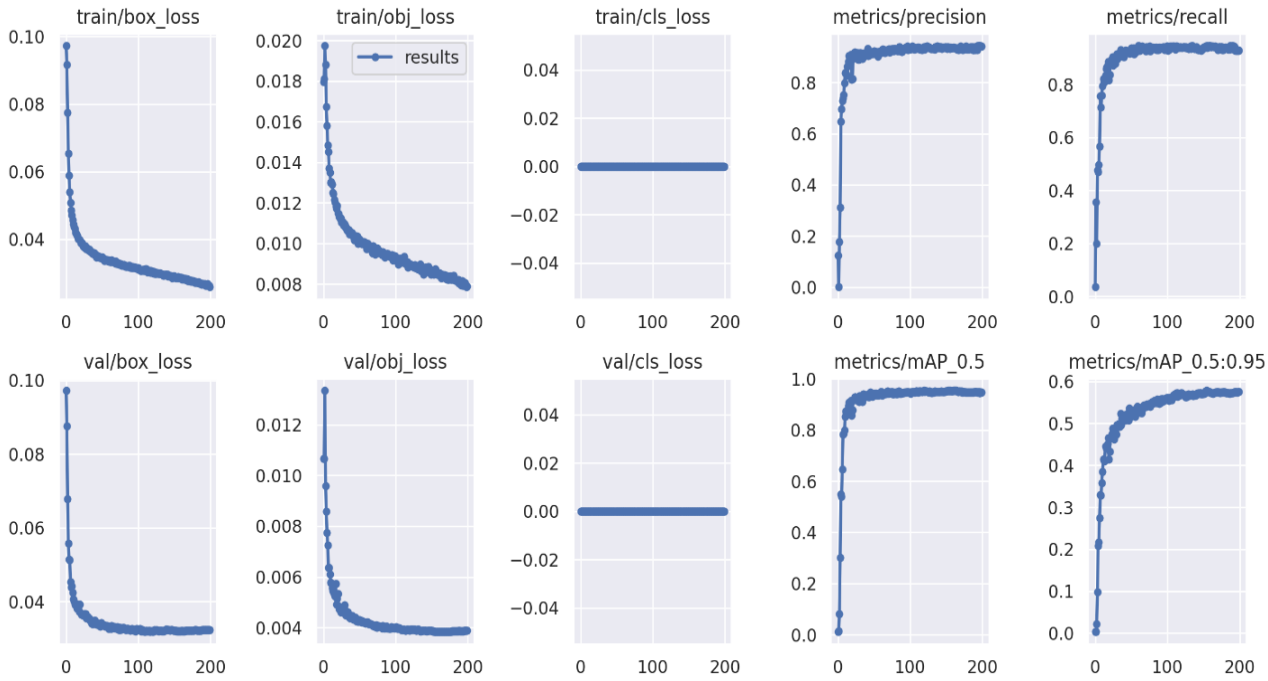



**Fig 4.** Training and Validation Loss

Training a YOLOv5x model can be computationally demanding, requiring a high-performance GPU and sufficient memory to effectively train the model. The x-axis in **Figure 4** represents the number of epochs, which is set to 200 for this work. The required number of epochs to train YOLOv5x for drone detection may vary depending on factors such as the dataset size, the complexity of the detection task, and the hardware configuration.

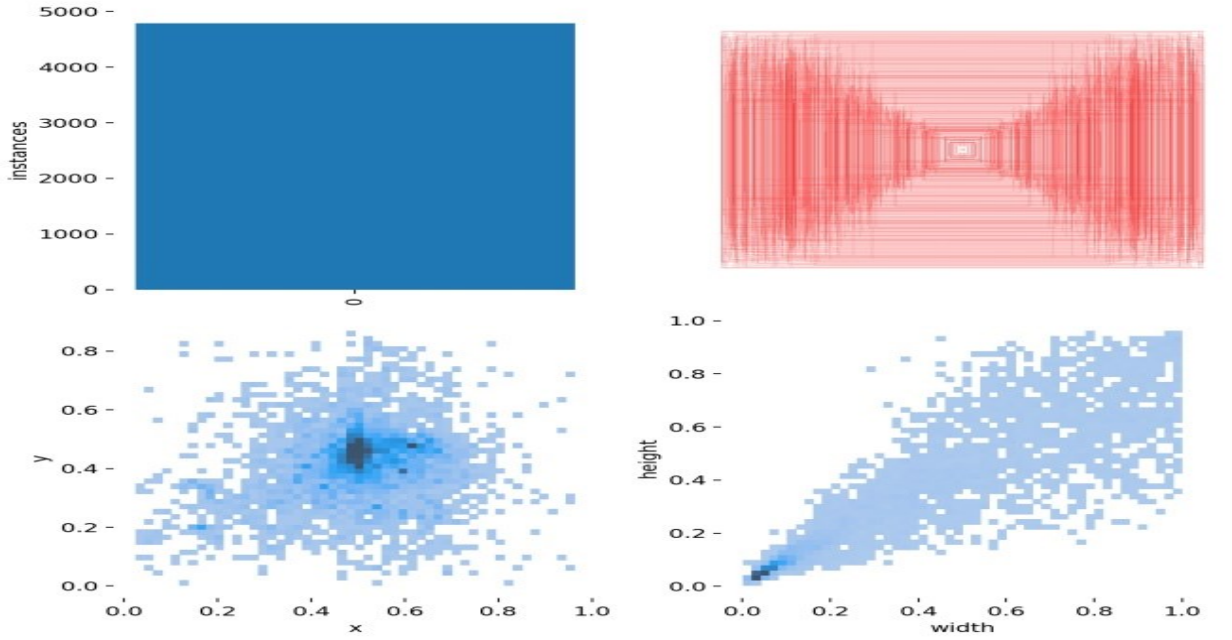



**Fig 5.** Distribution and Histogram of Drone in Training Set

Box loss, as represented in **Figure 4**, measures the change between the expected and true bounding box coordinates of an object. The Fig. 5 represents the distribution and histogram of the drone training set. Distribution means how data is stored in the dataset and histogram is used to represent this data graphically. Distribution helps in gaining knowledge about the features of data in the dataset. A histogram is also used to check the patterns in the dataset. The dataset is properly annotated and formatted as per the requirement of YOLOv5x. During training the number of input images, average number of drones per image, and maximum number of drones per image have been calculated. The occurrence of drones is retrieved and counted to check the patterns in the dataset. By using these patterns, it creates the histogram for the input. It is typically calculated using a regression loss function such as MSE or plane L1 loss. Eq. (1) and (2) represent the bounding box and object box mathematically.

$$L_{box} = \lambda_{coordinate} \sum_{n=0}^{k^2} \times \sum_{m=0}^{l} [\, 1_{n,m}^{obj} [(o_n - \hat{o}_n)^2 + (h_n - \hat{h}_n)^2]] + \lambda_{coordinate} \sum_{n=0}^{k^2} \times \sum_{m=0}^{l} [\, 1_{n,m}^{obj} [(\sqrt{c_n} - \sqrt{\hat{c}_n})^2 + (\sqrt{e_n} - \sqrt{\hat{e}_n})^2]] \quad (1)$$

where k is the numeral of grid cells in the input, l is the numeral of bounding boxes expected by each grid cell, ô, ĥ, ĉ, ê are the predicted coordinates of the center, width, and height of the box in cell n, and $o_n$, $h_n$, $c_n$, $e_n$ are the ground-truth coordinates of the center, width, and height of the box in cell n. The term $1_{n,m}^{obj}$ is an indicator function that equals 1 if the m bounding box in cell n is responsible for detecting an object, and 0 otherwise. The parameter $\lambda_{coordinate}$ controls the weight of the box localization loss.

$$L_{obj} = \sum_{n=0}^{k^2} \times \sum_{m=0}^{l} 1_{n,m}^{obj} \left[\log\left(\sigma(\hat{\imath}_{n,m})\right) + \left(1 - 1_{n,m}^{obj}\right) \log\left(1 - \sigma(\hat{\imath}_{n,m})\right)\right] \quad 1 \quad (2)$$

where $\hat{\imath}_{n,m}$ is the predicted objectness score for the $m^{th}$ bounding box in cell n, and σ is the sigmoid function. In the term $1 - 1_{n,m}^{obj}$ is an indicator function that equals 1 if the $m^{th}$ bounding box in cell n is not responsible for detecting an object and 0 otherwise.

$$loss_{YOLOv5} = loss_{bounding\ box} + loss_{classification} + loss_{confidence} \quad (3)$$

$$\text{loss}_{\text{bounding box}} = \lambda_{\text{if}} \sum_{a=0}^{b^2} \times \sum_{c=0}^{d} E_{a,c}^{g}\, h_g(2 - K_a X n_a) \left[(x_a - x_a^{'c})^2 + (y_a - y_a^{'c})^2 + (w_a - w_a^{'c})^2 + (h_a - h_a^{'c})^2\right] \quad (4)$$

$$loss_{classification} = \lambda_{classification} \sum_{a=0}^{b^2} \times \sum_{c=0}^{d} E_{a,c}^{g} \times \sum_{C \in c_l} L_a(c) \log\left(LL_a(c)\right) \quad (5)$$

$$loss_{confidence} = \lambda_{confidence} \sum_{a=0}^{b^2} \times \sum_{c=0}^{d} E_{a,c}^{confidence}\, (c_i - c_l)^2 + \lambda_g \sum_{a=0}^{b^2} \times \sum_{c=0}^{d} E_{a,c}^{g}\, (c_i - c_l)^2 \quad (6)$$

Loss function in YOLOv5x plays significant role in enhancing the outcomes. It is the combination of boundary box, classification and the confidence as represented by Eq. (3). Eq. (4) shows how bounding box can be calculated. Where h′ and w′ shows the height and width of target object in the input image. $x_a$ and $y_a$ represent the coordinates of the input image. $\lambda_{if}$ indicated whether the target object exists in the image or not. $\lambda_{confidence}$ in Eq. (5) and (6) represent loss coefficient. $\lambda_{classification}$ is classification loss coefficient, $c_l$ is class, and c is the confidence score.

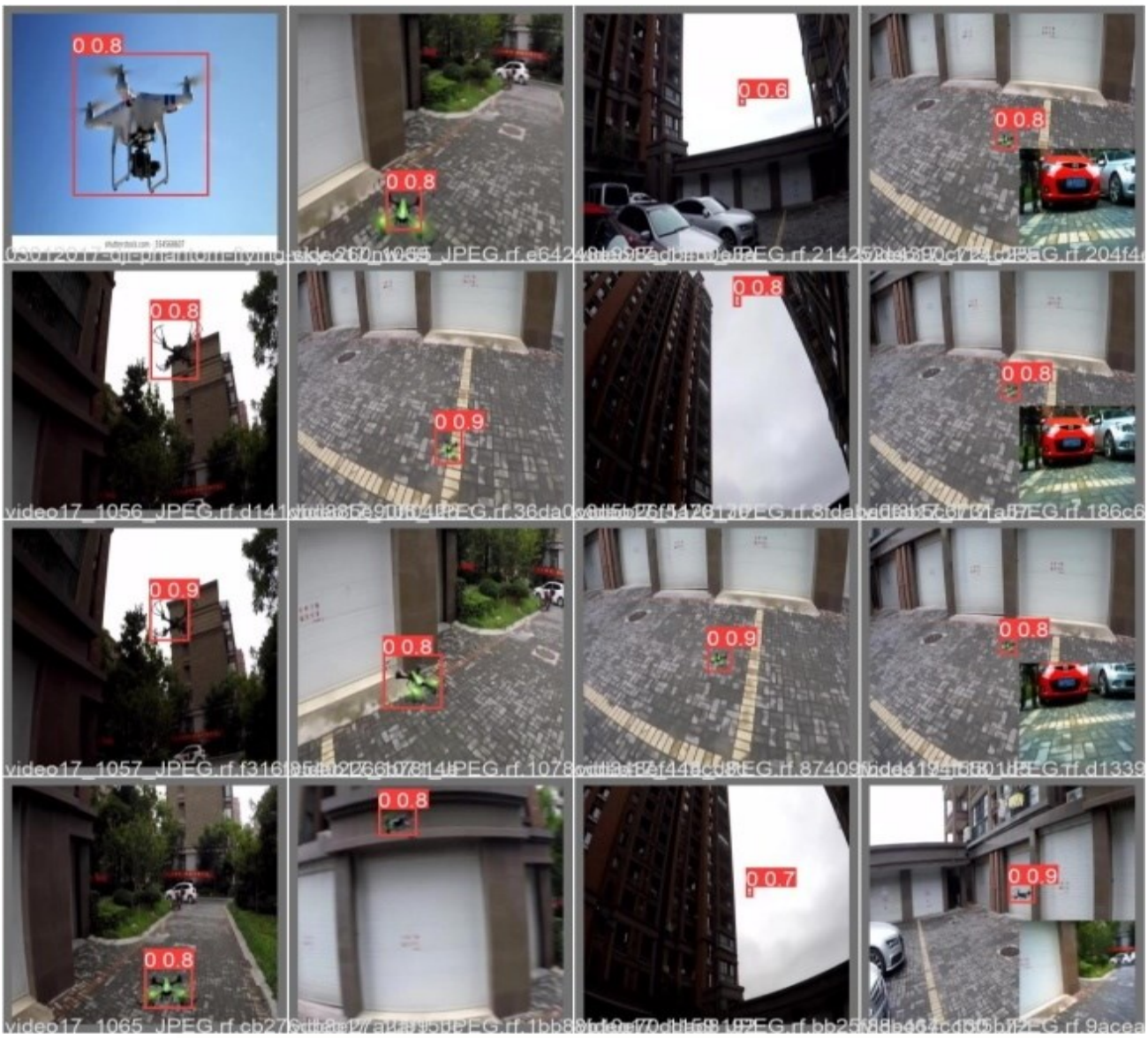


**Fig 6.** Drone Detection

Object loss is the loss function which is used to decide if an object exists or not in input image. It is calculated using a sigmoid activation function and binary cross-entropy loss Liu, X., et al. (2023) The purpose of object loss is to adjust the model's parameters so that it can more accurately detect objects in an image. Class loss is used for multi-class classification responsibilities. It measures the difference between the predicted probabilities and the true probabilities of an object belonging to different classes. The class loss is computed using a softmax activation function and categorical cross-entropy loss. The purpose of class loss is to enhance the model's parameters so that it can better classify objects into their respective categories. For the reduction of object loss in the proposed work, we have used high quality dataset and a regress training process.

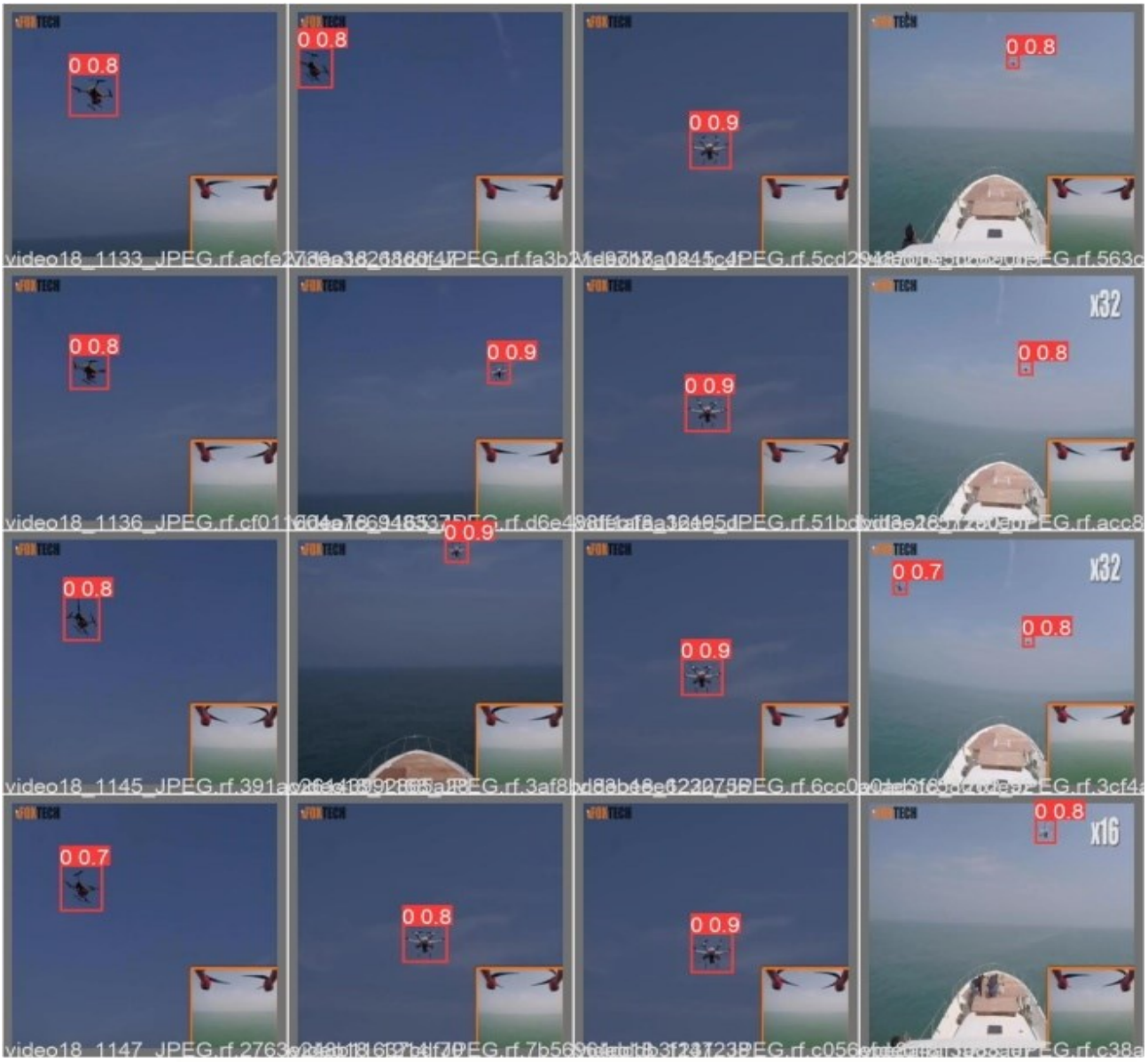



**Fig 7.**Drone Detection

False detection could be defined in terms of false positives which means the detection of a drone where it does not exist, false negative means the actual drone is present but the model is not able to detect it, and sometimes the model may label incorrectly to the detected object. As shown in Figure 7 the drones are detected but the accuracy of detection is not 100% which means the model is not 100% sure whether the detected object is a drone or some other object. As per the results model has a 4.7% false detection rate.

Confusion matrix is a commonly used table for assessing the enactment of a classification approach. It matches the expected labels of the approach with the actual labels in the validation or test set. In the context of object detection, a confusion matrix can be utilized to assess the model's ability to detect objects of various classes Abdallah, S. M. (2024). As shown in Figures 6 and 7, the results of YOLOv5x detection demonstrate that the deep learning method could detect drones with high accuracy.

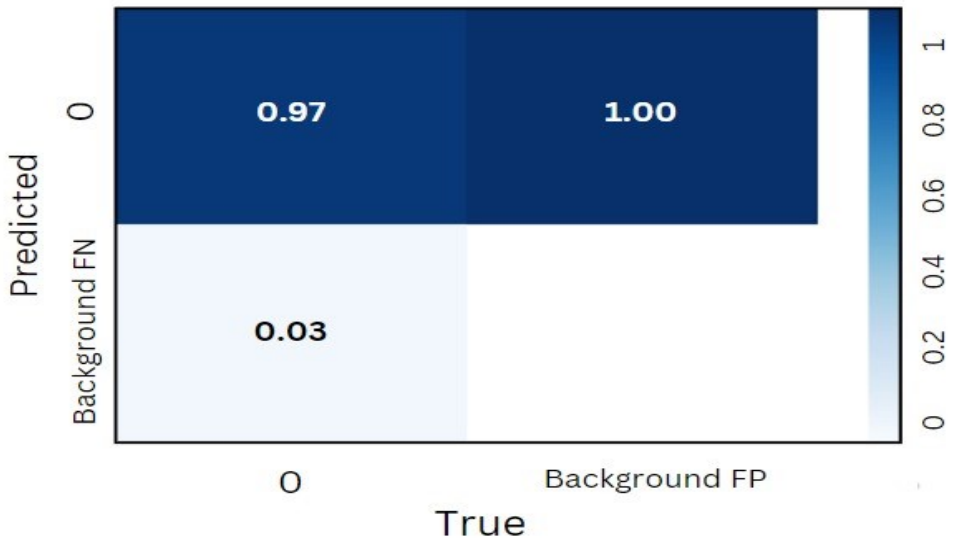
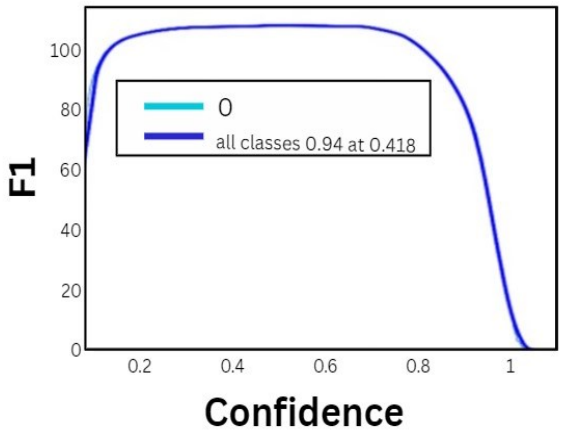



**Fig 8.** Confusion Matrix

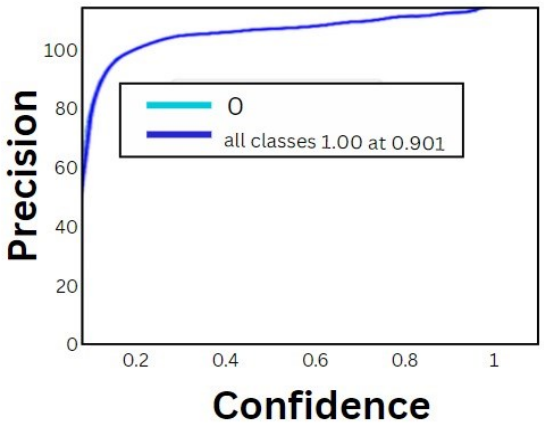
The validation or test set produces a confusion matrix that represents the numeral of True Positive (TP), false positive, false negative, and true negative expectations for each class, as illustrated in **Figure 8**. This information can be analyzed to evaluate the method's performance in identifying objects of diverse classes and identify opportunities for improvement Cheng, S., Zhu, Y., & Wu, S. (2023).

**i) Reliability and Accuracy Analysis**

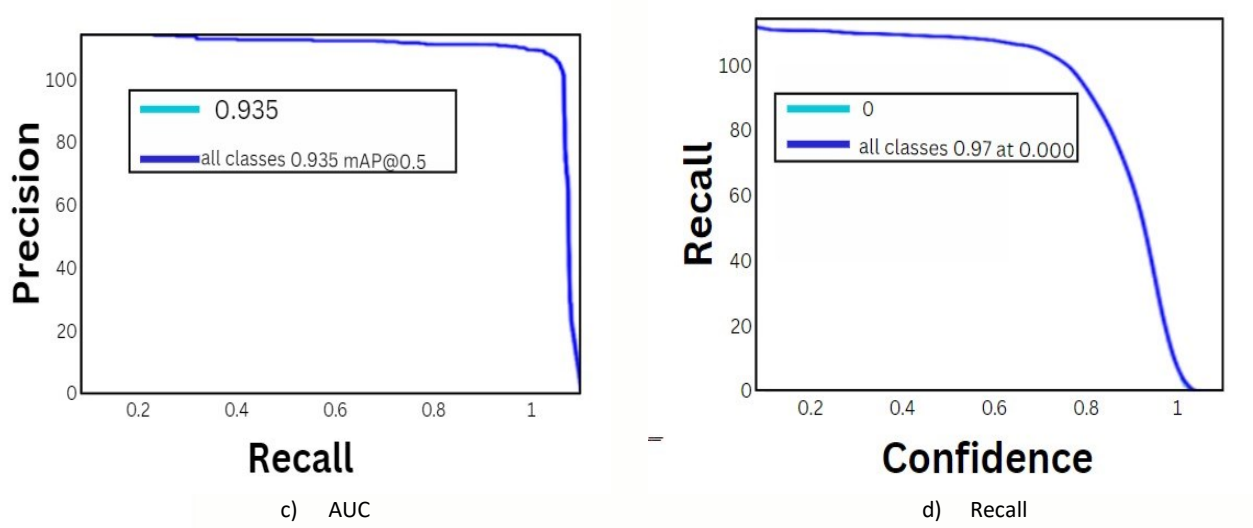
The reliability and accuracy analysis of drone detection using YOLOv5x assesses the model's ability to consistently perform over time and under different conditions. This analysis includes measuring the precision, F1-Score, AUC, and recall rates over multiple test runs, and checking for false positives and false negatives. Also, evaluating the model's performance under variable environmental circumstances such as changes in light, weather, drone size, and drone distance.




a) F1-Score




b) Precision




c) AUC d) Recall

**Fig 9.** Performance Metrics

The proposed work is found to be reliable and accurate based on values of performance metrics obtained from experimental results.

**F1-Score:** It is a performance metric which uses precision and recall in measuring a model's accuracy. It is the harmonic mean of precision and recall as per authors Li, S., et al (2023) in Eq. (7), and the F1-Score value was determined to be approximately 0.94 for the research work.

$$F1 = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad (7)$$

**Precision**: It measures the accuracy of the positive predictions made by the system, as represented in Eq. (8). In YOLOv5x, precision is utilized to assess the model's capability to accurately identify instances of a particular object class according to Alsanad, H. R.,. et al. (2023) The precision was calculated to be around 0.90 from the experimental work.

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)} \quad (8)$$

TP refers to a numeral of positive detections accurately made by the system, while False Positives refer to the numeral of positive findings incorrectly made by the system.

**AUC:** The AUC is a widely used metric to assess the Performance of machine learning systems in diverse applications, including drone detection. In drone detection, the AUC is commonly used to find the machine learning model's ability to accurately differentiate between the drone and non-drone signals as per authors Kumar, S., et al. (2024). At a learning rate of 0.5, the mAP or AUC was determined to be 0.95.

**Recall:** Recall measures the accuracy of the positive cases detected by the system. In the context of YOLOv5x, recall can be utilized to evaluate the model's ability to detect all instances of a particular object class in an image according to Kumar, S., et al (2023). It can be calculated using Eq. (9). The recall for the work was calculated to be equal to 0.97.

$$Recall = \frac{True\ Positives}{(True\ Positives + False\ Negatives)} \quad (9)$$

The TP denote to the numeral of accurate positive detections made by the system, while the False Negatives correspond to the numeral of positive detections missed by the model.

The computational outcomes of the work are illustrated in **Figure 9**. The F1-Score value was determined to be approximately 0.94, while the precision was calculated to be around 0.90. At a learning rate of 0.5, the mAP or AUC was determined to be 0.95. Furthermore, the recall for the work was calculated to be equal to 0.97. From the simulation, it has been determined that the proposed work is found to be reliable and accurate.

**ii) Complexity Analysis**

The complexity analysis of drone detection using YOLOv5x entails assessing the computational resources necessary to perform object detection on a given dataset. This includes measuring the computational workload per image, the memory consumption during processing, and the detection and tracking time for drones. Moreover, the analysis involves identifying any processing bottlenecks and optimizing the model to reduce computational complexity. Figure 10

depicts the detection time required for drones, which was found to be shorter than that of previous works in the field. For storing the data and instruction cloud google collab was used. Furthermore, YOLOv5x's image processing does not require significant memory compared to video databases.

**iii) Security Analysis**

YOLOv5x serves as a powerful tool for real-time object detection, extending its capability to both images and videos. To bolster the security of the proposed work, we meticulously curated a diverse dataset encompassing various critical variables, including backgrounds, lighting conditions, drone types, and other essential factors. Subsequently, we applied data augmentation techniques, introducing features like rotation, transformation, scaling, and lighting variations.

Our integration strategy extends to encompass security infrastructure elements, including CCTV cameras, alarm systems, and autonomous response mechanisms. Moreover, we diligently ensured that our proposed system adhered to all relevant regulations pertaining to privacy and surveillance.

**iv) Comparison of Proposed Work**

The proposed research has yielded superior results compared to the work in the paper Aydin, B., & Singha, S. (2023). In this research work, a mAP of 0.95 at a learning rate of 0.5 has been achieved. YOLOv5x, which employs a lightweight design and ResNet architecture, outperformed YOLOv5 in drone detection speed. YOLOv5x's PyTorch framework and fine-tuning the last layers of the architecture helped to optimize for the customized dataset. The values of the momentum, learning rate, and batch size are also customized. The proposed work underwent 200 iterations of training on top of reassigned weights for the roboflow drone 1 dataset.

The mAP is calculated 0.95 at a learning rate of 0.5 which is found to be better compared to the work by authors Aydin, B., & Singha, S. (2023). The precision is calculated 0.901 for the proposed work which is 0.01 smaller than the state of art work. The metric recall is calculated 0.97 which is much better than the work performed by authors Aydin, B., & Singha, S. (2023). F-1 score is calculated as 0.94 and for the state of art work, it is found to be 0.89.

**Figure 10** represents a comparison of the performance of the proposed system and the previous system in terms of 4 assessment metrics: precision, recall, F1 score, and mAP@0.5. We also calculated the AUC which is found to be 95.3% but in comparison work, it has not been calculated so we are not able to compare the AUC metric.

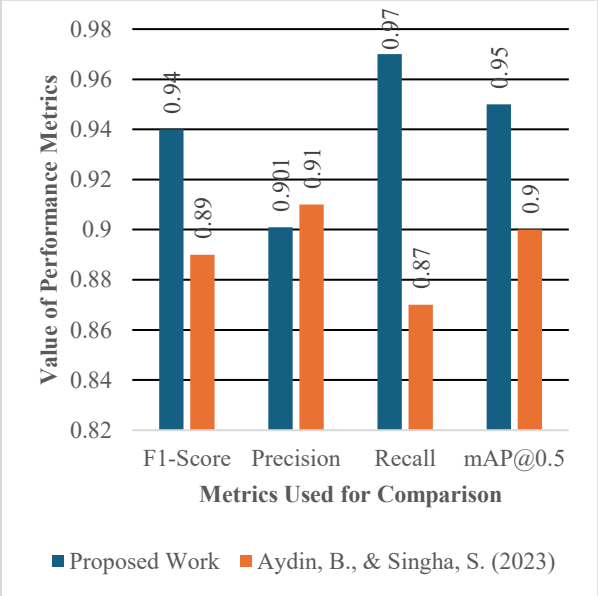



**Fig 10.** Comparison Graph

## 5. Conclusion

In conclusion, the proposed framework for the detection of drones using the YOLOv5x deep learning algorithm provides an effective solution for enhancing security surveillance systems. The framework leverages a combination of hardware and software components, including drones, cameras, and computer systems, to accurately detect unauthorized drone activity. The framework incorporates real-time monitoring capabilities, allowing for quick and effective responses to security threats. The YOLOv5x algorithm is chosen due to its high accuracy and low computational cost, making it an ideal solution for security surveillance systems. The mAP is calculated 0.95 at a learning rate of 0.5, precision is 0.901, recall is 0.97, AUC is 95.3, and F-1 score is 0.94.

The proposed framework is designed to be scalable, and customizable, enabling easy integration into existing security systems. It provides a reliable and effective method for detecting drone activity and enhancing the security of critical infrastructure, public events, and other sensitive locations.

Overall, the proposed framework is a valuable contribution to the field of security surveillance, providing a comprehensive solution for detecting unauthorized drone activity. Further development and customization of the framework can be done to meet specific security requirements.

## Reference

Abdallah, S. M. (2024). Real-Time Vehicles Detection Using a Tello Drone with YOLOv5 Algorithm. Cihan University-Erbil Scientific Journal, 8(1), 1-7.

Ahmad, T., Cavazza, M., Matsuo, Y., & Prendinger, H. (2022). Detecting human actions in drone images using YOLOv5 and stochastic gradient boosting. Sensors, 22(18), 7020.

Aker, C., & Kalkan, S. (2017). Using deep networks for drone detection. In 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS) (pp. 1-6). IEEE.

Al-Qubaydhi, Nader, Abdulrahman Alenezi, Turki Alanazi, Abdulrahman Senyor, Naif Alanezi, Bandar Alotaibi, Munif Alotaibi, Abdul Razaque, Abdelaziz A. Abdelhamid, and Aziz Alotaibi. (2022). Detection of unauthorized unmanned aerial vehicles using YOLOv5 and transfer learning, Electronics 11, no. 17 (2022): 2669.

Akyon, F. C., Eryuksel, O., Ozfuttu, K. A., & Altinuc, S. O. (2021). Track boosting and synthetic data aided drone detection. In 2021 17th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (pp. 1-5). IEEE.

Anwar, M. Z., Kaleem, Z., & Jamalipour, A. (2019). Machine learning inspired sound-based amateur drone detection for public safety applications. IEEE Transactions on Vehicular Technology, 68(3), 2526-2534.

Aydin, B., & Singha, S. (2023). Drone detection using yolov5. Eng, 4(1), 416-433.

Cheng, S., Zhu, Y., & Wu, S. (2023). Deep learning based efficient ship detection from drone-captured images for maritime surveillance. Ocean Engineering, 285, 115440.

De Galiza Barbosa, Felipe, Samuel J. Galgano, Ariel L. Botwin, Aline Bobato Lara Gongora, Giovanna Sawaya, Ronaldo H. Baroni, and Marcelo. (2023). A. Queiroz. "Genitourinary imaging." In Clinical PET/MRI, pp. 289-312. Academic Press, 2023.

Kumar, M., Pilania, U., Bhayana, T., & Thakur, S. (2024). Utilizing YOLOv5x for the Detection and Classification of Brain Tumors. In 2024 2nd International Conference on Disruptive Technologies (ICDT) (pp. 1343-1348). IEEE.

Kumar, M., Pilania, U., Thakur, S., & Bhayana, T. (2024). YOLOv5x-based Brain Tumor Detection for Healthcare Applications. Procedia Computer Science, 233, 950-959.

Kumar, S., Pilania, U., & Bala, R. (2024). A novel hybrid deep belief Google network framework for brain tumor classification. International Journal of Systematic Innovation, 8(2), 95-112.

Li, S., Yang, X., Lin, X., Zhang, Y., & Wu, J. (2023). Real-time vehicle detection from UAV aerial images based on improved YOLOv5. Sensors, 23(12), 5634.

Pilania, U., Kumar, M., Singh, S., & Adhalkha, P. (2023). A Deep Ensemble Learning Network-based Approach to Detect Neurodegenerative Diseases. In 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), 2023, pp. 1033-1038. IEEE.

Taha, B., & Shoufan, A. (2019). Machine learning-based drone detection and classification: State-of-the-art in research. IEEE access, 2019, 7, 138669-138682.

Wu, M., Xie, W., Shi, X., Shao, P., & Shi, Z. (2018). Real-time drone detection using deep learning approach. In Machine Learning and Intelligent Communications: Third International Conference, MLICOM 2018, Hangzhou, China, July 6-8, 2018, Proceedings 3 (pp. 22-32). Springer International Publishing.

Zhu, X., Lyu, S., Wang, X., & Zhao, Q. (2021). TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 2778-2788).

## AUTHOR BIOGRAPHIES

**Dr. Urmila Pilania is** currently working as Associate Professor in the Department of Computer Science & Technology at Manav Rachna University, Faridabad. She is having 14 years of experience in academics and research. Dr. Urmila Pilania received the Engineering degree (B. TECH. in Computer Science & Engineering) in 2007 from Guru Jambheshwar University of Science and Technology, Hisar - Haryana (India). Dr. Urmila has earned master's degree (M.TECH. in Computer Science & Technology) in 2009 from Vaish College of Engineering, Rohtak-Haryana. Dr. Pilania has been awarded with the doctoral degree (Ph.D. in Computer Science & Technology) degree in 2021 from Manav Rachna University, Faridabad, India. Her primary research interests are in Information Security, Image Processing, Machine Learning and Artificial Intelligence. Dr. Urmila has published many research papers related to Information Security, object detection and classification. She is author of more than 34 research publications with journals/Conferences of repute and some papers are under review or in press for publication with leading publisher.

**Dr. Manoj Kumar** is presently working as Associate Professor in the Department of Computer Science & Technology, at Manav Rachna University, Faridabad. He has over 14 years of experience including academics, industry & research. He has earned his Ph.D. in the year 2022, in the area of Computer Vision. The title of the thesis work is "Automatic Detection of Moving Object in a Video for smart Surveillance system". His area of expertise is Computer Vision, Artificial Intelligence, Machine Learning, Image Processing, IOT and Blockchain. Dr. Kumar has published many research papers related to moving object detection and tracking in videos and also extracts features moving object/vehicles from video in real-time video surveillance system. Dr. Kumar has published more than 10 papers in SCIE, SCOPUS, WEB OF SCIENCE indexed journals. Dr. Kumar has presented more than 20 papers in various Scopus Indexed International Conferences. The proceedings of the conferences are Scopus indexed and few of them published as a Book chapters. He has published 4 National Patents. He has participated in various workshops & Faculty development programs.

DOI: 10.6977/IJoSI.202409_8(3).0009
V. Jayashekar, R. Pandian.etc /Int. J. Systematic Innovation, 8(3), 115-124 (2024)

# Deep learning based classification of motor imagery EEG signals using an improved path finder optimization algorithm

Vishwesh Jayashekar[1], Raviraj Pandian[2], Rajashekar Mallajamma Basavarajegowda[3]

[1]Assosiate Professor, Department of Computer Science & Engineering,
GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated to VTU, Belagavi, Karnataka, India

[2]Professor, Department of Computer Science & Engineering,
GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated to VTU, Belagavi, Karnataka, India

3Assosiate Professor, Department of Computer Science & Engineering,
GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated to VTU, Belagavi, Karnataka, India

(Received 17 April 2024; Final version received 17 July 2024; Accepted 29 July 2024)

## Abstract

Motor Imagery Brain-Computer Interfaces (MI-BCIs) are systems based on AI that collect patterns of brain activities in mental movement and translate these movements through external devices. The identification of motor intention by evaluating Electroencephalogram (EEG) signals is an important issue in applications related to Brain-Computer Interfaces (BCI). In this paper, an Improved Path Finder Optimization Algorithm (IPFOA) is proposed to improve the process of selecting features. The data is collected from the BCI competition IV 2a dataset and the stage of preprocessing takes place through sliding windows and 6th order Butterworth filter. The feature will be extracted from the pre-processed data using the hybrid Convolutional Neural Network (CNN) and Common Spatial Pattern (CSP) method. After this stage, the required features are selected from the extracted features using the proposed IPFOA algorithm. Finally, the classification of EEG signals takes place using a Stacked autoencoder, a classifier based on a Deep Neural Network. The experimental results show that the proposed approach achieved a better accuracy of 98.40% which is comparatively higher than the existing approaches.

*Keywords: Brain Computer Interface, Convolutional Neural Networks, Deep Learning, Electroencephalogram, Motor Imagery, Path Finder Optimization.*

## 1. Introduction

A control system known as a brain-computer interface (BCI) facilitates communication between the brain and other devices by using signals produced by brain activity Electroencephalogram (EEG) vibrations created by the brain's response are typically computed based on cerebral actions (Xiao X et al. 2021 & Yu X et al. 2021). The vibrations are in the pattern of bioelectric signals which are easy to operate and safe. The system produces direct communication between the brain and the machine without any intermediates; it is also considered as a modern and extraordinary interaction method between computers and humans (Meng X et al. 2021 & Mahapatra et al. 2022).

A technology called the Motor Imagery Brain-Computer Interface (MI-BCI) modifies how the brain reacts to physical actions. EEG is used by systems based on MI-BCI to track every mental activity. Additionally, MI-BCI has the ability to accurately identify concepts of physical activities including the movement of the hands, legs, and feet, among others (Tibrewal et al. 2022). The applications of EEG are widely utilized for the applications of BCI because of its high temporal resolutions and simple implementation. The BCI system helps in recording the activities of the brain and processes it to an artificial limb or wheelchair. The applications of BCI play a major role in psychology, entertainment purposes, gaming, smart buildings, neuromarketing, etc. (Idowu et al. 2021 & Cherloo et al. 2021).

There are many conventional Machine Learning (ML) algorithms such as random forest, decision tree, support vector machine etc., employed in the classification of motor imagery of EEG (Yang et al. 2021). The conventional ML algorithms in MI-BCI are based on handcrafted features (Musallamet al. 2021), and the features are extracted using a general method such as Common Spatial Pattern (CSP). Generally, employing CSP on the time series of EEG consists of a bandpass filter with a wide range of frequencies (for example, 5-45 Hz) (Malan et al. 2021). The CSP approach extracts the multi-channel EEG signals for classifying the signals on


**Fig 1.** EEG signal classification using IPFOA

the right and left hands. Moreover, the CSP extracts feature spatially from each sub-band. For large amounts of sub-bands, an autoencoder (AE) is utilized to lessen the dimensionality of more sub-band features in CSP (Tang et al. 2021). However, there are certain limitations in CSP such as the sensitivity of noise and the issue of overfitting (Darvish et al. 2021 & Keerthi et al. 2021), which are overcome in the proposed method. Additionally, employing CSP in the field of ML to obtain accurate results remains as a challenging task (Mirzaei et al. 2021). So, Deep Learning methods (DL) are considered state-of-the-art in various fields such as recognition of speech, natural language processing, and so on. Since the DL method has developed a graphic processing unit, it attributes high-level feature extraction and classification methods (Phadikar et al. 2023). This reason made researchers and scholars initiate the MI-BCI-based classification of EEG signals. DL consists of a Deep Neural Network (DNN) that is broadly utilized in the field of recognition of MI using EEG signals (Hermosilla et al. 2021).

The motor imagery signals consist of undesired noises and blurred signals which diminish the accuracy of the classification model (Dagdevir et al. 2023). This research undertakes the aforementioned issues which rely on the existing algorithms as motivation, and focuses on developing a DNN classifier that efficiently classifies the EEG signals based on MI-BCI.

The main contributions of the paper are highlighted below:

The proposed IPFOA achieved better performance in selecting the features from the sensorimotor rhythms to classify the EEG signals.

The CNN-based CSP is utilized in extracting the desired features according to the needs of the user.

From various optimization methods, the proposed IPFOA achieved the highest metrics during evaluation and attained comparatively better results than the existing methods.

The remaining paper is organized as follows; the proposed method is discussed in Section 2. The results and discussion are provided in Section 3, and the conclusion of this research is presented in Section 4.

## 2. Proposed Method

The EEG classification involves four steps namely, pre-processing or artifacts removel using the 6th order Butterworth filter following the feature extraction using the hybrid CNN and CSP methods and later feature selection using the proposed method, and finally classification using the stacked autoencoder. It is represented in the **Figure 1.**

### 2.1 Dataset

This paper is based on the information obtained from the BCI Competition IV 2a dataset. The dataset is introduced by the Graz University, Austria, which is utilized to compute the efficiency of the motor imagery EEG encoding method (Ang et al. 2012). The collection of the dataset includes EEG data from nine healthy people who perform four different motor imagery tasks, including moving the tongue, both feet, both hands, and both right and left hands, and 22 EEG electrodes are used to record the data at a sample rate of 250 Hz. The signals are band-pass filtered between 0.5 Hz and 100 Hz, and notch filtered at 50 Hz. For each subject, 2 periods on various days are verified, with 288 trials in each session. Each trial has a sampling interval of 3s, yielding 750 sample points overall. Each trial is used as a sample in this research and is considered in a 2D matrix measuring 22 by 750. The 750 columns of a sample correspond to the EEG data from the 750 sample points, and 22 rows of a sample correspond to the recorded signal from 22 electrodes. The original input signal is represented in **Figure 2** and the input signal obtained from one electrode is represented in **Figure 3**. The EEG signal of the person's left hand (**Figure 4**), right hand (**Figure 5**), foot and tongue are represented.

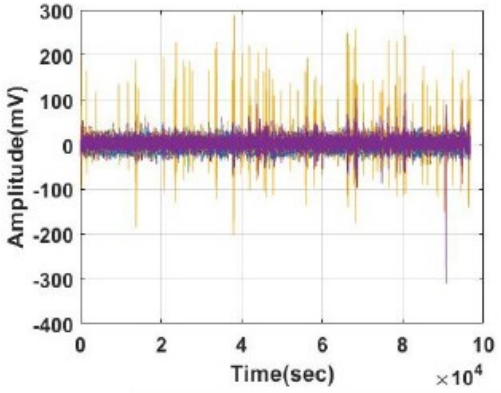



**Fig 2.** Input Signal

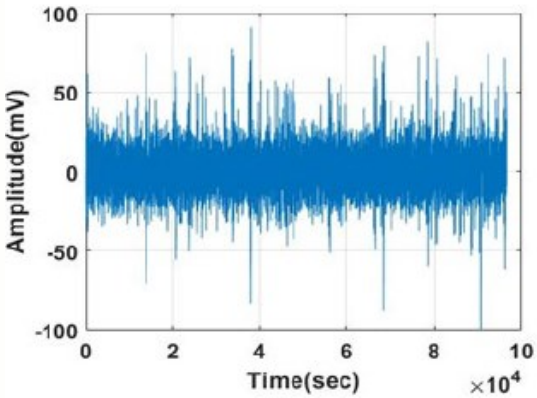



**Fig 3.** Input Signal for 1-Electrode

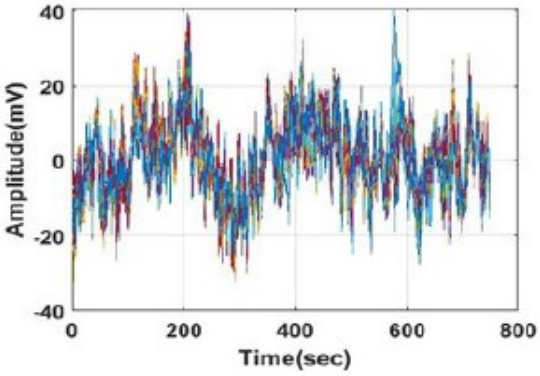



**Fig 4.** Class-1 for all Electrodes (Left hand)

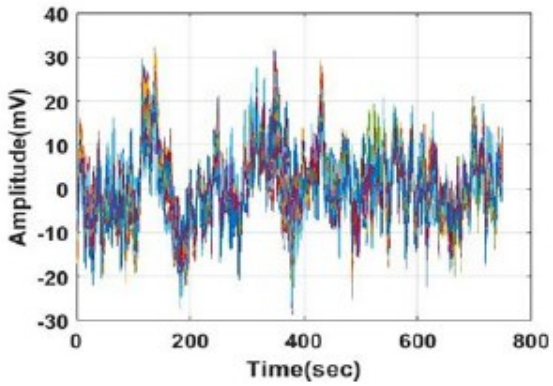



**Fig 5.** Class-2 for all Electrodes (Right hand)

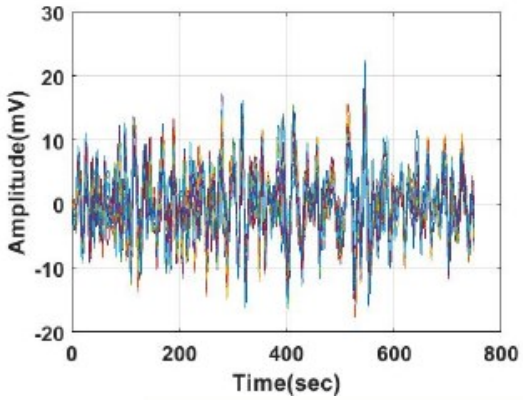



**Fig 6.** Pre-processing Class-1 (Left hand)

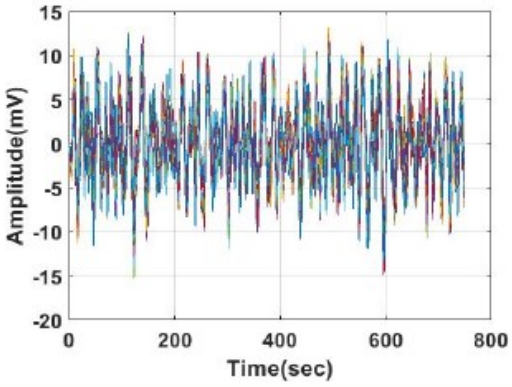



**Fig 7.** Pre-processing Class-2 (Right hand)

## 2.2 Pre-Processing

In the stage of pre-processing, the sliding window is used based on the Longest Consecutive Repetition (LCR). LCR (Gaur et al. 2021 & Saideepthi et al. 2023) is known as the longest repeating sequence of an element, for instance, in 5655566, '555' is considered the longest repeating sequence. Here, a total of 5 overlapping windows are used which are represented as $TT1$ to $TT5$. The first window starts ($TT1$) from the time period of 1.0 s, the second window ($TT2$) starts from 1.1 to 2.1, and the final window ($TT5$) starts from 4.4 to 5.5. The data is band-passed using the butter worth filter of order 6 and the signals of EEG occur in the rhythms of $\mu$ and $\beta$. The rhythms of $\mu$ and $\beta$ are attained from the signal of EEG to indicate the information regarding motor imagery. The equation of the Butterworth filter is represented in Eq. (1). The pre-processed signals of the left hand and right hand are represented in **Figures 6** and **7**, respectively.

$$H(z) = \frac{B(z)}{A(z)} = \frac{b(1) + b(2)z^{-1} + \cdots + b(n+1)z^{-n}}{a(1) + a(2)z^{-1} + \cdots + a(n+1)z^{-n}} \quad (1)$$

## 2.3 Feature Extraction

In this paper, the extraction of features takes place using a Convolutional Neural Network (CNN) and Common Spatial Pattern (CSP) (Rusnac et al. 2022, Rusnac et al. 2022, Sun et al. 2022 & Jin et al. 2021). The first layer in the CNN designs is an input layer, while the subsequent layers are made up of convolutional, max-pooling, and/or average pooling layers. To prevent overfitting, dropout layers are included in between the several convolutional layers. A thick layer, a fully linked layer and a soft maximum layer make up the final classification layer. The deeper layers of CNN are created to learn higher-level features by breaking down the input into complex structures, whereas the first layer of CNN learns fundamental features through filtering. The learned features from the previous layers are combined in the final, fully linked layer for classification. CNN uses backward propagation methods to reduce error, and forward propagation algorithms to discover the output.

The major goal of CSP is to investigate the spatial filters that reduce the variance of class and help in band pass filtering the multichannel signals of EEG. The method based on CSP attains optimum discrimination for the tasks based on MI-BCI on band power features. The optimized function of the CSP method using spatial filter P($x$) is represented in Eq. (2).

$$P(x) = \frac{x^T p_1^T p_1 x}{x^T p_2^T p_2 x} = \frac{x^T c_1 x}{x^T c_2 x} \quad (2)$$

Where the transpose of the matrix is represented as $T$, the training data of the matrix is represented as $P_i$ where the rows are considered spatial points and the channels are represented as Columns. For the particular class $i$, the spatial covariance matrix is represented as $C_i$. There are many ways to solve the problems in optimization, here the technique is based on visualizing the function $P(x)$ when the rescaling occurs in the filter $x$. The $P(kx) = P(x)$, where $k$ is denoted as a constant value of rescaling the filter. So, minimizing $P(x)$ is similar with $x^T c_1 x$to the constraint $x^T c_2 x$=1. The constrained optimization function using the method of Lagrange multiplier is represented in Eq. (3),

$$L(\beta, x) = x^T C_1 x - \beta(x^T C_2 x - 1) \quad (3)$$

Where the Lagrange multiplier is denoted as $\beta$, a derivative of $L$ concerning $x$ is denoted as 0. The filter $x$ minimizing $L$ is denoted using the Eq. (4),

$$\frac{\partial L}{\partial x} = 2x^T C_1 x - 2\beta x^T C_2 x = 0 \quad (4)$$

Where $\partial$ is the derivative related to the Lagrange multiplier. From this, the value of $\beta$ is attained as$C_2^{-1} C_1 x$, so the value of the Eigenvector is represented in Eq. (5),

$$z = C_2^{-1} C_1 x \quad (5)$$

Where, $z$ is the feature of CSP which is attained by linearly transforming the signals of EEG.

## 2.4 Feature Selection

The PFOA (Halder et al. 2022) is a swarm-intelligence-based meta-heuristic technique which was introduced in 2019. The PFA's computing process was motivated by the collective behaviour of any animal group in search of food. The animal population is divided into two groups, the leader and the followers, which imitates the leadership position and the followers to mathematically describe the algorithm. To find the ideal food zone, the leader is responsible for discovering new areas in the search area. The leader also leaves a footprint that aids the follower's upcoming

reorientation, and helps followers to find the path by following the footsteps. According to footsteps and perception, the followers follow the pathfinder. In the optimization of PFA, the part of followers and the leaders do not remain constant, it varies according to the search capability of individuals. So, the pathfinder may be a follower or leader. In the process of searching, the process of PFOA takes place in two stages. The first stage is the pathfinder phase, where the pathfinder is represented as $XX_{PP}$ and it is represented in Eq. (6),

$$X_p^{k+1} = X_p^k + 2 \times r_3 \times \left(X_p^k - X_p^{k-1}\right) + A, \quad \vec{r_3} \in \{R \text{ and } (1, D)\} \tag{6}$$

Where, $t$ represents the iteration counter and the dimensional vector set $D$ of the random numbers which are created uniformly within the range [0,1], and signified as $\{R \text{ and } (1,D)\}$. The $A$ is denoted as a parameter at each iteration and is represented in Eq. (7),

$$A = u_2 \times exp^{\left(\frac{-2k}{kmax}\right)}, \vec{u_2} \in \{R \text{ and } (1, D)\} \tag{7}$$

Where the maximum number of iterations is represented as $k$max.

In the second stage, every individual follower $ii$ varies its position which is represented in Eq. (8) as follows,

$$X_1^{k+1} = X_1^k + 2 \times R_1 \times \left(X_f^k - X_1^k\right) + R_2 \times \left(X_p^k - X_1^k\right) + \varepsilon, i \geq 2 \tag{8}$$

$$\varepsilon = \left(1 - \frac{k}{kmax}\right) \times u_1 \times Dist_{i,j}, \vec{u_1} \in \{R \text{ and } (1, D)\} \text{and } Dist_{i,j} = \left\|X_1^k - X_f^k\right\| \tag{9}$$

In Eq. (9), $X_f^k$ represents the adjacent individual of the $i^{th}$ follower and $R_1=\alpha.r_1$, $R_2=\alpha.r_2$. where $\vec{r_1}$ and $\vec{r_2} \in \{R \text{ and } (1,D)\}$. The coefficient of interface with a neighbor is denoted as $\alpha$ and the coefficient of attraction in a random distance is denoted as $\beta$. The value of $\varepsilon$ is computed using the formula given in Eq. (9). To perform a multi-dimensional search, the values of $A$ and $\varepsilon$ must be limited to a specific range.

## 2.5 Improvised Path Finder Optimization Algorithm (IPFOA)

The IPFOA is obtained by extending three features in existing PFA, the addition of features includes, (i) adding the mechanism of the wizard in PFA, (ii) adding the operators, (iii) improving the mechanism of mutation to branch out the individuals and increase the ability of algorithms. The following sections briefly explain the modifications to PFA.

### 2.5.1 Adding the mechanism of the wizard in PFA

In this stage, the pathfinder is considered a guide where the guide can be replaced using the iteration of the algorithm. During every iteration, the individual who has a small fitness value in the group is considered a guide. The guide provides useful information from the obtained experience to the followers. It is represented in Eq. (10).

$$X_1^{new} = X_1^{old} + rand.\left(X_1^{guide} - TF \times Mean\right) \tag{10}$$

In Eq. (10), $X_1^{new}$ gets updated using $X_1^{old}$ and the best individual in the group is represented as $X_1^{guide}$. The mean value of the follower present in the iteration is represented as $M$eanand the random value that lies between the range [0,1] is represented as $r$and. The passing weight is denoted as $T$F, where $T$F=$r$ound[1+$r$and(0,1){2−1}. When the fitness value of $X_1^{new}$ has better value than $X_1^{old}$, then $X_1^{old}$ is replaced with $X_1^{new}$. The guide offers better solutions for all the people and helps improve the speed during converging.

### 2.5.2 Adding the Operators

In this stage, the stage of the follower is improved and the addition of two operators takes place to regulate the search direction. Here, the value of accept operator $P$ is 0.8 and $e$tac is 20, and the values of $P$ and $e$tacare determined from numerous experiments. When the value of $P$ is greater than the provided random value, the position of the follower is organized consequently using the Eq. (11). In contrast, the addition of an exchange operator takes place and the position of the follower is organized using the Eq. (11-13).

$$X_1^{k+1} = X_1^k + 2(rand - 0.5) \times (X_1^{guide} - X_1^k) \quad (11)$$

$$X_1^{k+1} = X_1^k + \gamma \times R_1 \times \left(X_f^k - X_1^k\right) + \gamma \times R_2 \times \left(X_p^k - X_1^k\right) + \varepsilon, i \geq 2 \quad (12)$$

$$\beta = \begin{cases} 2r^{\frac{1}{etac+1}} r \leq 0.5 \\ \dfrac{1}{2(1-r)}^{\frac{1}{etac+1}} \quad else \end{cases} \quad (13)$$

Where the number of iterations are represented as $k$ and $X_1^{k+1}$ is upgraded as $X_1^k$. The best individual in the iteration is represented as $X_1^{guide}$and the random value lies between the range [0,1], represented as $r$and. $R1$ and $R2$ are denoted as random variables in the range of [0,1]. The neighboring individual is represented as $X_f^k$ and the interaction degree with neighboring individuals or leaders is represented as $\gamma$, which is determined by random variables $r$ and $e$tac. The best individual initiates its direction by learning and communicating with the leader and the neighboring individuals. This stage helps in increasing the mining capability of the algorithm.

### 2.5.3 Mechanism of Mutation

The mutation stage is employed in the search phase to maintain equilibrium between exploration and exploitation. The process of mutation is used to create a new population with fewer individual differences. People who are aware of their surroundings do searches at random inside the designated search zone, while others who appear to be less capable do not perform searches. To offer individuals an opportunity to increase the diversity of the population and keep the algorithm from reaching local optimality, the mutation mechanism is implemented. The introduction of the mutation process increases population diversity, gives an opportunity for novices to learn from more seasoned individuals, and keeps the algorithm from reaching local optimality. It uses a specific proportion of the information from multiple individuals as the amount of individual disturbance. The mutation probability ensures that the majority of people have the chance to learn ($p$cR). If the value of $p$cR is smaller than 0.5, it will be harder for weaker people to learn because the experimental effect won't be as noticeable. When the value of $p$cRis greater than 0.8, the mutation takes place among all individuals in the group, so the value of $p$cRis allotted between 0.5 and 0.8. The mathematical representation of the mutation mechanism is represented in Eq. (14),

$$X_1^{k+1} = X_m^k + pcR.(X_n^k - X_1^k) \quad (14)$$

Where $m$ and $n$ denotes the mutation which takes place in $m$th individual and $n$th individual of the population respectively.

## 2.6 Classification

In this stage, the classification of EEG signals takes place using Stacked Autoencoder (SAE) (Vafaei et al. 2023 & Li Q et al. 2022). An autoencoder (AE) is a network that consists of one input, hidden and output layer. The total neurons present in the output layer equals the total neurons in the layer of input. At the time of training, the input $xx$ is mapped with the hidden layer to give $yy$ as provided in Eq. (15), as a hidden output. Afterward, $yy$ gets mapped with the layer at the output to provide the value of $z$ (provided in Eq. (16). The mentioned steps are represented as,

$$y = f(W_y x + b_y) \quad (15)$$

$$z = f(W_z y + b_z) \quad (16)$$

Where $f$ is known as a function of activation and is represented in Eq. (17),

$$f(a) = 1/(1 + \exp(-a)) \quad (17)$$

$Wz$ and $Wy$ are known as the weight from the layers of output and weights from the layer of input to the hidden layer respectively. The bias values of hidden and output layers are represented as $b_y$ and $b_z$ respectively. When the value of $W_y$ is fixed as$W_z^{`}$, the weights are tied and this helps in obtaining the parameters of the model by reducing the cost function, which is represented in Eq. (18),

$$arg\ min_{W,b_y,b_z}[E(x,z)] \quad (18)$$

The reconstruction error is represented as $E(x,z)$, when the network is trained to reconstruct the values of output to the values of input. The model parameters can be reorganized by the Eq. (19-21),

$$W = W - \eta \frac{\delta E(x,z)}{\delta W} \quad (19)$$

$$b_y = b_y - \eta \frac{\delta E(x,z)}{\delta b_y} \quad (20)$$

$$b_z = b_z - \eta \frac{\delta E(x,z)}{\delta b_z} \quad (21)$$

The learning rate of the algorithm is denoted as η. After the training of AE, the features in the hidden layers are utilized for classification in the layer of the deep network of AE, which is referred to as a Stacked Auto Encoder (SAE). The SAE is composed of layers of input and output along with numerous AEs. Every individual layer of the AE (Zeng et al. 2023 & Wang et al. 2022) is trained separately and the output of the hidden layer in AE is utilized as input for the next layer in a deep neural network.

## 3. Result and Discussion

The results and discussion of the IPFOA are explained in this section. The design and implementation of the DL model using IPFOA are done by MATLAB R2020b version 9.9. The system utilized in classification involves specifications of an i5 processor with 6GB RAM. The major part of IPFOA is to attain precise and accurate classification of EEG motor imagery signals.

### 3.1 Performance Analysis

The performance of the classifier with feature selection in DNN is compared with the existing Multiclass Support Vector Machine (MSVM), K-Nearest Neighbor (KNN), Random Forest (RF), Decision Tree (DT), Neural Network (NN) and Deep Neural Network (DNN). The results are validated with 5-fold cross-validation with 20% testing and 80% training. The performance analysis of the classifier with feature selection is represented in **Table 1**.

**Table 1**. Performance Analysis of Various Classifiers with Feature Selection

| Classifier | Acc* (%) | Sen* (%) | Spe* (%) | F1* (%) | Pre* (%) |
|---|---|---|---|---|---|
| MSVM | 94.23 | 92.79 | 96.91 | 94.93 | 97.17 |
| KNN | 95.28 | 91.96 | 93.79 | 92.82 | 93.71 |
| RF | 95.55 | 96.49 | 93.96 | 95.24 | 94.03 |
| DT | 96.96 | 92.80 | 95.74 | 92.03 | 91.28 |
| NN | 96.96 | 94.14 | 93.39 | 94.09 | 94.05 |
| DNN | 98.40 | 99.00 | 98.91 | 99.46 | 99.92 |

*Acc*: Accuracy, Sen*: Sensitivity, Spe*: Specificity, F1*: F1-score, Pre*: Precision*

**Table 2.** Performance Analysis of Various Classifiers without Feature Selection

| Classifier | Acc* (%) | Sen* (%) | Spe* (%) | F1* (%) | Pre* (%) |
|---|---|---|---|---|---|
| MSVM | 90.87 | 89.14 | 88.60 | 91.09 | 87.61 |
| KNN | 92.69 | 91.77 | 93.71 | 91.26 | 90.75 |
| RF | 90.32 | 88.80 | 89.75 | 90.01 | 91.26 |
| DT | 93.95 | 92.19 | 90.28 | 91.85 | 91.51 |
| NN | 94.03 | 90.49 | 93.68 | 91.09 | 91.70 |
| DNN | 95.44 | 93.45 | 94.66 | 93.67 | 93.89 |

*Acc*: Accuracy, Sen*: Sensitivity, Spe*: Specificity, F1*: F1-score, Pre*: Precision*

From **Table 1** and **Table 2**. it is concluded that the performance of the Stacked Auto-Encoder (SAE), a DNN classifier provides better results in accuracy (95.44%), sensitivity (93.45%), specificity (94.66%), F1 score (93.67%), and precision (93.89%). **Table 3** represents the performance of the optimization algorithms such as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Fruit fly Optimization algorithm (FOA), Path Finder Optimization (PFOA), and the proposed IPFOA. From the results of Table 3, it is concluded that the proposed IPFOA has the highest evaluation metrics when compared with other optimization algorithms for feature selection

**Table 3.** Evaluation of Optimization Algorithms

| Algorithms | Acc* (%) | Sen* (%) | Spe* (%) | F1* (%) | Pre* (%) |
|---|---|---|---|---|---|
| PSO | 87.66 | 90.27 | 89.93 | 90.49 | 90.72 |
| ACO | 92.35 | 91.70 | 90.99 | 91.34 | 90.99 |
| FOA | 94.78 | 93.36 | 92.16 | 93.73 | 94.10 |
| PFOA | 94.96 | 94.25 | 93.25 | 94.91 | 95.57 |
| Proposed (IPFOA) | 98.40 | 99.00 | 98.91 | 99.46 | 99.92 |

*Acc*: Accuracy, Sen*: Sensitivity, Spe*: Specificity, F1*: F1-score, Pre*: Precision*

### 3.2 Comparative Analysis

The comparative analysis of the method used for feature extraction of the hybrid (combination of CNN-CSP) with the existing research is provided in this section. The existing t-distributed Stochastic Neighbor Embedding (t-SNE) and MFTL-LDA are compared to calculate the performance of the hybrid method. The hybrid method is utilized in extracting the features from the raw data which can reduce the redundancy of data and increase the speed of learning. The comparative analysis of the hybrid feature extraction method is provided in **Table 4**.

**Table 4.** Comparative Table

| Methods | Acc* (%) | Sen* (%) | Spe* (%) | F1* (%) | Pre* (%) |
|---|---|---|---|---|---|
| t-distributed Stochastic Neighbor Embedding (t-SNE) (Song et al. 2021) | 91.67 | 71.67 | 95 | 78.33 | 61.67 |
| MFTL-LDA (Liang et al. 2020) | 88.50 | 83.50 | 86 | 83 | 68.55 |
| IDEOA (Vishwesh et al. 2023) | 97.34 | NA | 98.01 | 98.72 | 98.54 |
| Hybrid (proposed) | 98.40 | 99.00 | 98.91 | 99.46 | 99.92 |

*Acc*: Accuracy, Sen*: Sensitivity, Spe*: Specificity, F1*: F1-score, Pre*: Precision*

### 3.3 Practical Applications and Implications of the Proposed Method

Using the Improved Path Finder Optimisation (IPFO) technique, BCI (Brain-Computer Interface) systems can be greatly improved. Mainly it can be used in medical rehabilitation as stroke recovery where BCIs can be used to interpret neural signals and control robotic arms, stroke patients can benefit from assistance in relearning motor abilities. To improve these systems' precision and responsiveness, IPFO can optimize signal processing. By using IPFO to increase the speed and accuracy of communication devices, BCIs can help people with severe speech or movement disabilities communicate more effectively. By optimizing the path finding in brain signal interpretation, IPFOA can greatly increase the signal-to-noise ratio in BCIs, resulting in more precise and dependable BCIs.

## 4. Conclusion

In this paper, the improved pathfinder optimization algorithm is proposed to efficiently classify the motor imagery EEG signals into four classes Left hand, right hand, feet, and tongue. The classification of EEG signals is performed using a DNN classifier called a stacked autoencoder. Here, the DL model is utilized to classify the Electroencephalogram (EEG) signals because they offer a method for the automatic extraction of spatiotemporal information from the signals. The features are extracted using the CNN-based CSP which efficiently extracts the features from the raw data which can reduce the redundancy of data. The extracted features undergo the process of feature selection using IPFOA. The proposed IPFOA performs better than previous methods, according to experimental findings when the data quantity from the subject is very small. It performs better than the older techniques even if the new training data is adequate. The IPFOA offers a new concept for regulating the features of EEG in the classification task of motor imagery signals. In the future, the proposed algorithm will be applied to the real-time applications of BCI.

## ACKNOWLEDGEMENTS

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis, and comparison, preparing and editing a draft, visualization have been done by first. The supervision, review of work, and project administration have been done by the second author.

## REFERENCES

Ang KK, Chin ZY, Wang C, Guan C, Zhang H. (2012). Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Frontiers in Neuroscience. 2012; 6:39. https://doi.org/10.3389/fnins.2012.00039.

Cherloo MN, Amiri HK, Daliri MR. (2021). Ensemble Regularized Common Spatio-Spectral Pattern (ensemble RCSSP) model for motor imagery-based EEG signal classification. Computers in Biology and Medicine. 2021; 135:104546. DOI: 10.1016/j.compbiomed.2021.104546.

Darvish ghanbar K, Rezaii TY, Farzamnia, A, Saad I. (2021). Correlation-based common spatial pattern (CCSP): A novel extension of CSP for classification of motor imagery signal. Plos one. 2021; 16(3):e0248511. https://doi.org/10.1371/journal.pone.0248511.

Dagdevir E, Tokmakci M. (2023) Determination of effective signal processing stages for brain computer interfaceon BCI competition IV data set 2b: a review study. IETE Journal of Research. 2023; 69(6):3144-55. DOI: 10.17485/IJST/v16i6.2076.

Gaur P, Gupta H, Chowdhury A, McCreadie K, Pachori RB, Wang H. (2021). A sliding window common spatial pattern for enhancing motor imagery classification in EEG-BCI. IEEE Transactions on Instrumentation and Measurement. 2021; 70:4002709. DOI:10.1109/TIM.2021.3051996.

Halder S, Dora BK, Bhat S. (2022). An Enhanced Pathfinder Algorithm based MCSA for rotor breakage detection of induction motor. Journal of Computational Science. 2022; 64:101870. DOI:10.1016/j.jocs.2022.101870.

Hermosilla DM, Codorniú RT, Baracaldo RL, Zamora RS, Rodriguez DD, Albuerne YL, Álvarez JRN. (2021). Shallow convolutional network excel for classifying motor imagery EEG in BCI applications. IEEE Access. 2021; 9:98275-86. DOI:10.1109/ACCESS.2021.3091399.

Idowu OP, Adelopo O, Ilesanmi AE, Li X, Samuel OW, Fang P, Li G. (2021). Neuro-evolutionary approach for optimal selection of EEG channels in motor imagery based BCI application. Biomedical Signal Processing and Control. 2021; 68:102621. https://doi.org/10.1016/j.bspc.2021.102621.

Jin J, Xiao R, Daly I, Miao Y, Wang X, Cichocki A. (2020). Internal feature selection method of CSP based on L1-norm and Dempster–Shafer theory. IEEE transactions on neural networks and learning systems. 2021; 32(11):4814-25. DOI: 10.1109/TNNLS. 2020.3015505.

Keerthi KK, Soman KP. (2021). CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image. Biomedical Engineering Letters. 2021; 11(3):235-47. DOI: 10.1007/s13534-021-00190-z.

Liang Y, Ma Y. (2020). Calibrating EEG features in motor imagery classification tasks with a small amount of current data using multisource fusion transfer learning. Biomedical Signal Processing and Control. 2020; 62:102101.

Li Q, Liu Y, Shang Y, Zhang Q, Yan F. (2022). Deep sparse autoencoder and recursive neural network for EEG emotion recognition. Entropy. 2022; 24(9):1187. https://doi.org/10.3390/e24091187.

Meng X, Qiu S, Wan S, Cheng K, Cui, L. (2021). A motor imagery EEG signal classification algorithm based on recurrence plot convolution neural network. Pattern Recognition Letters. 2021; 146:134-41. DOI: 10.1016/j.patrec.2021.03.023.

Mahapatra AK, Panda N, Pattanayak BK. (2022). An Improved Pathfinder Algorithm (ASDR-PFA) based on Adaptation of Search Dimensional Ratio for solvingconstrained optimization problems and optimal Feature Selection. 2022. DOI:10.21203/rs.3.rs-2115041/v1.

Musallam YK, AlFassam NI, Ghulam M, Amin SU, Alsulaiman M, Abdul W, Altaheri H, Bencherif MA, Algabri M. (2021). Electroencephalography-based motor imagery classification using temporal convolutional network fusion. Biomedical Signal Processing and Control. 2021; 69:102826. https://doi.org/10.3390/bioengineering9070323.

Malan NS, Sharma S. (2021). Time window and frequency band optimization using regularized neighbourhood component analysis for Multi-View Motor Imagery EEG classification. Biomedical Signal Processing and Control. 2021; 67:102550. DOI:10.1016/j.bspc.2021.102550.

Mirzaei S, Ghasemi P. (2021). EEG motor imagery classification using dynamic connectivity patterns and convolutional autoencoder. Biomedical Signal Processing and Control. 2021; 68:102584. https://doi.org/10.1016/j.bspc.2021.102584.

Phadikar S, Sinha N, Ghosh R. (2023). Unsupervised feature extraction with autoencoders for EEG based multiclass motor imagery BCI. Expert Systems with Applications. 2023; 213A:118901.https://doi.org/10.1016/j.eswa.2022.118901.

Rusnac AL, Grigore O. (2022). CNN architectures and feature extraction methods for EEG imaginary speech recognition. Sensors. 2022; 22(13):4679. https://doi.org/10.3390/s22134679.

Rusnac AL, Grigore O. (2022). Imaginary Speech Recognition Using a Convolutional Network with Long-Short Memory. Applied Sciences. 2022; 12(22):11873. https://doi.org/10.3390/app122211873.

Saideepthi P, Chowdhury A, Gaur P, Pachori RB. (2023). Sliding Window along with EEGNet based Prediction of EEG Motor Imagery. IEEE Sensors Journal. 2023; 23(15):17703-17713. https://dx.doi.org/10.1109/JSEN.2023.3270281.

Sun J, Wei M, Luo N, Li Z, Wang H. (2022). Euler common spatial patterns for EEG classification. Medical & Biological Engineering & Computing. 2022; 60(3):753-67. DOI: 10.1007/s11517-021-02488-7.

Song Y, Jia X, Yang L, Xie L. (2021). Transformer-based spatial-temporal feature learning for EEG decoding. arXiv preprint arXiv:2106.11170. 2021.

Tibrewal N, Leeuwis N, Alimardani M. (2022). Classification of motor imagery EEG using deep learning increases performance in inefficient BCI users. Plos one. 2022; 17(7):e0268880. https://doi.org/10.1371/journal.pone.0268880.

Tang R, Li Z, Xie X. (2021). Motor imagery EEG signal classification using upper triangle filter bank auto-encode method. Biomedical Signal Processing and Control. 2021; 68:102608. DOI:10.1016/j.bspc.2021.102608.

Vafaei E, Rahatabad FN, Setarehdan SK, Azadfallah P. (2023). Extracting a Novel Emotional EEG Topographic Map Based on a Stacked Autoencoder Network. Journal of Healthcare Engineering. 2023; 2023:9223599. https://doi.org/10.1155/2023/9223599.

Vishwesh J, Raviraj P. (2023). Improved Differential Evolution with Stacked Auto Encoder for EEG Motor Imagery Classification. Indian Journal of Science and Technology. 2023; 16(6):391-400. DOI: 10.17485/IJST/v16i6.2076.

Wang Y, Qiu S, Li D, Du, C., Lu, B.L., He, H. (2022). Multi-modal domain adaptation variational autoencoder for eeg-based emotion recognition. IEEE/CAA Journal of AutomaticaSinica. 2022; 9(9):1612-26. DOI: 10.1109/JAS.2022.105515.

Xiao X, Fang Y. (2021). Motor imagery EEG signal recognition using deep convolution neural network. Frontiers in Neuroscience. 2021; 15:655599. https://doi.org/10.3389/fnins.2021.655599.

Yang L, Song Y, Ma K, Xie L. (2021). Motor imagery EEG decoding method based on a discriminative feature learning strategy. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2021; 29:368-79. DOI: 10.1109/TNSRE.2021.3051958.

Yu, X., Aziz, MZ, Sadiq, MT, Fan Z, Xiao G. (2021). A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems. IEEE Transactions on Instrumentation and Measurement. 2021; 70:1006612. DOI: 10.1109/TIM.2021.3069026.

Zeng H, Xia N, Tao M, Pan D, Zheng H, Wang C, Xu F, Zakaria W, Dai G. (2023). DCAE: A dual conditional autoencoder framework for the reconstruction from EEG into image. Biomedical Signal Processing and Control. 2023; 81:104440. https://doi.org/10.1016/j.bspc.2022.104440.

## AUTHOR BIOGRAPHIES

**Vishwesh Jayashekar** is currently working as Associate Professor in the Department of Computer Science & Engineering at GSSS Institute of Engineering & Technology for Women, Mysore. He obtained Bachelor Degree in Computer Science & Engineering from PES College of Engineering, Mandya, Karnataka State, India in 2009. M. Tech from University of Mysore, Mysuru, Karnataka State, India in 2011. Ph D from VTU, Belagavi in 2023. He has 12 years of teaching experience. currently he published 4 international papers.

**Raviraj Pandian** completed his doctorate degree in Computer Science and Engineering in the area of Image Processing. He holds a position of Director-IQAC and Professor in the Department of Computer Science and Engineering at GSSS Institute of Engineering and Technology for Women, Mysore, Karnataka. He has 19 years of teaching and research experience. He has published more than 92 papers in international journals and conferences. At present he is guiding the Ph.D. research scholars in the areas of image processing, pervasive and cloud computing, and bio-inspired algorithms and robotics. He has received the project grant Rs.5 Lakhs from the VGST, Govt. of Karnataka for the "Underwater Robotic Fish for Surveillance and Pollution monitoring". He has received the awards and recognitions such as 'Rhastriya Gaurav Award-2015', 'Shri P.K. das Memorial Best Faculty Award-2012', 'Young Achiever Award-2016'. He can be contacted at email: raviraj@gsss.edu.in.

**Rajashekar Mallajamma Basavarajegowda** is currently working as Associate Professor in the Department of Computer Science & Engineering at GSSS Institute of Engineering & Technology for Women, Mysore. He obtained Bachelor Degree in Computer Science & Engineering from Coorg Institute of Technology, Ponnampet, Karnataka State, India in 2008. M. Tech from National Institute of Engineering, Mysore, Karnataka State, India. Ph D from VTU, Belagavi in 2023.He has 16 years of teaching experience. currently he published 6 international papers.

IJoSI

## 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 figures 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 Figures.* 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. Figure (or table) numbering and its subsequent caption must be below the figure (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:<br>Position:<br>School/Company: |
| **Order Information** | **I would like to order ___ copy(ies) of the** *International Journal of Systematic Innovation*:<br>**Period Start: 1st/ 2nd half ____, Year:____(Starting 2010)**<br>**Period End : 1st/ 2nd half ____, Year:**<br>**Price:**<br>**Institutions: US $150 (yearly) / NT 4,500 (In Taiwan only)**<br>**Individuals: US $50 (yearly) / NT 1500 (In Taiwan only)**<br>(Local postage included. International postage extra)<br>**E-mail to**: IJoSI@systematic-innovation.org or fax: +886-3-572-3210<br>Air mail desired ☐ (If checked, we will quote the additional cost for your consent) |
| **Total amount due** | **US$** |

**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.