Vol. 10 No. 1
February 2026

Opportunity Identification & Problem Solving

INTERNATIONAL JOURNAL OF SYSTEMATIC INNOVATION

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202602_10(1)

ACCSCIENCE PUBLISHING

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202602_10(1)

The International Journal of Systematic Innovation

Publisher:

The Society of Systematic Innovation

Editorial Office:

International Journal of Systematic Innovation Editorial Office

- *Address:* 9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619
- *Email:* ijosi.office@accscience.sg
- *Website:* https://ijosi.org/journal/IJOSI
- *Tel:* +65 8182 1586

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202602_10(1)

INTERNATIONAL JOURNAL OF SYSTEMATIC INNOVATION

CONTENTS

FEBRUARY 2026 VOLUME 10 ISSUE 1

FULL PAPERS

ACCSCIENCE PUBLISHING

International Journal of Systematic Innovation

ARTICLE

Natural language processing-based development of artificial intelligence-driven autonomous socially assistive robots

Murat Şimşek[1], and Demiral Akbar[2*]

[1]Department of Artificial Intelligence Engineering, Faculty of Engineering, OSTİM Technical University, Ankara, Turkey

[2]Department of Mechanical Engineering, Faculty of Engineering, OSTİM Technical University, Ankara, Turkey

Abstract

This study focuses on addressing the growing need for localized language support in socially assistive robots (SARs) due to rising labor costs and the limitations of human labor in developed countries. The research aims to develop a Turkish natural language processing (NLP) module to enhance SARs' social interaction capabilities and integration into smart living spaces. By leveraging advanced machine learning models, specifically XLM-RoBERTa Large (deepset/xlm-roberta-large-squad2), the study evaluated cross-lingual transfer learning for Turkish question answering, addressing specific linguistic challenges, including agglutinative morphology and vowel harmony. The model was evaluated on the Turkish Question-Answering Dataset (TQuAD 2.0) with 2,520 validation examples, achieving 79.37% F1-score and 56.67% exact match score. The research established a methodological framework connecting adaptive NLP design principles with control systems theory, demonstrating how concepts from adaptive fuzzy control and robust neural adaptive control inform the development of more stable and reliable NLP systems for SAR applications. These outcomes highlight the potential of cross-lingual NLP models for SAR applications in Turkish-speaking environments. The research contributes to the field by: (i) evaluating cross-lingual transfer learning for Turkish SAR applications, (ii) demonstrating the effectiveness of XLM-RoBERTa for low-resource language adaptation, (iii) establishing a framework that connects adaptive NLP design with control systems theory for enhanced robustness, and (iv) identifying real-world SAE applications in healthcare, smart homes, and industrial settings. Future work will focus on integrating this NLP module with speech recognition and synthesis components for complete voice-interactive SAR systems.

Keywords: Socially assistive robot; Natural language processing; Large language model; Question-answering; XLM-RoBERTa; Turkish Question-Answering Dataset; Cross-lingual transfer learning; Adaptive control systems

***Corresponding author:**
Demiral Akbar
(demiral.akbar@ostimteknik.edu.tr)

Citation: Şimşek, M. & Akbar, D. (2026). Natural language processing-based development of artificial intelligence-driven autonomous socially assistive robots. *Int J Systematic Innovation*. 10(1):1-10.
https://doi.org/10.6977/IJoSI.202602_10(1).0001

Received: December 21, 2024

Revised: December 13, 2025

Accepted: January 5, 2026

Published online: February 13, 2026

1. Introduction

There is a growing interest in smart homes and robotic technology in various aspects of life. However, limited studies have investigated the social interactions of socially assistive robots (SARs) with humans and their integration into smart homes and living spaces. With recent advancements in robotics, mobile service robots are becoming a part of our daily lives, accompanying us in human-interactive environments, such as airports, hospitals, schools, and offices. Initially, robots entered living spaces with the capability to perform relatively limited and repetitive tasks without the need for human interaction, such as cleaning (Seo *et al.*, 2019), surveillance (Nayyar *et al.*, 2018), and cooking. Recently, companion robots, such as Astra Amazon, Alexa on wheels, Jibo (Breazeal, 2017), and Temi (Hung *et al.*, 2021), have begun to spread in living spaces to assist with entertainment and comfort.

García-Haro *et al.* (2021) previously proposed 10 main classes of service robots, identifying SARs as an emerging area in the service robotics market. SARs are physical or digital entities capable of socially interacting and communicating with humans, designed to offer companionship and/or assistance through interactions that resemble human interactions (Toscano *et al.*, 2022). Researchers have examined how interacting with an embodied robot, compared to disembodied speech agents, enhances users' confidence and enjoyment (Maj & Zarzycki, 2019; Spitale *et al.*, 2020). The effectiveness of SARs depends on their ability to address both utilitarian and hedonic factors. According to Van der Heijden (2004), hedonic systems aim to provide self-fulfilling value to users, in contrast to utilitarian systems that aim to provide instrumental value.

1.1. Previous research in question-answering systems across languages

Question-answering (QA) systems have been developed for various languages using different approaches. For English, the Stanford Question-Answering Dataset (SQuAD) (Rajpurkar *et al.*, 2016) established benchmarks that led to significant advances in reading comprehension models. The Bidirectional Encoder Representations from Transformers (BERT) model achieved state-of-the-art results on English QA tasks by leveraging bidirectional context understanding (Devlin *et al.*, 2019). For Chinese, researchers developed the Chinese Machine Reading Comprehension (CMRC) dataset and fine-tuned multilingual models, achieving F1 scores above 80% (Cui *et al.*, 2019). Arabic QA systems have been developed using similar transfer learning approaches, with AraBERT achieving competitive results despite the language's complex morphology (Alkesaiberi *et al.*, 2024). The Kazakh language QA system demonstrated the applicability of BERT-based models for low-resource Turkic languages, achieving high Bilingual Evaluation Understudy (BLEU) scores (Mukanova *et al.*, 2024).

Cross-lingual transfer learning has emerged as a powerful approach for low-resource languages. XLM-RoBERTa (Conneau *et al.*, 2020), trained on 100 languages including Turkish, has shown remarkable zero-shot and fine-tuned performance across diverse linguistic tasks. Our approach leverages this cross-lingual capability by utilizing the deepset/xlm-roberta-large-squad2 model, which combines the multilingual representations of XLM-RoBERTa with task-specific fine-tuning on SQuAD 2.0. This enables effective transfer to Turkish QA without requiring extensive Turkish-specific pre-training.

Our approach follows similar methodologies to these previous works, specifically: (i) utilizing cross-lingual transfer learning from pre-trained multilingual language models, (ii) evaluating on native Turkish QA datasets (TQuAD 2.0), and (iii) adapting the models for the specific linguistic characteristics of the target language. Our work aims to integrate the natural language processing (NLP) module into a SAR platform for real-time voice-based interaction capabilities.

1.2. Linguistic characteristics of the Turkish language

Turkish presents unique challenges for NLP systems, distinguishing this research from work in other languages. Turkish is an agglutinative language, meaning that words are formed by adding suffixes to root words, creating complex word forms that can express entire sentences in other languages with a single word. For example, the word "*evlerinizden*" (from your houses) contains four morphemes: *ev* (house) + *ler* (plural) + *iniz* (your) + *den* (from). This characteristic significantly increases vocabulary size and data sparsity challenges (Eryiğit, 2014).

Additionally, Turkish exhibits vowel harmony, where suffixes must harmonize with the vowels in the root word according to front/back and rounded/unrounded distinctions. Turkish also has a relatively free word order (subject–object–verb as default but flexible), which affects how context must be processed in transformer models. These linguistic features require specialized tokenization strategies and model adaptations. The XLM-RoBERTa model we employed uses SentencePiece tokenization trained on multilingual data, effectively handling Turkish morphological complexity through subword segmentation (Conneau *et al.*, 2020).

1.3. Research contributions

This research makes the following specific contributions to the field:

(i) We evaluate cross-lingual transfer learning using XLM-RoBERTa Large for Turkish QA in the context of SAR applications.
(ii) We demonstrate the effectiveness of the deepset/xlm-roberta-large-squad2 model on the TQuAD 2.0 benchmark, achieving 79.37% F1-score and 56.67% exact match in a zero-shot transfer learning setting without Turkish-specific fine-tuning.
(iii) We establish a methodological framework connecting adaptive NLP design principles with control systems theory, showing how concepts from adaptive fuzzy control (Boulkroune *et al.*, 2017), robust neural adaptive control (Hayakawa *et al.*, 2008), and adaptive backstepping control (Zouari *et al.*, 2013) can inform the development of more stable and reliable NLP systems for robotics applications.
(iv) We provide detailed error analysis and sample predictions to understand model behavior on Turkish linguistic patterns and identify real-world applications across healthcare, smart home, and industrial domains.

2. Materials and methods

We employed the Transformer architecture, a novel neural network design introduced by Vaswani *et al.* (2017), that relies on self-attention mechanisms to process sequences of data. Unlike traditional sequential models such as recurrent neural networks, Transformers process data in parallel, significantly reducing training times and improving the handling of long-distance dependencies.

2.1. XLM-RoBERTa Large model architecture

For our Turkish QA system, we employed the XLM-RoBERTa Large model (deepset/xlm-roberta-large-squad2), a state-of-the-art cross-lingual language model. XLM-RoBERTa (Conneau *et al.*, 2020) is trained on 2.5 TB of filtered CommonCrawl data spanning 100 languages, including Turkish. The model architecture consists of 24 transformer layers with 1,024 hidden dimensions and 16 attention heads, totaling approximately 560 million parameters.

The key advantages of XLM-RoBERTa for our application include: (i) Cross-lingual transfer capability—the model learns language-agnostic representations that transfer effectively to Turkish without language-specific pre-training; (ii) Robust tokenization using SentencePiece with a 250,000 token vocabulary that effectively handles Turkish agglutinative morphology; (iii) Pre-training on diverse multilingual data that captures cross-linguistic patterns beneficial for understanding Turkish syntax and semantics.

The specific model variant (deepset/xlm-roberta-large-squad2) has been additionally fine-tuned on the SQuAD 2.0 English dataset, which includes both answerable and unanswerable questions. This task-specific fine-tuning provides the model with strong QA capabilities that transfer to Turkish through cross-lingual representations.

2.2. TQuAD 2.0

We evaluated our model on TQuAD 2.0, a native Turkish reading comprehension dataset. Unlike machine-translated datasets, TQuAD 2.0 contains originally Turkish content derived from Turkish Wikipedia articles, ensuring natural language patterns and culturally relevant content.

The dataset comprises 14,221 training examples and 2,520 validation examples across diverse topics, including history, geography, science, sports, and culture. Each example consists of a context paragraph, a question, and one or more answer spans extracted from the context. The dataset statistics are presented in Table 1.

Table 1. TQuAD 2.0 statistics

Dataset property	Value
Dataset name	TQuAD 2.0 (erdometo/tquad2)
Training examples	14,221
Validation examples	2,520
Language	Turkish (native)
Source	Turkish Wikipedia

Below are the dataset examples:

(i) Example 1: Context: "*Rollo'nun gelişinden önce popülasyonları Picardy'den veya Franklar olarak nitelendirilen Île-de-France'den farklı değildi. Daha önceki Viking yerleşimcileri 880'lerde gelmeye başlamıştı...*" Question: "*Kim geldiğinde orijinal viking yerleşimcilerine ortak bir kimlik vermiştir*?" Answer: "*Rollo.*"
(ii) Example 2: Context: "*Akdeniz'e ulaşmak için en önde gelen iki Norman ailesi, Hauteville'den Tancred ve Drengot ailesinin soyundan geliyordu...*" Question: "*III. Henry tarafından asillendirilmiş liderin adı nedir*?" Answer: "*Drogo.*"
(iii) Example 3: Context: "*Akdeniz'e ulaşmak için*

en önde gelen iki Norman ailesi, Hauteville'den Tancred ve Drengot ailesinin soyundan geliyordu..." Question: "*Melfi Kontu kimdi*?" Answer: "*William Iron Arm*."

2.3. Model configuration and hyperparameters

The XLM-RoBERTa Large model was configured using the hyperparameters detailed in Table 2 for evaluation.

Table 2. Model configuration parameters

Hyperparameter	Value
Maximum sequence length	384 tokens
Document stride	128 tokens
Batch size	4
N-best size	20
Max answer length	30 tokens
Precision	BF16 (mixed precision)
GPU	NVIDIA A100-SXM4-40GB
Transformers version	4.57.3

The tokenization process used SentencePiece with a vocabulary size of 250,000 tokens. For contexts longer than the maximum sequence length, overlapping chunks were created using a document stride of 128 tokens to ensure answer spans at chunk boundaries could be correctly identified.

2.4. Model evaluation approach

Important Clarification: This research evaluates the pre-trained deepset/xlm-roberta-large-squad2 model in a *zero-shot transfer learning* setting on TQuAD 2.0. No additional fine-tuning was performed on Turkish data. This approach was chosen to:

(i) Assess cross-lingual transfer capability: Evaluate how well multilingual pre-training on 100 languages (including Turkish) enables the model to perform Turkish QA without language-specific fine-tuning.
(ii) Resource efficiency: Demonstrate that effective Turkish NLP capabilities can be achieved without the extensive computational resources required for full model fine-tuning, making this approach more accessible for SAR development.
(iii) Baseline establishment: Provide baseline performance metrics that future research can improve through Turkish-specific fine-tuning or domain adaptation.

The model leverages two levels of pre-training:

(i) Multilingual pre-training: XLM-RoBERTa base training on 2.5 TB of data across 100 languages
(ii) Task-specific pre-training: Fine-tuning on English SQuAD 2.0 for QA capabilities

The evaluation on TQuAD 2.0 demonstrated the model's ability to transfer both multilingual representations and task-specific knowledge to Turkish QA contexts. The TQuAD 2.0 validation set contains only answerable questions (2,520 examples with answers, 0 unanswerable questions). This differs from SQuAD 2.0, which includes unanswerable questions to test model discrimination capabilities.

3. Results and discussion

This section presents the evaluation results of the XLM-RoBERTa Large model on TQuAD 2.0. The evaluation was conducted using the HuggingFace Transformers library with the pre-trained deepset/xlm-roberta-large-squad2 model in a zero-shot transfer learning setting.

3.1. Model evaluation results

The XLM-RoBERTa Large model was evaluated on the TQuAD 2.0 validation set comprising 2,520 examples. The evaluation results are presented in Table 3.

Table 3. TQuAD 2.0 evaluation results

Metric	Score
Total examples evaluated	2,520
Exact match	56.67%
F1 score	79.37%

The F1 score of 79.37% indicates strong performance in extracting relevant answer spans from Turkish context passages. The exact match score of 56.67% reflects the strict nature of this metric, which requires character-perfect matching with ground truth answers. The gap between F1 and EM scores is typical for extractive QA tasks and indicates that the model often identifies partially correct answers that overlap significantly with the ground truth.

3.2. Answer detection performance

Since TQuAD 2.0 contains only answerable questions (all 2,520 validation examples have answers), the results of the confusion matrix for answer detection are presented in Table 4. Table 5 describes answer detection metrics.

The model successfully identified that all questions in the TQuAD 2.0 validation set are answerable, achieving 100% accuracy in answer detection. This result indicates that the model's null-answer threshold is appropriately calibrated for this dataset.

Table 4. Confusion matrix for answer detection

Category	Prediction: No answer	Prediction: Has answer
Actual: No answer	0 (TN)	0 (FP)
Actual: Has answer	0 (FN)	2,520 (TP)

Abbreviations: FN: False negative; FP: False positive; TN: True negative; TP: True positive.

Table 5. Answer detection metrics

Detection metric	Score (%)
Precision	100.00
Recall	100.00
F1 (detection)	100.00
Accuracy	100.00

3.3. Sample prediction analysis

Analysis of model predictions revealed both successful extractions and error patterns. The following examples illustrate the model's behavior:

(a) Correct predictions (exact match=1):

- (i) Q: "*Panthers savunması kaç sayı bırakmıştır*?" Gold: "308." Predicted: "308."
- (ii) Q: "*Jared Allen'ın kaç tane kariyer sack edişi vardır*?" Gold: "136." Predicted: "136."
- (iii) Q: "*Pro Bowl için kaç tane Panthers savunma oyuncusu seçilmiştir*?" Gold: "*dört*." Predicted: "*dört*."
- (iv) Q: "*Hangi oyuncu sezonun en çok top kapmasına sahiptir*?" Gold: "*Kurt Coleman*." Predicted: "*Kurt Coleman*."

(b) Incorrect predictions (exact match=0):

- (i) Q: "*Josh Norman kaç tane top çalmıştır*?" Gold: "*dört*." Predicted: "88." The model extracted a related number from context, but not the correct answer.
- (ii) Q: "*Bu sezon takımdaki en fazla sack etmeyi kim kaydetmiştir*?" Gold: "*Kawann Short*." Predicted: "*Jared Allen*." Entity confusion between players mentioned in the same context.
- (iii) Q: "*Panthers savunması 2015 yılında kaç top çalma ile kayda geçmiştir*?" Gold: "24." Predicted: "(118)." Extracted wrong numerical value from context.

3.4. Error analysis

Analysis of incorrect predictions revealed the following error categories:

- (i) Numerical confusion (approximately 35% of errors): The model sometimes extracted incorrect numbers when multiple numerical values appeared in the context. This was particularly common for statistical questions about sports data, where many numbers were mentioned.
- (ii) Entity confusion (approximately 30% of errors): When multiple named entities of the same type (e.g., player names and location names) appeared in the context, the model occasionally selected the wrong entity. This suggests challenges in resolving coreference and understanding fine-grained semantic distinctions.
- (iii) Partial match errors (approximately 25% of errors): The model extracted answers that overlapped with the ground truth but included extra tokens or missed some tokens. For example, extracting "(118)" instead of "118."
- (iv) Context misalignment (approximately 10% of errors): In some cases, the model extracted text from a different part of the context that seemed superficially related to the question but did not contain the correct answer.

3.5. Theoretical framework: Adaptive control principles for robust natural language processing in socially assistive robots

The design of robust NLP modules for SARs shares fundamental conceptual parallels with adaptive control theory. Both domains address the problem of maintaining stable and reliable system performance under uncertainty, nonlinearities, and time-varying operating conditions. In control systems, these challenges arise from model uncertainties, external disturbances, and unmodeled dynamics; in NLP-driven robotic interaction, they emerge from linguistic ambiguity, morphological variability, contextual uncertainty, and user-dependent interaction dynamics.

3.5.1. Adaptive and robust control analogies in natural language processing systems

Adaptive control strategies, such as adaptive fuzzy control and robust output-feedback control, have been extensively studied for uncertain nonlinear dynamical systems and chaotic synchronization problems (Boulkroune *et al.*, 2017). These methods rely on real-time parameter adaptation and feedback mechanisms to guarantee boundedness and stability in the presence of nonlinear input characteristics and unknown system dynamics.

An analogous requirement exists in SAR-oriented NLP systems. The linguistic input to the NLP module is inherently uncertain and nonstationary, particularly for agglutinative languages such as Turkish. In this context, the self-attention mechanism of transformer-based models can be interpreted as a form of adaptive gain scheduling, where attention weights are dynamically adjusted based on the relevance and uncertainty of linguistic tokens. This adaptive allocation of representational capacity mirrors the role of adaptive control laws in regulating system response under varying operating conditions.

Robust neural adaptive control approaches (Hayakawa *et al*., 2008) further strengthen this analogy. In such systems, neural networks compensate for modeling uncertainties while guaranteeing asymptotic stability. Similarly, the XLM-RoBERTa model leverages learned cross-lingual representations to compensate for linguistic variability and sparse data conditions, thereby maintaining stable QA performance across diverse Turkish sentence structures.

3.5.2. Backstepping control perspective on transformer layer hierarchies

Adaptive backstepping control provides a recursive methodology for stabilizing nonlinear systems by systematically designing control laws from lower-order subsystems to higher-order dynamics (Zouari *et al*., 2013). This layered design philosophy closely aligns with the hierarchical structure of transformer architectures used in NLP.

In transformer-based NLP models, lower layers predominantly capture lexical and syntactic features, while higher layers encode semantic and contextual relationships. This hierarchical abstraction process parallels the recursive stabilization steps in backstepping control, where intermediate virtual control signals are constructed to ensure global system stability. The validation of NLP models on structured benchmarks such as TQuAD 2.0 serves an analogous role to Lyapunov-based stability analysis in control theory, providing empirical guarantees of bounded and reliable system behavior.

The relevance of this analogy is further reinforced by adaptive backstepping control studies in flexible robotic manipulators (Zouari *et al*., 2013), where nonlinearities, elastic dynamics, and actuator uncertainties must be addressed simultaneously. SAR platforms integrating NLP modules face similar multi-domain uncertainties spanning perception, interaction, and actuation layers.

3.5.3. Optimal control and computational resource allocation

Nonlinear optimal control frameworks aim to achieve performance objectives while minimizing resource consumption under system constraints, as demonstrated in industrial applications such as gas compressors driven by induction motors (Rigatos *et al*., 2023). This optimization principle is directly applicable to NLP deployment in real-time robotic systems.

The use of mixed-precision computation (BF16) and constrained sequence lengths in this study represents an implicit optimal control strategy, balancing inference accuracy against computational latency and energy consumption. Such trade-offs are critical for SARs operating under real-time constraints and limited onboard computational resources.

3.5.4. Implications for control-informed socially assistive robot natural language processing design

The control-theoretic interpretation of NLP robustness suggests several design guidelines for future SAR systems:

(i) Closed-loop interaction: Incorporating user feedback mechanisms enables adaptive adjustment of dialogue strategies, analogous to feedback control loops.
(ii) Stability monitoring: Confidence-aware answer selection acts as a stability margin, preventing unreliable or unsafe system responses.
(iii) Adaptive tuning: Domain-specific fine-tuning parallels parameter adaptation in adaptive controllers, allowing performance improvement without loss of stability.
(iv) Fault tolerance: Detection of degraded NLP confidence can trigger fallback interaction modes, similar to fault-tolerant control architectures.

Overall, this framework demonstrates that established principles from adaptive and robust control theory—uncertainty compensation, hierarchical stabilization, and optimal resource allocation—provide a rigorous foundation for designing stable and reliable NLP-driven SARs operating in complex real-world environments.

4. Real-world applications and future directions

4.1. Real-world applications and adaptive socially assistive robot deployment

The integration of localized Turkish NLP capabilities into SARs extends beyond educational settings to numerous real-world applications where human–robot collaboration requires linguistic and cultural adaptation.

4.1.1. Healthcare applications

In healthcare environments, Turkish-speaking SARs can provide crucial support:

(i) Patient monitoring and companionship: SARs equipped with Turkish NLP can conduct routine health assessments through conversational interaction, asking patients about symptoms, medication adherence, and daily activities. The QA capability demonstrated in this research enables the robot to respond to patient queries about medication schedules, dietary restrictions, and appointment information.

(ii) Elderly care: Turkey's aging population (projected to reach 20.8% by 2040) creates significant demand for assistive technologies. SARs with culturally appropriate Turkish language support can provide companionship, remind elderly users to take medications, and alert caregivers to emergencies. The adaptive nature of our NLP system, informed by control theory principles (Section 3.5), enables the robot to adjust interaction complexity based on the user's cognitive status.

(iii) Rehabilitation support: SARs can guide patients through physical therapy exercises using natural Turkish instructions, answer questions about recovery procedures, and provide motivational feedback. The system's ability to handle Turkish's agglutinative morphology ensures understanding of varied phrasing from patients with different educational backgrounds or dialects.

4.1.2. Smart home integration

Turkish-speaking SARs can serve as intelligent interfaces for smart home ecosystems:

(i) Home automation control: Users can issue natural language commands in Turkish to control lighting, temperature, security systems, and appliances. The QA capability allows users to query system status ("What temperature is the living room?") and receive contextualized responses.

(ii) Energy management: SARs can provide real-time information about energy consumption, recommend optimization strategies, and answer questions about utility usage patterns—all in culturally appropriate Turkish communication styles.

(iii) Security and monitoring: Integration with smart home sensors enables SARs to answer questions about home security status, provide alerts about unusual activities, and serve as mobile surveillance platforms with natural language interfaces.

4.1.3. Industrial human–robot collaboration

In manufacturing and service industries, Turkish NLP-enabled SARs facilitate safer and more efficient operations:

(i) Workplace safety: SARs can monitor work environments, provide safety reminders in Turkish, answer worker questions about hazard protocols, and report safety violations. Drawing parallels to adaptive backstepping control for robot manipulators (Zouari *et al.*, 2013) and nonlinear optimal control for industrial systems (Rigatos *et al.*, 2023), these SARs must balance multiple objectives, including information delivery, physical safety, and operational efficiency.

(ii) Training and onboarding: New employees can interact with SARs to receive training information, ask procedural questions, and practice skills in a low-pressure environment—all conducted in their native Turkish language, which is especially valuable in Turkey's diverse industrial workforce.

(iii) Quality control: SARs equipped with vision systems and Turkish NLP can answer operator questions about quality standards, provide real-time feedback on product specifications, and explain defect classifications in an accessible language.

4.1.4. Adaptive feedback mechanisms

The control systems perspective (Section 3.5) informs how adaptive feedback from users can continuously improve SAR behavior across these applications:

(i) User interaction analytics: Similar to system identification in adaptive control, analyzing user interaction patterns (question types, rephrasing frequency, and interaction duration) provides data for model refinement. For example, if healthcare users frequently rephrase questions about medications, this indicates areas where model fine-tuning is needed.

(ii) Contextual adaptation: Just as adaptive fuzzy control adjusts parameters based on system state (Boulkroune *et al.*, 2017), SARs can adjust their language complexity, response verbosity, and interaction pace based on user profiles and real-time feedback. An elderly user might receive simpler, slower responses, while a technical professional receives detailed, technical information.

(iii) Performance monitoring: Implementing confidence-based feedback mechanisms (analogous to stability monitoring in robust neural adaptive control [Hayakawa *et al*., 2008]) allows SARs to request clarification when understanding is uncertain, improving interaction reliability.
(iv) Multi-domain transfer: Insights from our Turkish NLP development can be transferred to other low-resource Turkic languages (e.g., Azerbaijani, Uzbek, and Kazakh) through similar cross-lingual approaches, expanding SAR accessibility across Central Asia and the Turkic-speaking regions.

4.1.5. Connection to intelligent control and automation

The integration of localized NLP into SARs represents a convergence of intelligent control and human-centered automation:

(i) Hierarchical control architecture: Modern industrial automation employs hierarchical control structures where high-level planning, mid-level coordination, and low-level execution operate at different timescales. Similarly, SAR systems integrate strategic planning (task goals), tactical execution (dialogue management), and operational control (NLP inference), all coordinated through natural language interfaces.
(ii) Human-in-the-loop systems: Unlike fully autonomous systems, SARs exemplify human-in-the-loop intelligent automation where natural language serves as the primary human–machine interface. This aligns with Industry 4.0 principles, emphasizing collaborative rather than replacement automation.
(iii) Safety-critical adaptation: Drawing from safety-critical control systems literature, SARs in healthcare and industrial settings must guarantee bounded performance even under unusual inputs. The theoretical framework connecting adaptive control to NLP robustness (Section 3.5) provides design principles for ensuring fail-safe behavior.

4.1.6. Scalability and deployment challenges

Real-world SAR deployment faces challenges analogous to control systems implementation:

(i) Environmental variability: Like control systems operating under varying conditions, SARs must handle acoustic noise, diverse user accents, and non-standard linguistic inputs. Robust control strategies (Boulkroune *et al*., 2017; Hayakawa *et al*., 2008) suggest incorporating conservative design margins and fault detection mechanisms.
(ii) Long-term stability: Deployed systems must maintain performance over extended periods. Continual learning approaches, validated through principles from adaptive control theory, can enable SARs to improve without catastrophic forgetting or instability.
(iii) Integration complexity: SARs integrate multiple subsystems (perception, navigation, manipulation, and dialogue). This multi-component integration parallels multivariable control systems (Hayakawa *et al*., 2008), requiring careful interface design and coordination mechanisms to ensure system-wide stability and performance.

4.2. Future work

Future research directions include:

(i) Speech integration: Integrating the NLP module with Turkish automatic speech recognition (ASR) and text-to-speech (TTS) systems for complete voice-interactive capabilities.
(ii) Domain-specific fine-tuning: Fine-tuning the model on domain-specific Turkish datasets (medical, elderly care, and industrial) to improve performance in targeted applications.
(iii) Multimodal integration: Combining NLP with computer vision and gesture recognition for richer human–robot interaction.
(iv) Continual learning: Implementing adaptive learning mechanisms that allow the SAR to improve from user interactions while maintaining stability guarantees informed by control theory.
(v) Cross-lingual extension: Extending the approach to other Turkic languages to expand SAR accessibility across Central Asia.

5. Conclusion

This study evaluated cross-lingual transfer learning for Turkish QA in SAR applications. The XLM-RoBERTa Large model (deepset/xlm-roberta-large-squad2) was evaluated in a zero-shot setting on the TQuAD 2.0, achieving an F1 score of 79.37% and an exact match score of 56.67% on 2,520 validation examples without any Turkish-specific fine-tuning.

The results demonstrate the effectiveness of multilingual transformer models for Turkish QA tasks. The model successfully handled Turkish linguistic challenges, including agglutinative morphology and free word order, through its SentencePiece tokenization and cross-lingual representations learned from 100 languages.

Error analysis revealed that the primary challenges are numerical confusion (35%), entity confusion (30%), partial match errors (25%), and context misalignment (10%). These findings suggest directions for future improvement, including domain-specific fine-tuning and enhanced numerical reasoning capabilities.

Significantly, this research established a methodological framework connecting adaptive NLP design with control systems theory, demonstrating how principles from adaptive fuzzy control, robust neural adaptive control, and adaptive backstepping control inform the development of more stable and reliable NLP systems for robotics. This theoretical grounding provides design principles for implementing robust, fault-tolerant SARs that can maintain stable performance under uncertain and dynamic real-world conditions.

The identified applications in healthcare (patient monitoring, elderly care, and rehabilitation), smart homes (automation control, energy management, and security), and industrial settings (workplace safety, training, and quality control) demonstrate the broad potential impact of localized Turkish NLP in SAR systems. The adaptive feedback mechanisms informed by control theory enable continuous improvement while maintaining stability guarantees essential for safety-critical applications.

This work establishes a foundation for developing culturally and linguistically appropriate SARs for Turkish-speaking populations, with implications for other low-resource Turkic languages. Future work will focus on speech integration, domain-specific fine-tuning, and multimodal interaction capabilities to realize complete voice-interactive SAR systems for real-world deployment.

Acknowledgments

None.

Funding

This research was supported by internal funding from OSTİM Technical University Scientific Research Programs (Grant No. BAP202301).

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions

Conceptualization: Murat Şimşek, Demiral Akbar
Formal analysis: Murat Şimşek
Investigation: Murat Şimşek
Methodology: Murat Şimşek, Demiral Akbar
Writing–original draft: Murat Şimşek
Writing–review & editing: Demiral Akbar, Murat Şimşek

Availability of data

The data used in this study are publicly available. TQuAD 2.0 (erdometo/tquad2) and the pre-trained model (deepset/xlm-roberta-large-squad2) are available via Hugging Face. Any additional materials (e.g., evaluation configurations/scripts) can be provided by the corresponding author upon reasonable request.

References

Alkesaiberi, A., Alkhathlan, A., & Abdelali, A. (2024). Ara-CANINE: Character-based pre-trained language model for Arabic language understanding. *International Journal on Cybernetics & Informatics,* 13(2), 45-59.

https://doi.org/10.5121/ijci.2024.130204

Boulkroune, A., Hamel, S., Zouari, F., Boukabou, A., & Ibeas, A. (2017). Output-feedback controller based projective lag-synchronization of uncertain chaotic systems in the presence of input nonlinearities. *Mathematical Problems in Engineering,* 2017, Article 8045803.

https://doi.org/10.1155/2017/8045803

Breazeal, C. (2017). Social robots: From research to commercialization. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; March 6-9, 2017; Vienna, Austria.

https://doi.org/10.1145/2909824.3020258

Conneau, A., Khandelwal, K., Goyal, N., *et al.* (2020). Unsupervised cross-lingual representation learning at scale. In: Jurafsky, D., Chai, J., Schluter, N., & Tetreault, J., (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; July 5-10, 2020; Online. pp. 8440-8451.

https://doi.org/10.18653/v1/2020.acl-main.747

Cui, Y., Liu, T., Che, W., *et al.* (2019). A span-extraction dataset for Chinese machine reading comprehension. In: Inui, K., Jiang, J., Ng, V., & Wan, X. (Eds.), Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP); November 3-7, 2019; Hong Kong, China, pp. 5883–5889.

https://doi.org/10.18653/v1/D19-1600

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In: Burstein, J., Doran, C., & Solorio, T. (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; June 2-7, 2019; Minneapolis, MN, USA. Volume 1, pp. 4171-4186.

https://doi.org/10.18653/v1/N19-1423

Eryiğit, G. (2014). ITU Turkish NLP web service. In: Bentivogli, L., Koehn, P., & Manning, C. (Eds.), Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics; April 26-30, 2014, Gothenburg, Sweden. pp. 1-4. Association for Computational Linguistics.

https://doi.org/10.3115/v1/E14-2001

García-Haro, J. M., Oña, E. D., Hernandez-Vicen, J., Martinez, S., & Balaguer, C. (2021). Service robots in catering applications: A review and future challenges. *Electronics*, 10(1), Article 47.

https://doi.org/10.3390/electronics10010047

Hayakawa, T., Haddad, W. M., & Hovakimyan, N. (2008). Neural network adaptive control for a class of nonlinear uncertain dynamical systems with asymptotic stability guarantees. *IEEE Transactions on Neural Networks*, 19(1), 80–89.

https://doi.org/10.1109/TNN.2007.902704

Hung, C.-F., Lin, Y., Ciou, H.-J., Wang, W.-Y., & Chiang, H.-H. (2021). Foodtemi: The AI-oriented catering service robot. In: Proceedings of the 2021 IEEE International Conference on Consumer Electronics–Taiwan (ICCE-TW); September 15-17, 2021; Penghu, Taiwan. pp. 1–2.

https://doi.org/10.1109/ICCE-TW52618.2021.9603096

Maj, K., & Zarzycki, P. (2019). Meeting with social robots like the cat-cucumber meeting? An integrated model of human-robot first contact: Psychological perspective. *Paladyn, Journal of Behavioral Robotics*, 10(1), 454–465.

https://doi.org/10.1515/pjbr-2019-0026

Mukanova, A., Barlybayev, A., Nazyrova, A., Kussepova, L., Matkarimov, B., & Abdikalyk, G. (2024). Development of a geographical question-answering system in the Kazakh language. *IEEE Access*, 12, 105460–105469.

https://doi.org/10.1109/ACCESS.2024.3433426

Nayyar, A., Puri, V., Nguyen, N. G., & Le, D.-N. (2018). Smart surveillance robot for real-time monitoring and control system in environment and industrial applications. In: Bhateja, V., Nguyen, B. L., Nguyen, N. G., Satapathy, S. C., & Le, D.-N. (Eds.), *Information systems design and intelligent applications*. Berlin: Springer. pp. 229–243.

https://doi.org/10.1007/978-981-10-7512-4_23

Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). SQuAD: 100,000+ questions for machine comprehension of text. In: Su, J., Carreras, X., & Duh, K. (Eds.), Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing; November 1-4, 2016; Austin, Texas, USA. pp. 2383-2392.

https://doi.org/10.18653/v1/D16-1264

Rigatos, G., Abbaszadeh, M., Sari, B., Siano, P., Cuccurullo, G., & Zouari, F. (2023). Nonlinear optimal control for a gas compressor driven by an induction motor. *Results in Control and Optimization*, 11, 100226.

https://doi.org/10.1016/j.rico.2023.100226

Seo, T., Jeon, Y., Park, C., & Kim, J. (2019). Survey on glass and façade-cleaning robots: Climbing mechanisms, cleaning methods, and applications. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 6(2), 367-376.

https://doi.org/10.1007/s40684-019-00079-4

Spitale, M., Silleresi, S., Cosentino, G., Panzeri, F., & Garzotto, F. (2020). "Whom would you like to talk with?": Exploring conversational agents for children's linguistic assessment. In: Santos, O. C., Dillenbourg, P., & Börner, K. (Eds.), Proceedings of the Interaction Design and Children Conference; June 21-24, 2020; London, UK. pp. 262-272.

https://doi.org/10.1145/3392063.3394421

Toscano, E., Spitale, M., & Garzotto, F. (2022). Socially assistive robots in smart homes: Design factors that influence the user perception. In: Proceedings of the 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI); March 7-10, 2022; Sapporo, Japan. pp. 1075-1079.

https://doi.org/10.1109/HRI53351.2022.9889467

van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly*, 28(4), 695–704.

https://doi.org/10.2307/25148660

Vaswani, A., Shazeer, N., Parmar, N., *et al*. (2017). Attention is all you need. *arXiv*. Preprint.

https://doi.org/10.48550/arXiv.1706.03762

Zouari, F., Ben Saad, K., & Benrejeb, M. (2013). Adaptive backstepping control for a class of uncertain single input single output nonlinear systems. In: Proceedings of the 10th International Multi-Conference on Systems, Signals & Devices (SSD); May 6-10, 2022; Setif, Algeria. pp. 1-6.

https://doi.org/10.1109/SSD.2013.6564134

International Journal of Systematic Innovation

ARTICLE

Identifying mineral potentials related to geological structures using deep learning

Hadi Shahraki[1]*, and Mohsen Jami[2]

[1]Department of Computer Engineering, Faculty of Industry and Mining, University of Sistan and Baluchestan, Khash, Sistan and Baluchestan, Iran

[2]Department of Mining, Faculty of Industry and Mining, University of Sistan and Baluchestan, Khash, Sistan and Baluchestan, Iran

***Corresponding author:**
Hadi Shahraki
(Hadi_shahraki@eng.usb.ac.ir)

Citation: Shahraki H, Jami M. (2026). Identifying mineral potentials related to geological structures using deep learning. *Int J Systematic Innovation*. 10(1):11-18. https://doi.org/10.6977/IJoSI.202602_10(1).0002

Received: January 27, 2025

Revised: November 4, 2025

Accepted: January 5, 2026

Published online: February 13, 2026

Abstract

Artificial intelligence is increasingly being used as a powerful tool in various industries, including earth sciences. Geological structures play an undeniable role in the formation of mineral potentials. Investigating these structures relies on satellite imagery, together with expert interpretation, which can be a time-consuming process. Artificial intelligence can serve as a valuable tool to expedite this process and enhance the accuracy of mineral potential identification. This article presents a new model based on deep neural networks for identifying mineral potentials. The unique feature of the proposed method is the incorporation of morphological data alongside multispectral data to identify mineral potentials. To evaluate the effectiveness of the proposed method, advanced spaceborne thermal emission and reflection radiometer satellite images from a region in the southeast of Iran were utilized. The results demonstrate an improvement in the accuracy of the proposed method compared to similar approaches.

Keywords: Artificial intelligence; Deep learning; Deep neural network; Mineral potential identification; Machine learning

1. Introduction

Today, artificial intelligence, as a new technology, plays an important role in industries. Artificial intelligence can help improve processes and decisions by analyzing data to extract patterns and relationships that support decision-making. This can lead to an increase in productivity in various industries. Therefore, the use of artificial intelligence in various applications such as medicine (Chen *et al*., 2022), Internet of Things (Zhou *et al*., 2023), robotics (Soori *et al*., 2023), weather forecasting (Bochenek & Ustrnul, 2022), and intelligent vehicle systems (Suhaib *et al*., 2022) has grown significantly.

Artificial intelligence algorithms can also be used in geological and mineral mapping using multispectral satellite images (Shirmard, Farahbakhsh, Müller, *et al*., 2022). Oneof the most important advantages of using artificial intelligence in the preparation of geological maps is enhanced accuracy and speed of mapping (Sun *et al*., 2019), which has stimulated increased research in this area. Various algorithms have been used in geological and mineral mapping, including random forest (Kuhn *et al*., 2018; Radford *et al.*, 2018) support vector machine (SVM) (Cardoso-Fernandes *et al.*, 2020; Othman & Gloaguen, 2014), and neural networks (Latifovic *et al.*, 2018; Shirmard, Farahbakhsh, Heidari, *et al.*, 2022).

The SVM is one of the most widely used algorithms in the preparation of geological maps (Cardoso-Fernandes *et al.*, 2020; Othman & Gloaguen, 2014). Li *et al.* (2004) previously presented an automatic lithology classification method using advanced spaceborne thermal emission and reflection radiometer (ASTER) images using an SVM. Additionally, Bachri *et al.* (2019) reported a lithology classification method using this algorithm. Moreover, Bachri *et al.* (2019) and Cracknell & Reading (2014) compared different types of machine learning methods in the preparation of geological maps.

Neural networks are another widely used algorithm in the preparation of geological maps. Rigol-Sanchez *et al.* (2003) employed a neural network model to identify mineral potentials of gold in various areas in the southeast of Spain. Sergi *et al.* (1995) used neural networks to classify Landsat multispectral images. Moreover, Otele *et al.* (2021) used neural networks to prepare lithology maps.

The strong performance of deep learning or deep neural networks in various applications has led to the use of these methods in remote sensing processes (Ma *et al.*, 2019). Additionally, deep learning algorithms have been used in the preparation of geological maps. Latifovic *et al.* (2018) utilized deep neural networks to accelerate the production of geological maps and increase their accuracy. Shirmard, Farahbakhsh, Heidari, *et al.* (2022) compared the performance of convolutional neural networks (CNNs) and SVMs, demonstrating that CNNs produce superior results compared to other methods.

In this study, a new deep learning method to identify mineral potentials is presented. Considering the importance of morphological data along with the multispectral data to identify mineral potentials, both types of data are used in the proposed method. For this purpose, the important features are first extracted from the morphological data and then used with the multispectral data to identify mineral potentials. The organization of this article is as follows: Section 2 introduces the proposed algorithm. Experimental results are reported in Section 3, and finally, Section 4 concludes the paper.

2. Methodology

In the proposed algorithm, prediction was made based on two categories of data: information related to multispectral data and morphological data.

A multispectral image is a collection of several image layers of an area, each of which is obtained in a specific wavelength band. For example, the ASTER satellite has the ability to create images with 14 different bands, in which nine are optical, and the rest are thermal. Six out of nine light bands produced by the ASTER satellite are called short-wave bands, and three other bands are visible. Multispectral data describe a feature vector for a pixel in a multispectral image. In geology, multispectral data can be used to identify different types of rocks and minerals based on their unique spectral signatures. This information can be used to identify areas of interest for mineral exploration and to map the distribution of minerals.

Petrological and morphological information are used by an expert to identify mineral potentials in remote sensing. Morphological data could provide information about the surface of the Earth, focusing on its terrain, elevation, and structural features. In this study, a band of multispectral satellite images is considered as a simple morphological snapshot of the Earth's surface.

Considering the importance of morphological and multispectral data, both were used in the proposed algorithm to identify mineral potentials. Because the morphological input is higher-dimensional than the multispectral data, the important features in the image containing the morphological data were first extracted. Subsequently, the extracted morphological features along with the multispectral data were combined and classified. In the proposed method, these steps were implemented using deep learning algorithms.

A multilayer perceptron (MLP) neural network, one of the most popular deep neural networks, consists of several layers, such as the input layer, the output layer, and the hidden layers. In the proposed method, this model was used for data classification.

Convolutional neural networks are also among the most commonly used models in deep learning methods, designed for processing grid-like data, such as images. The neurons in convolutional networks employ convolution operations on the input data. Convolution operations apply filters to input data to extract features from it. This can lead to the recognition of local patterns in the images, which can be used to recognize patterns and identify targets in the images. CNNs have been widely used in different applications. For this purpose, different models of these neural networks have been introduced, the most important of which are LeNet-5, AlexNet, VGGNet, ResNet, and convolutional self-encoder networks.

Convolutional autoencoder networks are deep learning models based on convolutional networks, which have attracted considerable attention in recent years due to their ability to extract features from high-dimensional data, such as images. Convolutional autoencoder networks consist of two main components: an encoder and a decoder. In the first part of these networks, there is an encoder, which is responsible for compressing the input data, or in other words, extracting features from the input, while in the decoder part, the original input is reconstructed using the extracted features. In the proposed method, a convolutional autoencoder network was used to extract features from the morphological data.

In the proposed algorithm, to access the morphological data of each pixel, a neighborhood image with dimensions of 36 × 36 pixels was considered. To reduce the dimensions and extract important features from the image, a convolutional encoder was used. The details of the convolutional autoencoder network in the proposed algorithm are presented in Table 1. In the proposed algorithm, after training the encoder network, the encoder part was used to extract the features of the images. By using the convolutional encoder, the morphological data were reduced to 200 features. Then, the extracted morphological features along with the multispectral data for that pixel were put together and classified by an MLP network. In the output of the MLP network, each pixel was assigned to one of the desired classes. The parameters of the MLP network in the proposed algorithm are listed in Table 2. The general structure of the proposed algorithm is presented in Figure 1.

3. Results

In this study, multispectral satellite images of ASTER were used for mineral potential mapping. These images, presented in Shirmard, Farahbakhsh, Heidari, *et al.* (2022), are related to an area in the southeast of Iran, near Mirjaveh city, within the Sistan and Baluchestan province, with an area of about 66 km². The dimensions of the image are 513 × 577 pixels, with each pixel containing 225 m². In this image, nine different classes were considered, as shown in Figure 2. The details of this satellite image are mentioned in the study by Shirmard, Farahbakhsh, Heidari, *et al.* (2022). The pixels related to the nine classes were randomly divided into two parts: training and testing. A total of 70% of the pixels were considered as the training data and 30% as the test data. The training data were used in the learning process of the algorithms, which were then evaluated using the test data.

To clearly evaluate the performance of the proposed algorithm, it was compared with a conventional CNN model, which is one of the standard deep learning approaches widely used for lithological mapping. In addition, the results were also compared with two reference methods: the SVM, representing a traditional machine learning approach, and the MLP, representing a simple neural network model. This comparison followed the evaluation framework presented in Shirmard, Farahbakhsh, Heidari, *et al.* (2022). The distinctive advantage of the proposed model compared with the CNN approach lies in its ability to combine morphological features, extracted via a convolutional autoencoder, with multispectral information, allowing the network to capture both spectral and structural patterns.

We compared the proposed method with SVM, MLP, and CNN baselines reported by Shirmard, Farahbakhsh, Heidari, *et al.* (2022).

The results were compared using the overlap ratio of the sets (IOU), considering accuracy. The IOU criterion obtains the degree of overlap between the real pixels and the pixels predicted by the algorithms for class c, and calculates the accuracy of the percentage of pixels that are correctly segmented compared to the total pixels. These criteria are calculated as follows using Equations (1) and (2):

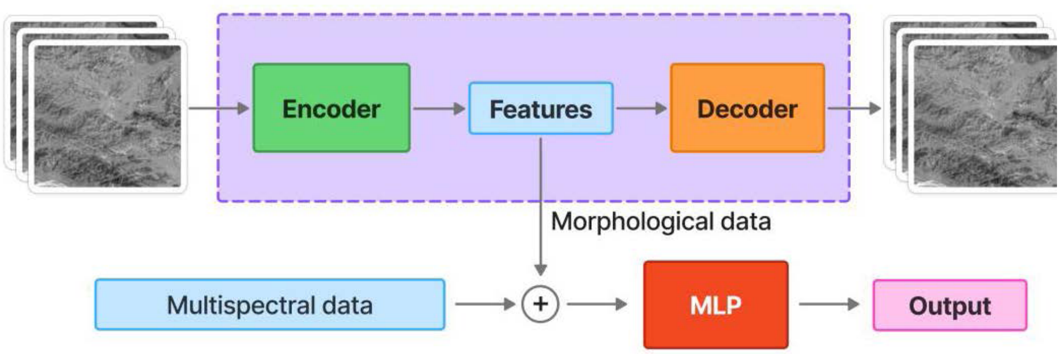


Figure 1. Outline of the proposed algorithm
Abbreviation: MLP: Multilayer perceptron.

Accuracy = The number of correct pixels/Total number of pixels (1)

IOU (c) = The number of pixels in the overlap area/The number of pixels in the union area (2)

Figure 3 shows the result of the proposed algorithm. The proposed algorithm was able to identify the mineral potentials presented in Figure 2 with high accuracy. Quantitative results obtained from the performance of the proposed algorithm, including IOU and accuracy criteria, are shown in Tables 3–5. Table 3 shows the accuracy value of different algorithms for the testing and training data. The proposed algorithm achieved the highest level of accuracy among the considered algorithms. Using the test data, the model achieved an accuracy 1.1 percentage points higher than that of the CNN algorithm.

For a more detailed comparison, the results of the IOU of different classes are shown in Tables 4 and 5. According to the results, the proposed algorithm exhibited superior performance to other algorithms. The proposed algorithm achieved its best performance in identifying class 8, with an IOU of 0.99 on the test data, significantly outperforming other comparable methods.

The obtained results demonstrate that integrating morphological and multispectral information significantly improved the accuracy of identifying mineral potential zones. This integration is crucial for practical mineral exploration applications, where field data acquisition is often time-consuming and expensive.

The proposed method can serve as an efficient tool for the preliminary identification of promising areas based on

Table 1. Details of the convolutional autoencoder network

Activation function	Window size	Number of filters	Type	Layer
ReLU	3 × 3	16	Conv2D	1
-	2 × 2	-	Max pooling 2D	2
ReLU	3 × 3	8	Conv2D	3
-	2 × 2	-	Max pooling 2D	4
ReLU	3 × 3	8	Conv2D	5
-	2 × 2	-	Max pooling 2D	6
ReLU	3 × 3	8	Conv2D	7
-	2 × 2	-	UpSampling2D	8
ReLU	3 × 3	8	Conv2D	9
-	2 × 2	-	UpSampling2D	10
ReLU	3 × 3	16	Conv2D	11
-	2 × 2	-	UpSampling2D	12
Sigmoid	3 × 3	1	Conv2D	13

Abbreviations: 2D: Two-dimensional; Conv: Convolutional; ReLU: Rectified linear unit.

Table 2. Hyperparameters of the multilayer perceptron neural network

Value	Hyperparameter
Rectified linear unit	Activation function
Adaptive moment estimation	Optimizer
15	Number of hidden layers
100	Number of hidden neurons
400	Epochs

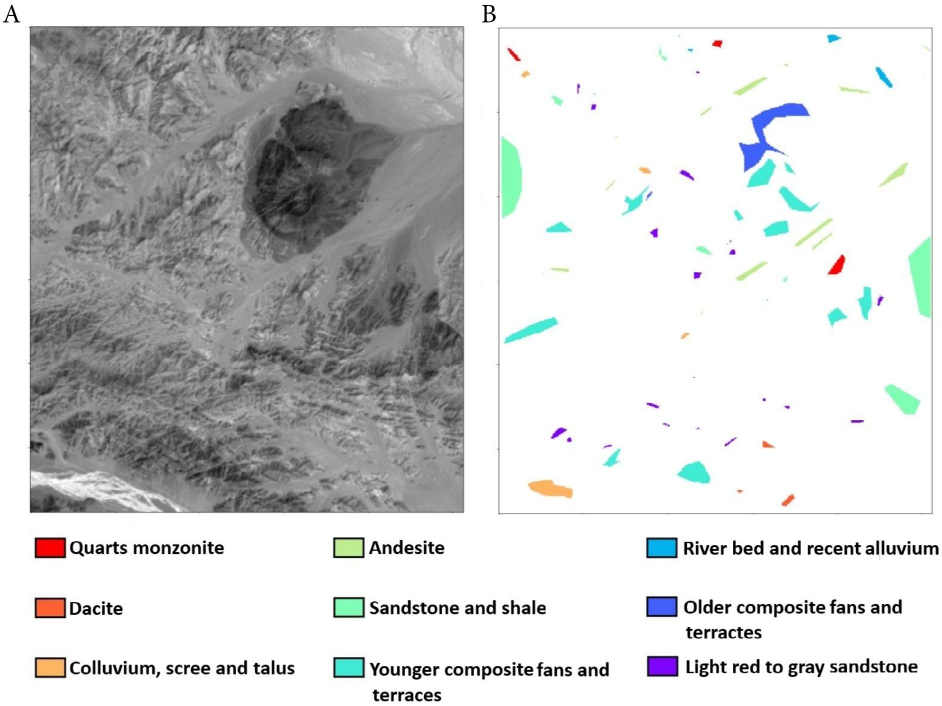


Figure 2. ASTER multispectral imagery and lithological information used for mineral potential analysis in the study area near Mirjaveh, southeastern Iran. (A) Satellite image related to the region in the southeast of Iran, along with the (B) mineral potentials in it. Images adapted from Shirmard, Farahbakhsh, Heidari, et al. (2022).

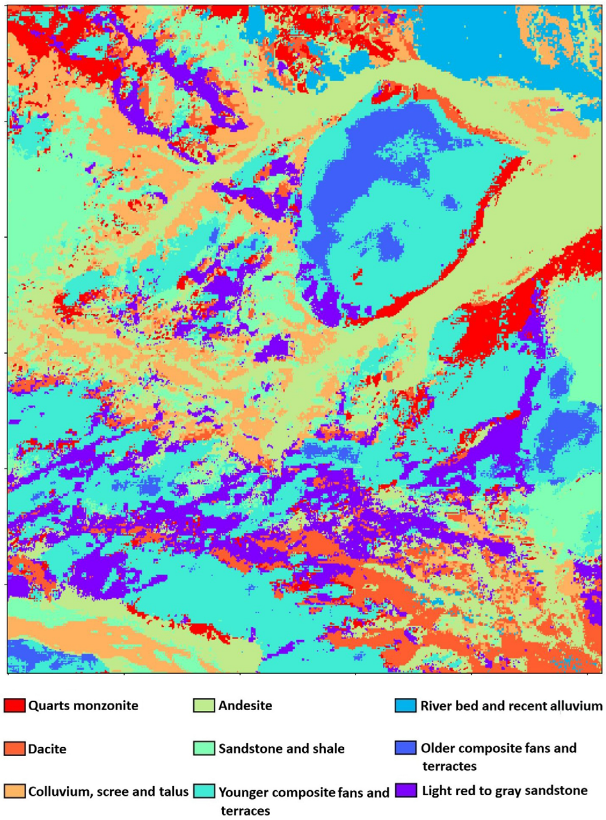


Figure 3. The result obtained using the proposed method to identify mineral potentials in Figure 2

Table 3. Comparison of the accuracy obtained from different methods

Proposed method	Convolutional neural network	Multilayer perceptron	Support vector machine	Data
99.4	98.2	96	94.6	Train
99.1	98	95.7	94.4	Test

Table 4. Comparison of the overlap ratio of the set values for different classes on the training data

Proposed method	Convolutional neural network	Multilayer perceptron	Support vector machine	Class
1.00	1.00	1.00	1.00	Class 1
0.99	0.99	0.94	0.93	Class 2
0.99	0.90	0.95	0.89	Class 3
0.99	0.96	0.92	0.90	Class 4
0.99	0.97	0.91	0.90	Class 5
0.97	0.94	0.88	0.84	Class 6
1.00	1.00	0.99	0.99	Class 7
0.99	0.77	0.79	0.60	Class 8
0.99	0.96	0.92	0.84	Class 9
0.99	0.94	0.92	0.87	Average

Table 5. Comparison of the overlap ratio of the set values for different classes on the test data

Proposed method	Convolutional neural network	Multilayer perceptron	Support vector machine	Class
0.99	1.00	0.99	0.99	Class 1
0.99	0.99	0.95	0.94	Class 2
0.95	0.85	0.97	0.90	Class 3
0.99	0.97	0.91	0.89	Class 4
0.98	0.97	0.91	0.89	Class 5
0.96	0.93	0.87	0.82	Class 6
0.99	1.00	0.99	1.00	Class 7
0.99	0.73	0.74	0.57	Class 8
0.99	0.97	0.91	0.83	Class 9
0.99	0.94	0.87	0.92	Average

satellite imagery, helping geologists concentrate their field investigations on the most prospective zones and thereby reducing both cost and time.

Moreover, the morphological feature extraction module, which captures structural and terrain characteristics of the Earth's surface, can be reused across different mineral exploration projects, as morphological patterns are not specific to any single mineral type. Consequently, for the exploration of other minerals or new geographic regions, only the multispectral data component of the model requires retraining using the new spectral datasets, while the morphological extractor can remain unchanged. This design makes the approach versatile and easily adaptable to various remote sensing-based geological mapping tasks.

4. Conclusion

This study presented a new method based on deep learning to identify mineral potentials and create lithology maps. The special feature of this algorithm is the use of morphological data in addition to the multispectral data. To use the morphological data, the features related to the neighboring image of each pixel were first extracted by an auto-encoder model, and then analyzed by an MLP neural network along with the multispectral data. The results obtained from the proposed algorithm were compared with those of three similar methods and outperformed the baseline methods. Future extensions could include uncertainty evaluation and testing the model across different geological regions to further assess its robustness and generalization capability.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that they have no conflicts of interest.

Author contributions

Conceptualization: All authors
Formal analysis: All authors
Investigation: All authors
Methodology: All authors
Visualization: All authors
Writing–original draft: All authors
Writing–review & editing: All authors

Availability of data

The dataset used in this study was originally published by Shirmard *et al.* (2022) and is publicly available at https://github.com/sydney-machine-learning/deeplearning-lithology.

References

Bachri, I., Hakdaoui, M., Raji, M., Teodoro, A. C., & Benbouziane, A. (2019). Machine Learning Algorithms for Automatic Lithological Mapping Using Remote Sensing Data: A Case Study from Souk Arbaa Sahel, Sidi Ifni Inlier, Western Anti-Atlas, Morocco. *ISPRS International Journal of Geo-Information*, 8(6), 248.

https://doi.org/10.3390/ijgi8060248

Bochenek, B., & Ustrnul, Z. (2022). Machine Learning in Weather Prediction and Climate Anayses—Applications and Perspectives. *Atmosphere*, 13(2), 180.

https://doi.org/10.3390/atmos13020180

Cardoso-Fernandes, J., Teodoro, A. C., Lima, A., & Roda-Robles, E. (2020). Semi-Automatization of Support Vector Machines to Map Lithium (Li) Bearing Pegmatites. *Remote Sensing*, 12(14), 2319.

https://doi.org/10.3390/rs12142319

Chen, X., Wang, X., Zhang, K., Fung, K.-M., Thai, T. C., Moore, K., Mannel, R. S., Liu, H., Zheng, B., & Qiu, Y. (2022). Recent advances and clinical applications of deep learning in medical image analysis. *Medical Image Analysis*, 79, 102444.

https://doi.org/10.1016/j.media.2022.102444

Cracknell, M. J., & Reading, A. M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Computers & Geosciences*, 63, 22–33.

https://doi.org/10.1016/j.cageo.2013.10.008

Kuhn, S., Cracknell, M. J., & Reading, A. M. (2018). Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. *Geophysics*, 83(4), B183–B193.

https://doi.org/10.1190/geo2017-0590.1

Latifovic, R., Pouliot, D., & Campbell, J. (2018). Assessment of Convolution Neural Networks for Surficial Geology Mapping in the South Rae Geological Region, Northwest Territories, Canada. *Remote Sensing*, 10(2), 307.

https://doi.org/10.3390/rs10020307

Li, D., Deogun, J., Spaulding, W., & Shuart, B. (2004). Towards Missing Data Imputation: A Study of Fuzzy K-means Clustering Method. In: *Rough Sets and Current Trends in Computing (Lecture Notes in Computer Science)*. Heidelberg: Springer Berlin Heidelberg, pp. 573–579.

https://doi.org/10.1007/978-3-540-25929-9_70

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166–177.

https://doi.org/10.1016/j.isprsjprs.2019.04.015

Otele, C. G. A., Onabid, M. A., Assembe, P. S., & Nkenlifack, M. (2021). Updated Lithological Map in the Forest Zone of the Centre, South and East Regions of Cameroon Using Multilayer Perceptron Neural Network and Landsat Images. *Journal of Geoscience and Environment Protection*, 9(6), 120–134.

https://doi.org/10.4236/gep.2021.96007

Othman, A. A., & Gloaguen, R. (2014). Improving Lithological Mapping by SVM Classification of Spectral and Morphological Features: The Discovery of a New Chromite Body in the Mawat Ophiolite Complex (Kurdistan, NE Iraq). *Remote Sensing*, 6(8), 6867–6896.

https://doi.org/10.3390/rs6086867

Radford, D. D. G., Cracknell, M. J., Roach, M. J., & Cumming, G. V. (2018). Geological Mapping in Western Tasmania Using Radar and Random Forests. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(9), 3075–3087.

https://doi.org/10.1109/JSTARS.2018.2855207

Rigol-Sanchez, J. P., Chica-Olmo, M., & Abarca-Hernandez, F. (2003). Artificial neural networks as a tool for mineral potential mapping with GIS. *International Journal of Remote Sensing*, 24(5), 1151–1156.

https://doi.org/10.1080/0143116021000031791

Sergi, R., Solaiman, B., Mouchot, M. C., Pasquariello, G., & Posa, P. (1995). LANDSAT-TM image classification using principal components analysis and neural networks. In: Proceedings of the 1995 International Geoscience and Remote Sensing Symposium, IGARSS '95. Quantitative Remote Sensing for Science and Applications; 10–14 July 1995; Firenze, Italy. Volume 3, pp. 1927–1929.

https://doi.org/10.1109/IGARSS.1995.524069

Shirmard, H., Farahbakhsh, E., Heidari, E., *et al.* (2022). A Comparative Study of Convolutional Neural Networks and Conventional Machine Learning Models for Lithological Mapping Using Remote Sensing Data. *Remote Sensing*, 14(4), 819.

https://doi.org/10.3390/rs14040819

Shirmard, H., Farahbakhsh, E., Müller, R. D., & Chandra, R. (2022). A review of machine learning in processing remote sensing data for mineral exploration. *Remote Sensing of Environment*, 268, 112750.

https://doi.org/10.1016/j.rse.2021.112750

Soori, M., Arezoo, B., & Dastres, R. (2023). Artificial intelligence, machine learning and deep learning in advanced robotics, a review. *Cognitive Robotics*, 3, 54–70.

https://doi.org/10.1016/j.cogr.2023.04.001

Suhaib Kamran, S., Haleem, A., Bahl, S., Javaid, M., Prakash, C., & Budhhi, D. (2022). Artificial intelligence and advanced materials in automotive industry: Potential applications and perspectives. *Materials Today: Proceedings*, 62, 4207–4214.

https://doi.org/10.1016/j.matpr.2022.04.727

Sun, T., Chen, F., Zhong, L., Liu, W., & Wang, Y. (2019). GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China. *Ore Geology Reviews*, 109, 26–49.

https://doi.org/10.1016/j.oregeorev.2019.04.003

Zhou, X., Hu, Y., Wu, J., Liang, W., Ma, J., & Jin, Q. (2023). Distribution Bias Aware Collaborative Generative Adversarial Network for Imbalanced Deep Learning in Industrial IoT. *IEEE Transactions on Industrial Informatics*, 19(1), 570–580.

https://doi.org/10.1109/TII.2022.3170149

International Journal of Systematic Innovation

ARTICLE

Variability in gesticulation patterns: A robust framework for recognizing self co-articulated dynamic gestures

Shweta Sharda[1*], **Ritu Vyas**[1], **and Joyeeta Singha**[2]

[1]Department of Electronics & Communication, Jaipur Engineering College and Research Center, Jaipur, Rajasthan, India

[2]Department of Electronics & Communication Engineering, The LNM Institute of Information Technology, Jaipur, Rajasthan, India

***Corresponding author:**
Shweta Sharda
(shweta.saboo.y18pg@lnmiit.ac.in)

Citation: Sharda, S., Vyas, R. & Singha, J. (2026). Variability in gesticulation patterns: A robust framework for recognizing self co-articulated dynamic gestures. *Int J Systematic Innovation*. 10(1):19-34.
https://doi.org/10.6977/IJoSI.202602_10(1).0003

Received: February 3, 2025

Revised: October 27, 2025

Accepted: January 5, 2026

Published online: February 13, 2026

Abstract

Dynamic hand gesture recognition has become an important research area in human–computer interaction, virtual reality, sign language interpretation, and intelligent surveillance systems. With the increasing demand for natural and contactless communication interfaces, gesture-based systems are gaining significant attention due to their intuitive and user-friendly nature. However, one of the major challenges in dynamic gesture recognition is inter-user variability, where differences in speed, style, and articulation patterns among users reduce the overall robustness and accuracy of recognition systems. Another critical issue is self co-articulation, which occurs when gestures overlap or influence each other during continuous motion, making feature extraction more complex. This study presents a dynamic hand gesture recognition system that addresses inter-user variability in gesticulation patterns. In our proposed system, a new set of features was employed, which divides the gesture into two halves, and feature extraction was performed after the removal of self-co-articulation. The efficiency of the proposed system was validated on a new set of gestures recorded in the LNM Institute of Information Technology Dynamic Hand Gesture Dataset-4, which consists of videos recorded according to different patterns. The performance of the proposed system was calculated with different features combined with individual as well as combinations of classifiers, such as support vector machine, k-nearest neighbor, naïve Bayes, adaptive neuro-fuzzy inference system, and discriminant analysis classifiers. The recognition accuracy of the naïve Bayes classifier was 93.13%, which is the best among all the classifiers. Recognition accuracy improved by about 10% with an increase in the number of features.

Keywords: Hand gesture recognition; Pattern variation; Self-co-articulation; Trajectory features

1. Introduction

A promising area in pattern-variation research is the systems developed for recognizing hand gestures. Various hand gesture systems have been designed and proposed in the literature, which can make human–computer interaction easier and simpler. Moreover, challenges arise when there is variation in gestures according to the user, which produces a hindrance in the flexibility of various practical applications. Additionally, the presence of self-co-articulation raises the problem of misclassifying gestures, which leads to incorrect recognition. If the challenges associated with pattern variation and self-co-articulation can be addressed by the hand gesture recognition systems, it will be easier to use the hand gesture systems in applications such as sign language communication, video-based surveillance, and the use of a mouse.

Static hand gestures are easy to recognize and use; however, the major challenge lies in dynamic hand gestures, as they are recorded by moving the hand in space according to the user's comfort level. In literature, users typically follow the reference pattern of the gestures provided to them. Li *et al.* (2018) used the Mobile Gesture Database, which consists of A–F letters and 1–6 numerals. Bhuyan *et al.* (2014) and Pun *et al.* (2011) developed a hand gesture recognition system using only 10 numerals from 0–9, which were recorded with a fixed set of patterns. Tang *et al.* (2018) used lowercase letters a–z for designing the hand gesture recognition system using dynamic time warping. Beh *et al.* (2014) considered numerals ranging from 1–9, a few letters like A, B, C, D, X, Y, and Z, along with five direction-based symbols. Singha *et al.* (2016) included the gestures from A–Z and 0–9 with pattern variation using a colored marker. However, previous studies have not included the bare hand gesture pattern variation. In this study, some of the natural variations of the hand gestures have been recorded and used for detection, tracking, and recognition. Figure 1 shows the gestures with the self-co-articulated strokes that need to be detected and removed to obtain the original gesture. Specifically, stroke 3 in "A," strokes 3 and 5 in "E," stroke 3 in "F," stroke 2 in "T," and stroke 2 in "X" correspond to self-co-articulated strokes (**Figure 1**). Self-co-articulation refers to strokes that occur either at the beginning of a gesture or between strokes during the formation of the gesture shape.

The main contributions of this paper are as follows:

(i) The design of a new dataset named the LNM Institute of Information Technology (LNMIIT) Dynamic Hand Gesture Dataset-4. This dataset mainly consists of numbers and letters recorded by different users. As very few datasets consisting of videos with pattern variation are publicly available, we developed our own dataset.

(ii) The detection and removal of self-co-articulation strokes occurring at the start of the gesture and those occurring at the time of gesture completion.

(iii) The proposal of a new feature that divides the stroke into two parts horizontally, such as the ratio of the number of points in the first half, the ratio of stroke length, the distance between the starting and ending points of each half, and the curliness of both halves.

(iv) Recognition accuracy was calculated for each of the gestures using different classifiers, including support vector machine (SVM), artificial neural network (ANN), k-nearest neighbor (k-NN), naïve Bayes, adaptive neuro-fuzzy inference system (ANFIS), and discriminant analysis (DA).

(v) A combination of different sets of features with classifiers was performed and analyzed to reduce the misclassification of the gestures.

The paper is organized as follows: Section 2 reviews related work. Section 3 provides an explanation of the detailed architecture of the proposed system, and describes the dataset used and the effect of pattern variation in gestures. Section 4 presents the recognition results, analyzed using different classifiers. Finally, the paper is concluded in Section 5.

2. Related works

In this section, we review various models used for tracking and summarize the features extracted from dynamic gesture trajectories for accurate recognition. Various traditional methods have been used to track gesture trajectories under certain conditions. Comaniciu *et al.* (2003) used a color histogram to develop a hand tracker model in which hand detection is determined by calculating a color histogram, which is used as a mean shift and locates the hand in video frames along with tracking. The CamShift algorithm is able to track any feature distribution representing the target in a successful way (Bradski & Kaehler, 2008). There are many techniques where the CamShift was combined with various other tracking methods, which led to improved tracking efficiency. For example, the CamShift algorithm was combined with a Kalman filter (Huang & Hong, 2011; Wang & Li, 2010). The possible positions of a target were predicted by the Kalman filter, and CamShift was subsequently used to search and match the target in the predicted areas. Kolsch & Turk (2004) introduced an algorithm using a Kanade–Lucas–Tomasi-based tracker. However, the Kanade–Lucas–Tomasi tracker did not yield good results when the hand shape changed during gesticulation.

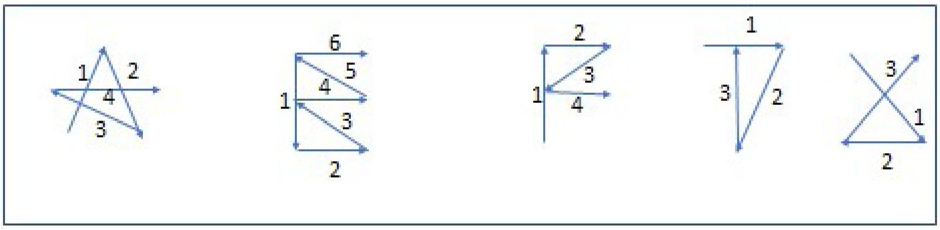


Figure 1. Gestures showing self-co-articulated strokes

Accurate tracking of the trajectory is a prerequisite for effective feature extraction, which in turn improves recognition accuracy. For developing a gesture recognition system, a single feature was used in several studies. Elmezain *et al.* (2008) tried to recognize both isolated and continuous gestures with the help of an orientation feature. With the help of this feature, gesture motion direction can be calculated using trajectory points. The quantization process was performed using code words from 1 to 18 on the angle of orientation. Orientation feature was also used by Kao & Fahn (2011) to design a real-time hand gesture recognition system. Gestures were classified using a Hidden Markov Model after the quantization process. Location, orientation, and velocity features were among the most used features by researchers (Elmezain *et al.*, 2009; Xu *et al.*, 2015; Yoon *et al.*, 2001). Xu *et al.* (2015) proposed a novel hand gesture recognition system for robotic applications using features such as orientation, with a chain code of 1–8 code words, location, and velocity. Yoon *et al.* (2001) used the combination of features such as location, orientation, and velocity.

Singla *et al.* (2019) used the normalized sequence of captured three-dimensional space coordinates as input, and a sequence of features was computed along the trajectory. Gesture direction, curvature, aspect, curliness, slope, and lineness are some of the features that have been calculated and used for the development of feature space and recognition. Misra and Laskar (2019) presented novel spatio-temporal trajectory features that provide output as structural values of the gestures. These features included the area of two halves, local geometrical area ratios, and curve-area features. The gesture was divided into two halves equally, and the area of each half provided the output as the area of the two halves. For the calculation of local geometrical area ratios, the ratio between enclosed areas in each case of stroke length was measured. Variation in the patterns of bare-hand dynamic gestures has not been adequately addressed in the existing literature, as users are typically required to record videos according to predefined gesture patterns. In this study, we developed a hand gesture recognition system that is capable of recognizing gestures with varying patterns.

Anish *et al.* (2021) proposed a gesture formation model for the removal of self-co-articulation gestures. Features like Euclidean distance, velocity, and the minimum-maximum-polarity algorithm were used as global and local measures. The experiments were done on the NITS hand gesture dataset, which showed an improvement in accuracy of about 40%. Although self-co-articulation strokes were identified and removed, pattern variations were not considered in the work. The effects of variable illumination were considered for real-time recognition of twenty-six American Sign Language signs. The model employed scale-invariant feature transform features to represent gesture patterns with inherent symmetry, with the ANN model showing a recognition accuracy of 97.03%. Cheng *et al.* (2023) used graph convolutional networks along with the path signature theory for extracting features from various skeleton joints. Since the model relied on differential features for recognition, variations in gesture patterns were not taken into account.

3. Methodology

The proposed system is mainly divided into seven phases: data acquisition, detection, tracking, trajectory smoothening, self-co-articulated stroke detection due to pattern variation, feature extraction, and the recognition of the gesture being provided as input.

Figure 2 shows the detailed flowchart of the proposed hand gesture recognition system. The detailed description of each phase is given below in subsections.

3.1. Data acquisition

Figure 3 shows a sample of the LNMIIT Dynamic Hand Gesture Dataset-4, which has videos recorded using pattern variation. Various numerals and letters have been considered for pattern variation, such as 5, A, B, D, E, F, and P. Self-co-articulation represents the unwanted strokes when recording the gestures. These unwanted movements can change the shape of the character; hence, they have to be recognized. It has been observed that the gestures that show the pattern variation consist of an extra stroke at the start of the gesture trajectory. This extra stroke, if

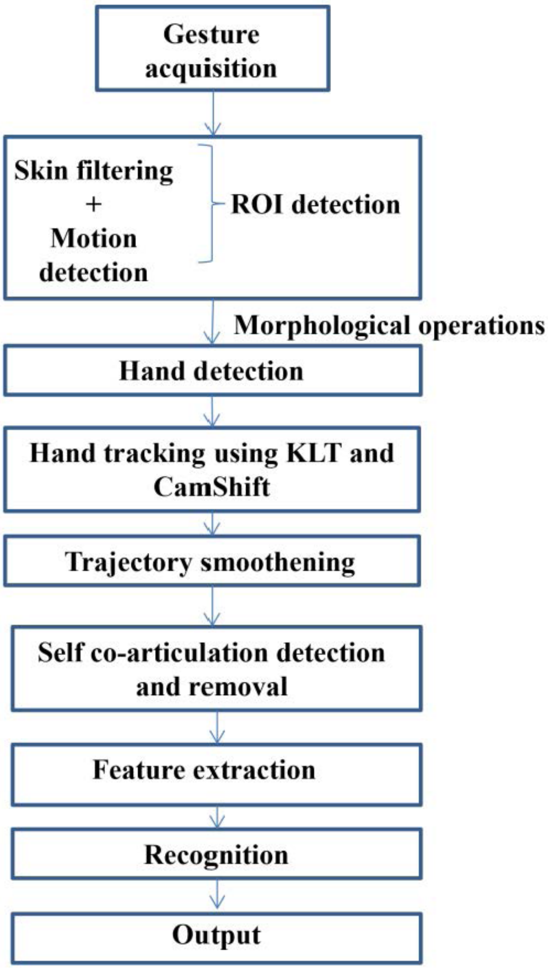


Figure 2. Flowchart of the proposed system
Abbreviations: KLT: Kanade–Lucas–Tomasi; ROI: Region of interest.

removed, will lead to an ordinary gesture that can be easily recognized by the proposed system. The self-co-articulated stroke present will extract the changed features, such as the number of points, stroke length, and angle, leading to a misclassification of the gesture and hence, an incorrect recognition. However, there is a requirement to remove the other self-co-articulated strokes also.

We developed the LNMIIT Dynamic Hand Gesture Dataset-4 with the help of various users who were using different ways of gesturing alphanumeric characters. Acquisition of gestures was performed in an indoor environment using a Logitech (Switzerland) C922 Pro Stream Webcam (640 × 360 pixels) with an aspect ratio of 16:9 and 30 frames per second. Figure 4 shows two patterns for different alphanumeric characters. Pattern 1 shows the gestures without any extra stroke, while pattern 2 represents gestures with self-co-articulated strokes. In numeral 5, pattern 2 consists of stroke number 2, which represents the self-co-articulated stroke. In the letter "A," stroke 2 and stroke 4 represent the two different self-co-articulated strokes. Only some of the alphanumeric characters have been shown in the figure. Both types of self-co-articulated strokes must be detected and removed to obtain the proper gesture, which can then be appropriately recognized.

Gestures from the LNMIIT Dynamic Hand Gesture Dataset-3 have also been used for the proposed system. This dataset consists of numbers from 0–9, letters from A–Z, and lowercase letters from a–z.

The LNMIIT Dynamic Hand Gesture Dataset-4 consists of alphanumeric characters showing pattern variation. Details of various datasets are given in Table 1. The constraints used for recording gestures in the LNMIIT Dynamic Hand Gesture Dataset-4 are:

(i) At the start of recording the gesture, the hand must be placed in the correct position.
(ii) The palm should exhibit clearer motion than the forearm to illustrate the gesture distinctly.
(iii) The movement of the hand must be smooth and continuous.
(iv) Lighting must be adequate at the time of recording.
(v) The hand must be kept in a static gesture position for a few seconds before the completion of the gesture.
(vi) The users were asked to record a video in their natural way of writing the characters.

3.2. Detection

Detection plays the most important role in the proper recognition of the gesture. After separating the region of interest with the help of a skin detection process, morphological operations were applied to remove non-hand skin regions and refine the detected hand mask. Skin filtering was followed by motion detection, giving only the hand region as the output. A bounding box was created around the hand region so that its coordinates can be used for obtaining the trajectory points. A combination of HSV and YCbCr color models was used for this purpose (Saboo *et al*., 2021). Figure 5 shows the various steps involved in the process of hand detection.

3.3. Tracking and trajectory smoothing

The motion of the detected hand must be tracked to obtain the trajectory of the recorded gesture. All the frames of the gesture video were taken into consideration, and the center of the hand region of each frame was marked to form the hand gesture trajectory. As the plotted trajectory contained some irregular strokes due to the movement of the hand at the start of the gesture, such as the diversion

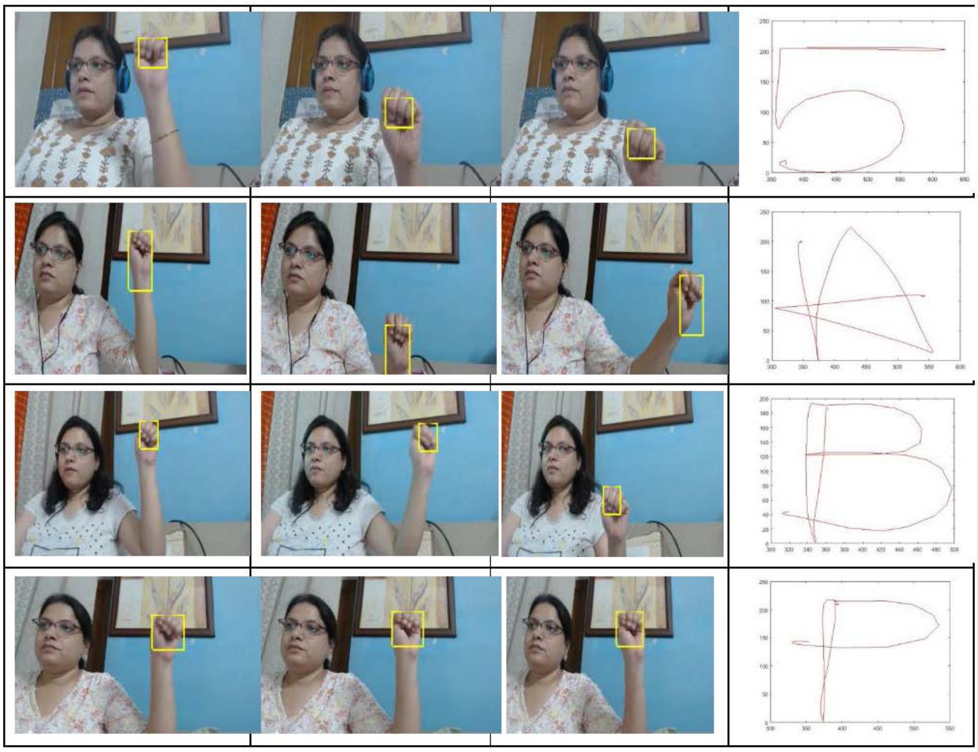


Figure 3. Video frames of some of the gestures from the LNM Institute of Information Technology Dynamic Hand Gesture Dataset-4

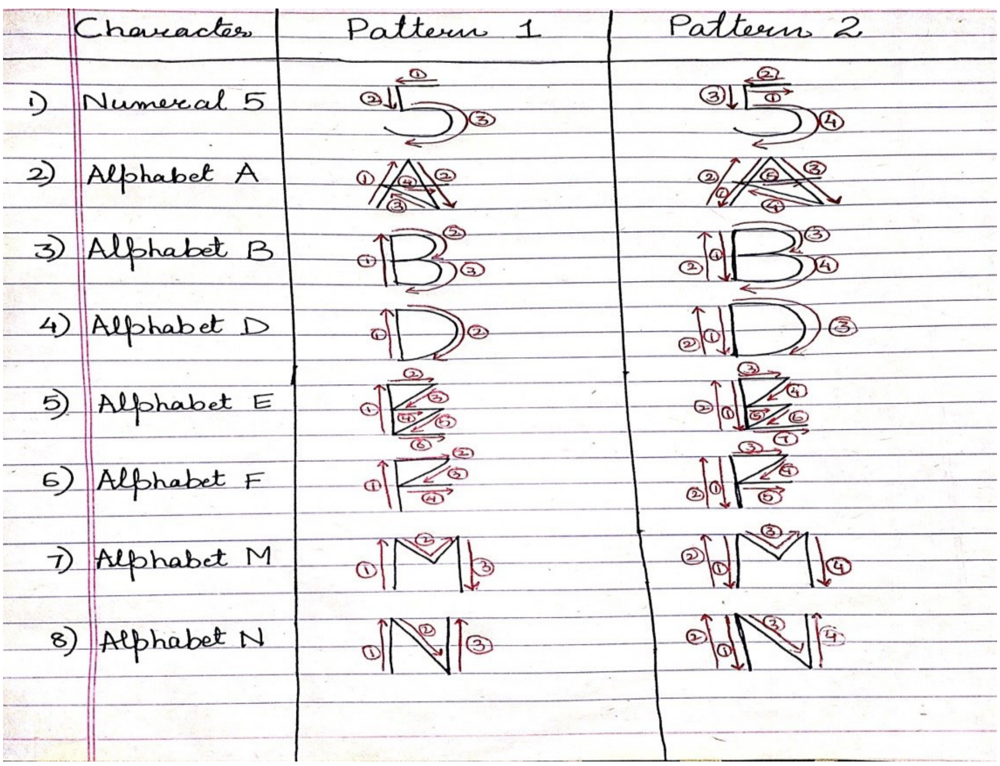


Figure 4. A few samples of gesture pattern variations

Table 1. Description of the hand gesture dataset

Characteristics	LNMIIT Dynamic Hand Gesture Dataset-1	LNMIIT Dynamic Hand Gesture Dataset-2	LNMIIT Dynamic Hand Gesture Dataset-3	LNMIIT Dynamic Hand Gesture Dataset-4
No. of videos	500	900	1,860	720
Details	0–9, A, C, S	0–9 numbers, a–z excluding f, i, j, k, t, and x	0–9 numbers, A–Z, and a–z	0–9 numbers, A–Z
Variations	Hand shape variation, illumination variation, and the presence of multiple persons	No variation	No variation with self-co-articulation	Pattern variation with self-co-articulation
Acquisition device	ArcSoft Web Companion	Logitech C922 Pro Stream Webcam	Logitech C922 Pro Stream Webcam	Logitech C922 Pro Stream Webcam
Resolution	640 × 480 pixels	640 × 360 pixels	640 × 360 pixels	640 × 360 pixels
No. of users	3–4	5	10	12
Frame rate (frames per second)	30	30	30	30

Abbreviation: LNMIIT: LNM Institute of Information Technology.

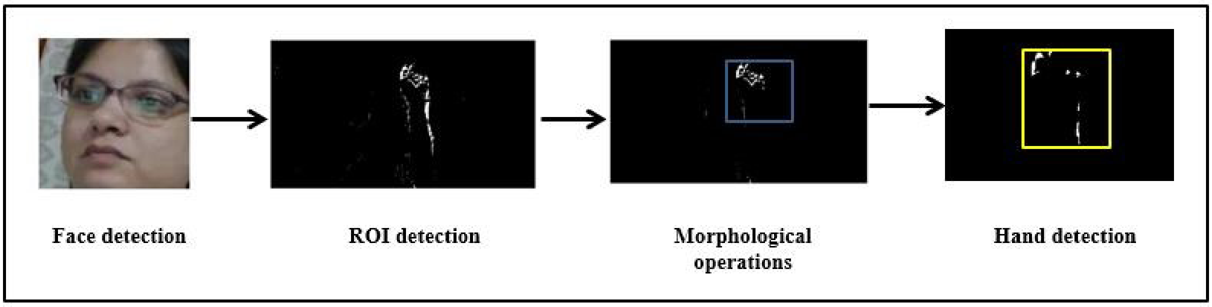


Figure 5. Steps of hand detection
Abbreviation: ROI: Region of interest.

of the bounding box, trajectory smoothening becomes a necessary process. A smoothed trajectory was obtained by averaging the previous, current, and next points and replacing all three of them with the calculated point (Saboo *et al.*, 2022). Figure 6 shows the frame-to-frame tracking, indicating the original trajectory of the gesture. Removal of the starting noise and averaging of the gesture points produced a smoothed gesture trajectory as the output (Equations 1–3).

$$x(t), y(t) = \frac{x(i-1)+x(i)+x(i+1)}{3} ; \frac{y(i-1)+y(i)+y(i+1)}{3} \quad (1)$$

$$x(s), y(s) = \frac{x(1)+\cdots x(5)}{5} ; \frac{y(1)+\cdots y(5)}{5} \quad (2)$$

$$x, y = (x(s) : x(t)) ; (y(s) : y(t)) \quad (3)$$

where $x(s)$ and $y(s)$ were used to remove starting noise; $x(t)$ and $y(t)$ represent the averaged points, and x and y represent the current array of trajectory points. Smoothed trajectory helps in the proper recognition of the gesture and hence is an important process in the steps of a hand gesture recognition system.

3.4. Self-co-articulation detection and removal

A gesture comprises different types of self-co-articulated strokes. In this work, we focused on removing two such strokes: one arising due to pattern variation and another occurring during the formation of the gesture itself. These self-co-articulated strokes can introduce unwanted feature variations and lead to incorrect gesture recognition. Thus, proper detection and removal of self-co-articulated strokes become an important phase in the proposed hand gesture recognition system. As illustrated in Figure 3, the self-co-articulation caused by pattern variation typically occurs at the beginning of the gesture. To address this, experiments were conducted to analyze the number of points in the majority of gestures, which was then used to determine an appropriate threshold for point consideration.

Table 2 consists of the various values of the number of points that help in deciding the threshold value required for extracting the points to be removed. Hence, we considered starting from 15 points and checking the

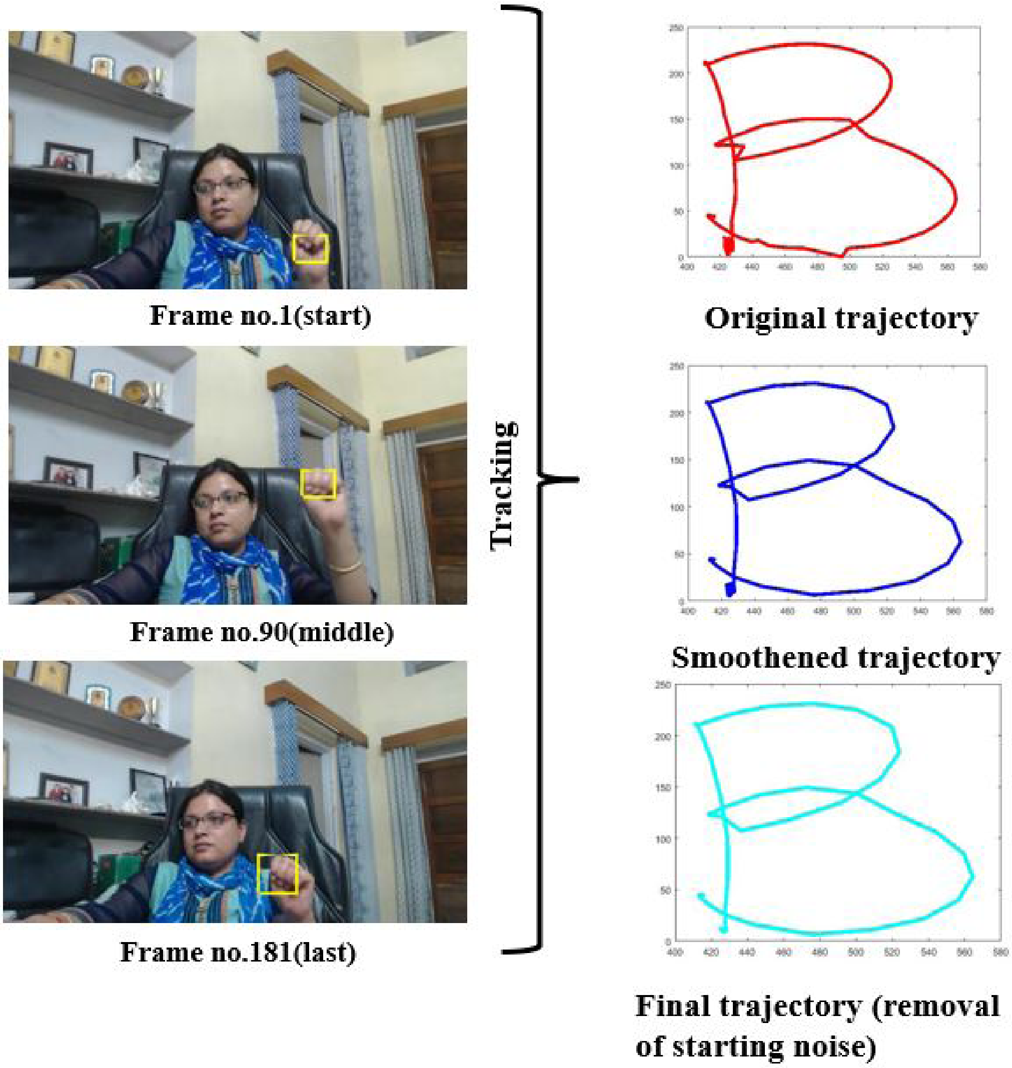


Figure 6. Gesture tracking showing the original and smoothed trajectory

difference between each point and the next point.

A point is considered the start of the gesture when the difference changes sign compared to previous values. All points occurring before this sign-change point are regarded as part of self-co-articulated strokes and should be removed.

Another type of self-co-articulation is the stroke occurring due to the formation of the gesture. It can be seen from Figure 7 that all the gestures have a co-articulation stroke in the negative x or y direction. The methodology for removing self-co-articulated strokes began by calculating the differences between consecutive trajectory points. Points with negative difference values were identified and separated (Equations 4 and 5). Figure 7 shows the flowchart of the detection and removal of self-co-articulated strokes.

$$diff = y(t+1) - y(t) \quad (4)$$

$$Y = Y(i)\ if\ diff > 0 \quad (5)$$

where $y(t)$ represents the current point, and all those points should be removed for which the difference is negative.

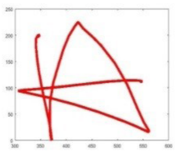
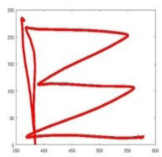
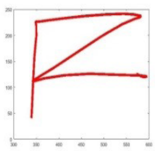
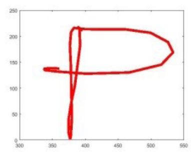
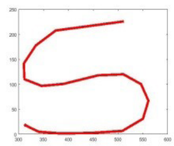
In continuous hand gesture movements, self-co-articulation strokes encounter problems during movement from one gesture to another. It is observed that the duration of a gesture is longer than the duration of the co-articulated strokes.

3.5. Feature extraction

Extraction of the features from the trajectory points is required for matching the trajectory with the gesture.

These features are provided as the input to various

Table 2. The number of points for starting the self-co-articulation stroke

Gesture					
No. of self-co-articulation strokes present	Two (one at the start of the gesture and another after three strokes)	Three (one at the start of the gesture and two after three strokes)	One (After two strokes)	One (At the start of the stroke)	No self-co-articulation stroke
No. of points for starting self-co-articulation	12	15	17	14	16

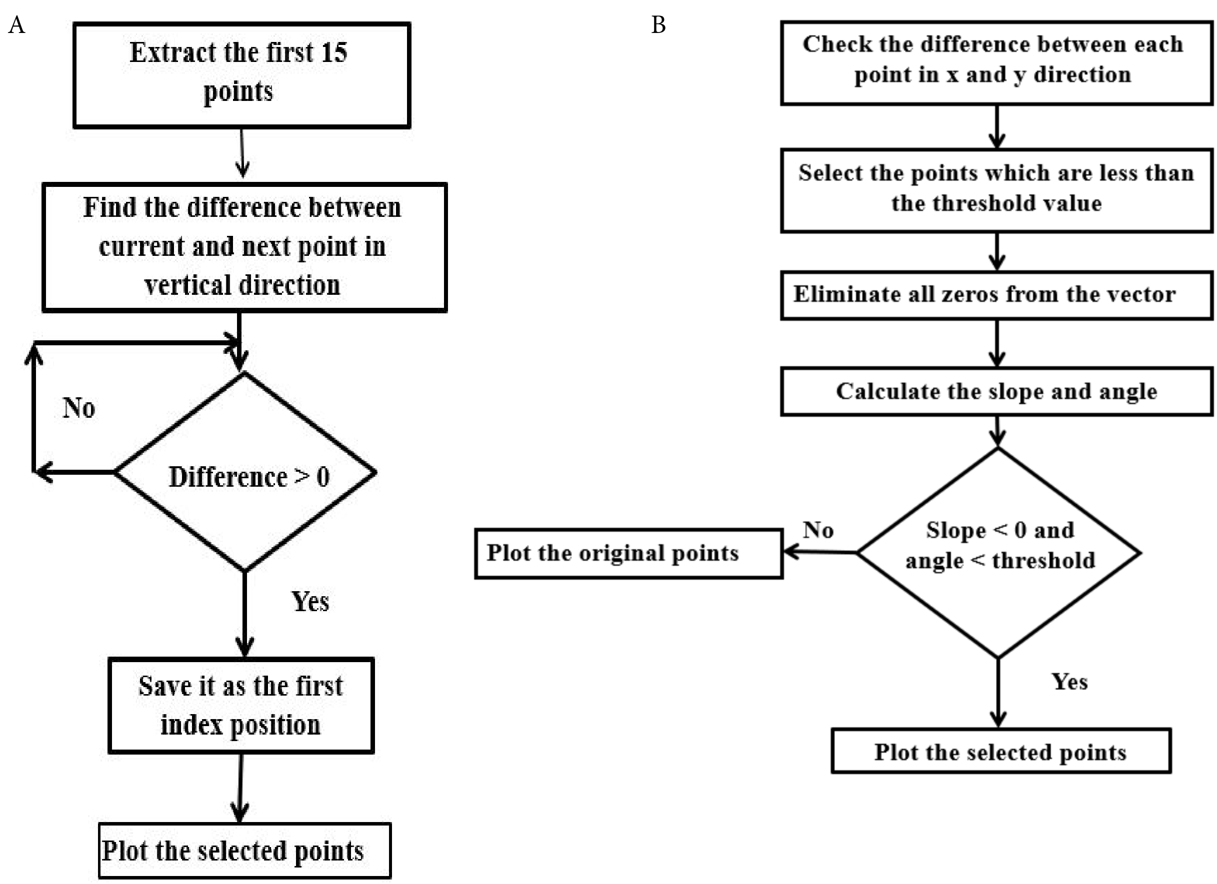


Figure 7. Flowcharts for the detection and removal of self-co-articulation strokes. (A) Starting of the self-co-articulated strokes. (B) Self-co-articulated strokes at the time of formation of the gesture shape.

classifiers, and hence they should be appropriately determined. In general, features that exhibited consistency across instances of the same gesture were considered for analysis. A total of 12 features were considered in this study, among which some were taken from the literature, and some were proposed by dividing the trajectory into two halves horizontally. The features, such as the distance feature (Rubine, 1991b), the location feature (Xu *et al.*, 2015), and the close figure test feature (Misra *et al.*, 2018), were taken from the literature. A subset of features was selected from the available feature set provided by Rubine (1991a). Table 3 provides the list of proposed and existing features.

3.5.1. The ratio of the number of points in the first and second half, and the ratio of the midpoint of the first and second half

After removing the self-co-articulated strokes, the resulting gesture shape matches the original gesture. The gesture trajectory was divided into two halves in the horizontal direction, and the number of points was calculated for the two halves. The ratio of these numbers was considered as one of the features. For each half of the trajectory, the midpoint was calculated, and the ratio of both values was taken as a new feature, which matches the pattern variation gestures.

3.5.2. Stroke length of half/Distance between first and last point of half

This feature was calculated by first identifying the complete distance traversed by the trajectory and dividing it by the distance between the first and last points, as shown in Equation 6:

$$\frac{\sum_{i=1}^{n}\sqrt{[x(i+1)-x(i)]^2+[y(i+1)-y(i)]^2}}{\sqrt{(x_{p-1}-x_o)^2+(y_{p-1}-y_o)^2}} \quad (6)$$

where, x_{p-1} represents the last point and x_o represents the first point of the trajectory. A feature reflects more than one entry in the taxonomy; for example, the entropy feature was considered a measure of density.

3.5.3. Perimeter efficiency of halves

To find the value of this feature, all the two-dimensional coordinates were used to create an alpha shape, and the perimeter was calculated. Area of the alpha shape when multiplied by pi and divided by the perimeter gives the perimeter efficiency as Equation 7:

$$PE = 2 * \frac{\sqrt{\pi * A}}{P} \quad (7)$$

where A and P represent the area and perimeter, respectively. The selected value of the alpha radius gives a scalar quantity specifying the radius of the alpha disk or sphere used to recover the alpha shape. Figure 8 represents different gestures with some of the features taken into consideration.

3.5.4. Ratio of convex hull area of the first and second half

A convex hull can be defined as the shape of the smallest convex set that contains it. To calculate the convex hull, the point with the minimum x coordinate value or the leftmost point was taken as the starting point, and the points were wrapped up in a counterclockwise direction.

3.5.5. Curliness of halves

The curliness feature states the deviation of a gesture from a straight line in the vicinity of a curl in two dimensions, as shown in Equation 8:

$$\text{Curliness, } C(t) = \frac{P}{\max(\Delta x, \Delta y)} - 2 \quad (8)$$

3.5.6. Minimum bounding rectangle

This feature finds the average of the rectangles bounding the various trajectory points. The area of the minimum bounding rectangle is calculated by finding the ratio of the difference between each trajectory and the minimum value by the difference between the maximum value of the trajectory points and the minimum value of the trajectory points, as shown in Equations 9 and 10:

Table 3. List of total features used in the proposed system

Existing features	Proposed features
• Length feature (Rubine, 1991a) • Angle feature (Rubine, 1991a) • Distance feature (Rubine, 1991b) • Close figure test (Misra *et al.*, 2018) • Location feature (Xu *et al.*, 2015)	• Ratio of the number of points in the 1st and 2nd half • Stroke length of half/Distance between first and last point of half • Perimeter efficiency of halves • Curliness of halves • Ratio of the midpoint of the 1st half and the 2nd half • Ratio of the convex hull area of the 1st and 2nd half • Minimum bounding rectangle

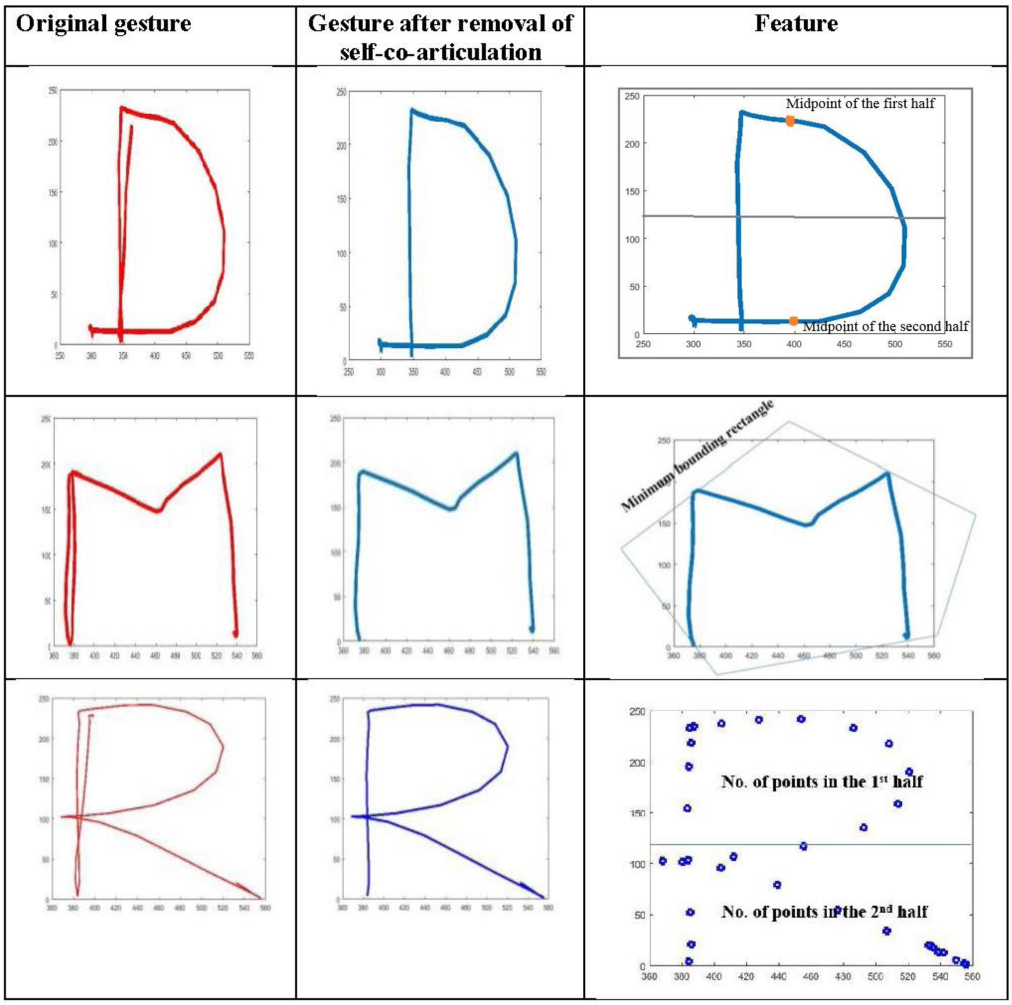


Figure 8. Gestures representing some of the proposed features

$$x_{max} = \max_{t=1} n\,(X_t)\ ; \quad x_{min} = \min_{t=1} n\,(X_t)$$

$$y_{max} = \max_{t=1} n\,(Y_t); \quad y_{min} = \min_{t=1} n\,(Y_t)$$

$$x_t = \frac{X_t - x_{min}}{x_{max} - x_{min}}\ ; \quad y_t = \frac{Y_t - y_{min}}{y_{max} - y_{min}} \tag{9}$$

$$mbr = x_t * y_t \tag{10}$$

3.6. Recognition

Support vector machine can be used to classify for both supervised and unsupervised learning. SVM uses various types of kernels like linear, Gaussian, and polynomial. It projects the data into a higher-dimensional space, enabling the separation of linearly non-separable data and reducing classification errors (Hsu & Lin, 2002). k-NN is a classifier that solves both classification and regression problems by using supervised machine learning methods. This algorithm assumes that similar things are available in proximity and thus captures the idea of similarity.

Naïve Bayes classifications assume that, given the class, the features are conditionally independent, even though the class itself may depend on multiple features (Singha *et al.*, 2016). An ANN is a gesture recognition approach used by various researchers. ANN comprises an input, hidden, and output layer with neurons used according to the available dataset (Bamwenda & Özerdem, 2019).

Adaptive neuro-fuzzy inference system (Subasi, 2007) uses a hybrid learning algorithm and applies a combination of the least-squares method and the back-propagation gradient descent method for training FIS membership function parameters to emulate a given training dataset. The quadratic DA (Tharwat, 2016) model is a discriminant model of classification used for predicting the response of the test data. This type of classifier can be linear or quadratic in nature.

4. Results and discussion

Tracking accuracy was calculated for various classifiers individually as well as with different combinations of features. The LNMIIT Dynamic Hand Gesture Dataset-3 and the LNMIIT Dynamic Hand Gesture Dataset-4 were used, which consist of 20 samples each of numbers and letters; hence, a total of 720 gestures were used in this experiment. Out of 720 gestures, the ratio of training and testing sets was separated as 70% and 30%, respectively. Different kernel functions, such as linear, Gaussian, and polynomial, were used in SVM. The polynomial kernel provided the best accuracy of 85.88% for the complete dataset (Table 4). Tests were conducted for numbers and letters separately, and then both were considered together.

We also evaluated the performance of the ANN classifier, and the results are shown in Table 4. Training, testing, and validation accuracies were also calculated along with the overall classifier accuracy. The highest accuracy for 20 hidden neurons and network structure 22–36–20 was obtained, which was 100%, 94.56%, and 90.21% for numbers, letters, and alphanumeric characters, respectively. The number of hidden neurons and layers was changed, and different permutations and combinations were tested to obtain reliable accuracy.

For different values of k, the k-NN classifier was also used to calculate the recognition accuracy. The highest accuracy was calculated for k = 1. Recognition accuracy for

Table 4. Recognition accuracy of the classifiers

Classifiers	Accuracy measures		
	Numerals	Letters	Alphanumeric characters
Support vector machine (SVM): polynomial kernel			
SVM + existing features	65.52%	63.31%	60.24%
SVM + 15 features	74.22%	72.23%	69.87%
SVM + 20 features	88.33%	86.36%	85.88%
Artificial neural network (ANN): network structure 22–36–20			
ANN + existing features	74.45%	70.23%	65.52%
ANN + 15 features	82.24%	81.12%	78.86%
ANN + 20 features	100%	94.56%	90.21%
k-nearest neighbor (k-NN): k = 1			
k-NN + existing features	74.25%	72%	68.8%
k-NN + 15 features	82.2%	75.05%	70.01%
k-NN + 20 features	95%	91.67%	88.89%
Naïve Bayes (NB): Data = kernel			
NB + existing features	80.1%	79.5%	75.44%
NB + 15 features	87.74%	83.34%	81.14%
NB + 20 features	96.43%	95.94%	93.13%
Adaptive neuro-fuzzy inference system (ANFIS): Linear			
ANFIS + existing features	67.52%	62.33%	60.01%
ANFIS + 15 features	75.29%	72.55%	70.14%
ANFIS + 20 features	96.50%	92.23%	89.96%
Discriminant analysis (DA): Linear			
DA + existing features	80.21%	79.96%	78.08%
DA + 15 features	86.65%	82.22%	81.14%
DA + 20 features	95%	95.51%	90.28%

numbers, letters, and a combination of both was obtained as 95%, 91.67%, and 88.89%, respectively (Table 4). The value of k is generally taken as an odd number. Naïve Bayes classifier, when used with data distribution as "kernel," yielded the best accuracy of 96.43% for the number dataset. Using the existing features, the classifier achieved a maximum accuracy of 80.1% for numerals. Increasing the number of features to 15 improved the accuracy by approximately 7%. Finally, expanding the feature set to 20 yielded a satisfactory recognition accuracy of 96.43%. When performed with an alphanumeric dataset, the accuracy obtained was 93.13% (Table 4).

The ANFIS classifier was used to solve the problems regarding fuzzy classification. Table 4 reports the accuracy of 89.96% when the function used was linear. This type of classifier helps in recognizing similar types of gestures like "o" and "0," "2," and "Z." Another type of classifier used for checking the effectiveness of the proposed system is the quadratic type of linear DA, which helps in classifying the gestures that are generated based on Gaussian distributions. As the number of features increased, the recognition accuracy increased to 95%, which is better than the classification conducted with 15 features (Table 4). Analysis of the individual classifiers shows that naïve Bayes achieved the highest accuracy, while SVM yielded the lowest (Figure 9).

The reduced accuracy is attributed to confusing gestures such as O, 0, 2, and Z. Excluding these misclassified cases improved recognition accuracy by approximately 3–4%. Table 5 shows the results with and without misclassified cases. This misclassification arises because these characters share similar shapes and patterns, as shown in Figure 10. The proposed system is designed for individuals who have become impaired due to accidents and are unfamiliar with

Table 5. Results of classifiers with and without misclassification

Classifiers	With misclassified gestures	Without misclassified gestures
SVM	85.88%	91.26%
ANN	90.21%	96.62%
k-NN	88.89%	94.44%
Naïve Bayes	93.13%	96.36%
ANFIS	89.96%	92.22%
Quadratic DA	90.28%	92.37%

Abbreviations: ANFIS: Adaptive neuro-fuzzy inference system; ANN: Artificial neural network; DA: Discriminant analysis; k-NN: k-nearest neighbor; SVM: Support vector machine.

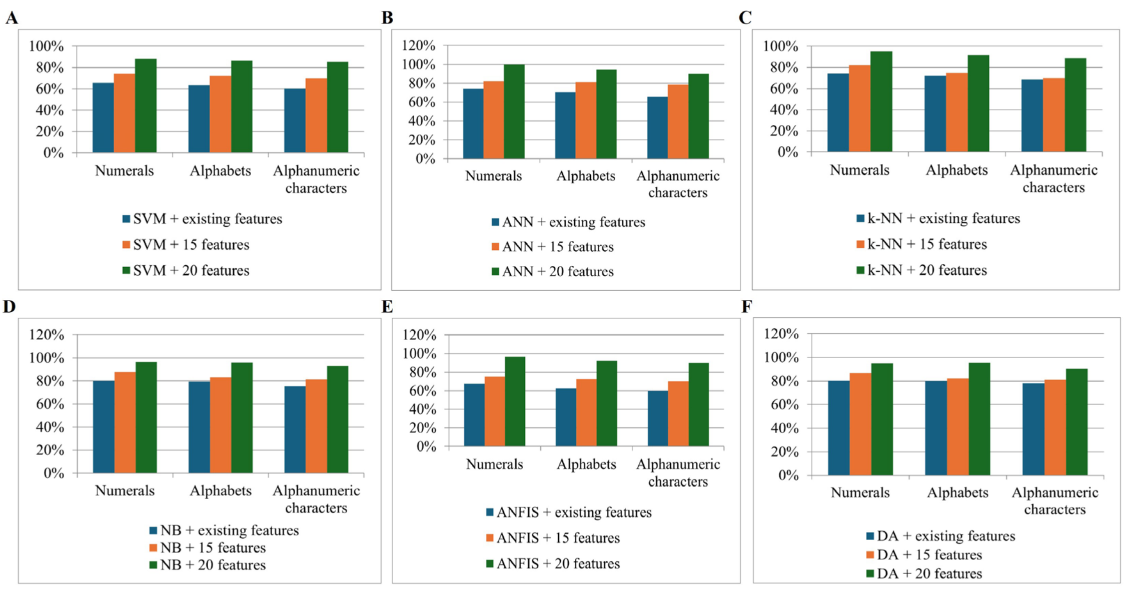


Figure 9. Performance in terms of accuracy using different feature combinations. (A) Support vector machine (SVM), (B) artificial neural network (ANN), (C) k-nearest neighbor (k-NN), (D) naïve Bayes (NB), (E) adaptive neuro-fuzzy inference system (ANFIS), and (F) quadratic discriminant analysis (DA).

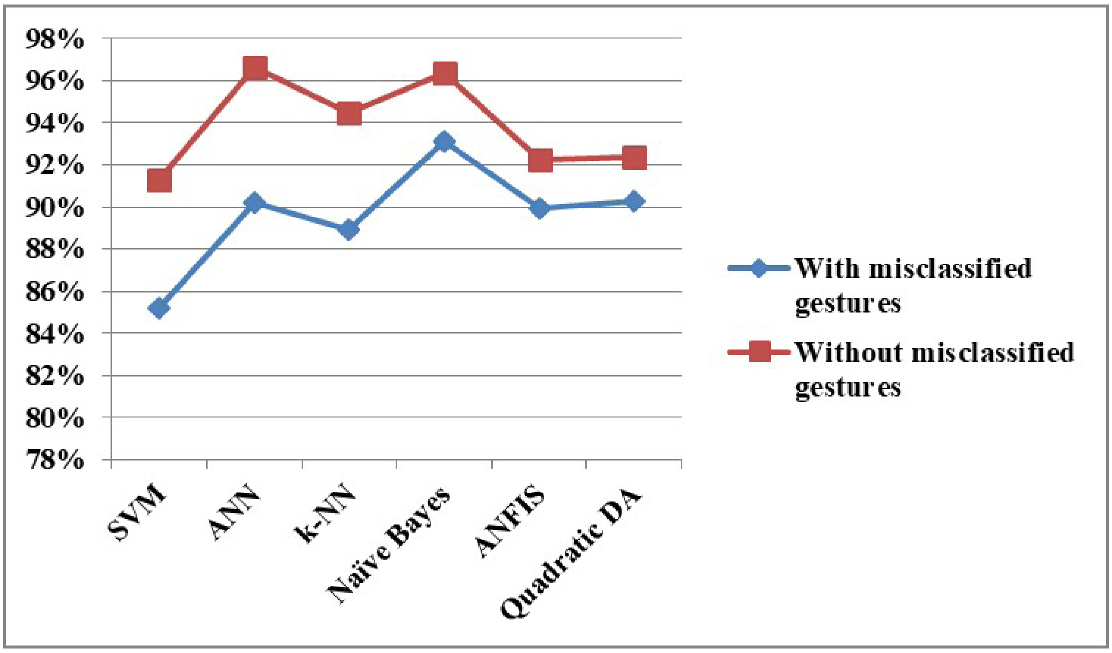


Figure 10. Output graph showing the change in efficiency with respect to misclassified gestures

Abbreviations: ANFIS: Adaptive neuro-fuzzy inference system; ANN: Artificial neural network; DA: Discriminant analysis; k-NN: k-nearest neighbor; SVM: Support vector machine.

sign language. Figure 11 shows one of the applications implemented using the proposed system.

To validate the results of the proposed system, a comparison with existing literature was conducted, and the results are shown in Table 6. Accuracies calculated in these papers using different machine learning algorithms, such as SVM, ANN, k-NN, and naïve Bayes, were compared with the proposed system. It can be observed that the proposed system achieved the highest mean accuracy for all the individual classifiers.

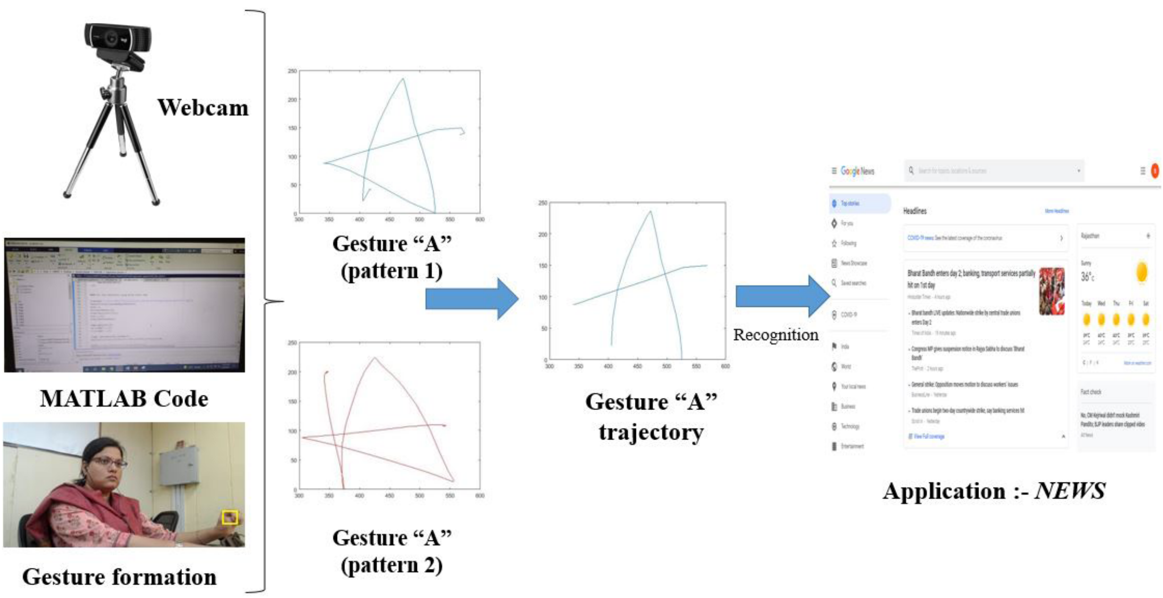


Figure 11. Block diagram for the application interface

Table 6. Comparison of existing literature with the proposed system

Method	SVM accuracy (%)	ANN accuracy (%)	k-NN accuracy (%)	Naïve Bayes accuracy (%)
Bhuyan et al. (2014)	60.31	62.77	59.64	55.98
Singha et al. (2016)	85.34	88.69	82.21	79.65
Proposed	85.88	90.21	88.89	93.13

Abbreviations: ANN: Artificial neural network; k-NN: k-nearest neighbor; SVM: Support vector machine.

5. Conclusion

Few studies in the literature have focused on gestures exhibiting pattern variations. Publicly available hand gesture datasets contain only a limited number of dynamic gestures, which are typically recorded following a predefined pattern that serves as the reference. To calculate the recognition accuracy of the various hand gestures, a dataset named LNMIIT Dynamic Hand Gesture Dataset-4 was developed, which contains all the alphanumeric characters with different variations in the pattern of the gestures. Experiments were conducted on 720 samples recorded with more than 10 people. Hand detection, hand tracking, and feature extraction were supported with the detection and removal of self-co-articulated strokes, which improved the recognition accuracy of the proposed hand gesture recognition system. An algorithm was developed to remove the starting self-co-articulated strokes, followed by the removal of other self-co-articulated strokes that occurred during the formation of some of the gestures. Several new features, such as the ratio of the points contained in the two halves of the gesture trajectory and the curliness feature, were introduced. Different features and their combination with classifiers, including SVM, k-NN, ANN, naïve Bayes, ANFIS, and DA classifiers, were explored in this study. The naïve Bayes classifier achieved the best accuracy of 93.13%, which is an improvement of 3% compared to ANN and quadratic classifiers. The proposed system was tested on different gestures of numerals and letters with pattern variation, gestures without pattern variation, and the total set of gestures. It can be observed that reduced accuracy was due to misclassification of gestures, such as 2, Z, O, 0, 5, and S. Future work should include the development of a new set of features to improve the discrimination of the gestures from one another. Additionally, recognition of gestures to form words may be included in future research work.

Acknowledgments

We would like to express our sincere gratitude to all those who contributed to the successful completion of this research project. Their support, guidance, and expertise were invaluable in shaping the outcome of this study.

Funding

None.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: Shweta Sharda
Formal analysis: Joyeeta Singha
Investigation: Shweta Sharda
Methodology: Shweta Sharda
Visualization: Shweta Sharda
Writing–original draft: Shweta Sharda
Writing–review & editing: Joyeeta Singha, Ritu Vyas

Ethics approval and consent to participate

Ethics approval was not required for this study. All participants provided informed consent before participating in this study.

Consent for publication

All participants provided consent for the publication of the study's findings and any accompanying images.

Availability of data

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Anish Monsley, K., Yadav, K. S., Misra, S., Khan, T., Bhuyan, M. K., & Laskar, R. H. (2021). Segregation of meaningful strokes, a pre-requisite for self co-articulation removal in isolated dynamic gestures. *IET Image Processing,* 15(5), 1166-1178.

https://doi.org/10.1049/ipr2.12095

Bamwenda, J., & Özerdem, M. S. (2019). Recognition of static hand gesture with using ANN and SVM. *Dicle University Journal of Engineering*, 10, 561-568.

https://doi.org/10.24012/dumf.569357

Beh, J., Han, D., & Ko, H. (2014). Rule-based trajectory segmentation for modeling hand motion trajectory. *Pattern Recognition*, 47(4), 1586-1601.

Bhuyan, M. K., Ajay Kumar, D., MacDorman, K. F., & Iwahori, Y. (2014). A novel set of features for continuous hand gesture recognition. *Journal on Multimodal User Interfaces*, 8(4), 333-343.

https://doi.org/10.1007/s12193-014-0165-0

Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. Sevastopol, CA: O'Reilly Media, Inc..

Cheng, J., Shi, D., Li, C., *et al.* (2023). Skeleton-based gesture recognition with learnable paths and signature features. *IEEE Transactions on Multimedia*, 26, 3951-3961.

https://doi.org/10.1109/TMM.2023.3318242

Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on pattern analysis and machine intelligence*, 25(5), 564-577.

https://doi.org/10.1109/TPAMI.2003.1195991

Elmezain, M., Al-Hamadi, A., Appenrodt, J., & Michaelis, B. (2008). A hidden markov model-based continuous gesture recognition system for hand motion trajectory. In: Proceedings of the 2008 19th international conference on pattern recognition; December 8-11, 2008, Tampa, FL, USA. pp. 1-4.

https://doi.org/10.1109/ICPR.2008.4761080

Elmezain, M., Al-Hamadi, A., & Michaelis, B. (2009). Hand gesture recognition based on combined features extraction. *International Journal of Electrical and Computer Engineering*, 3(12), 2389-2394.

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, 13(2), 415-425.

https://doi.org/10.1109/72.991427

Huang, S., & Hong, J. (2011). Moving object tracking system based on camshift and Kalman filter. In: Proceedings of the 2011 International conference on consumer electronics, communications and networks (CECNet); April 16-18, 2011, Xianning, China. pp. 1423-1426.

https://doi.org/10.1109/CECNET.2011.5769081

Kao, C. Y., & Fahn, C. S. (2011). A human-machine interaction technique: hand gesture recognition based on hidden Markov models with trajectory of hand motion. *Procedia Engineering*, 15, 3739-3743.

https://doi.org/10.1016/j.proeng.2011.08.700

Kolsch, M., & Turk, M. (2004). Fast 2d hand tracking with flocks of features and multi-cue integration. In: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop; June 27-July 2, 2004; Washington, DC, USA. pp. 158-158.

https://doi.org/10.1109/CVPR.2004.345

Li, C., Xie, C., Zhang, B., Chen, C., & Han, J. (2018). Deep Fisher discriminant learning for mobile hand gesture recognition. *Pattern Recognition*, 77, 276-288.

https://doi.org/10.1016/j.patcog.2017.12.023

Misra, S., Singha, J., & Laskar, R. H. (2018). Vision-based hand gesture recognition of alphabets, numbers, arithmetic operators and ASCII characters in order to develop a virtual text-entry interface system. *Neural Computing and Applications*, 29(8), 117-135.

https://doi.org/10.1007/s00521-017-2838-6

Misra, S., & Laskar, R. H. (2019). Development of a hierarchical dynamic keyboard character recognition system using trajectory features and scale-invariant holistic modeling of characters. *Journal of Ambient Intelligence and Humanized Computing*, 10(12), 4901-4923.

https://doi.org/10.1007/s12652-019-01189-2

Pun, C. M., Zhu, H. M., & Feng, W. (2011). Real-time hand gesture recognition using motion tracking. *International Journal of Computational Intelligence Systems*, 4(2), 277-286.

https://doi.org/10.2991/ijcis.2011.4.2.15

Rubine, D. (1991a). Specifying gestures by example. *ACM SIGGRAPH computer graphics*, 25(4), 329-337.

https://doi.org/10.1145/127719.122753

Rubine, D. (1991b). The automatic recognition of gestures [PhD thesis]. Carnegie Mellon University.

Saboo, S., & Singha, J. (2021). Vision based two-level hand tracking system for dynamic hand gestures in indoor environment. *Multimedia Tools and Applications*, 80(13), 20579-20598.

https://doi.org/10.1007/s11042-021-10669-7

Saboo, S., Singha, J., & Laskar, R. H. (2022). Dynamic hand gesture recognition using combination of two-level tracker and trajectory-guided features. *Multimedia Systems*, 28(1), 183-194.

https://doi.org/10.1007/s00530-021-00811-8

Singha, J., Misra, S., & Laskar, R. H. (2016). Effect of variation in gesticulation pattern in dynamic hand gesture recognition system. *Neurocomputing*, 208, 269-280.

https://doi.org/10.1016/j.neucom.2016.05.049

Singla, A., Roy, P. P., & Dogra, D. P. (2019). Visual rendering of shapes on 2D display devices guided by hand gestures. *Displays*, 57, 18-33.

https://doi.org/10.1016/j.displa.2019.03.001

Subasi, A. (2007). Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction. *Computers in biology and medicine*, 37(2), 227-244.

https://doi.org/10.1016/j.compbiomed.2005.12.003

Tang, J., Cheng, H., Zhao, Y., & Guo, H. (2018). Structured dynamic time warping for continuous hand trajectory gesture recognition. *Pattern Recognition*, 80, 21-31.

https://doi.org/10.1016/j.patcog.2018.02.011

Tharwat, A. (2016). Linear vs. quadratic discriminant analysis classifier: a tutorial. *International Journal of Applied Pattern Recognition*, 3(2), 145-180.

https://doi.org/10.1504/IJAPR.2016.079050

Wang, X., & Li, X. (2010, December). The study of MovingTarget tracking based on Kalman-CamShift in the video. In: Proceedings of the 2nd International Conference on Information Science and Engineeringl; October 26-28, 2012; Chongqing, China. pp. 1-4.

https://doi.org/10.1109/ICISE.2010.5690826

Xu, D., Wu, X., Chen, Y. L., & Xu, Y. (2015). Online dynamic gesture recognition for human robot interaction. *Journal of Intelligent & Robotic Systems*, 77(3), 583-596.

https://doi.org/10.1007/s10846-014-0039-4

Yoon, H. S., Soh, J., Bae, Y. J., & Yang, H. S. (2001). Hand gesture recognition using combined features of location, angle and velocity. *Pattern recognition*, 34(7), 1491-1501.

https://doi.org/10.1016/S0031-3203(00)00096-0

ACCSCIENCE PUBLISHING

International Journal of Systematic Innovation

ARTICLE

Entrepreneurship in electronic waste management: A design thinking approach

Jurry Hatammimi[1,2]*, Astri Ghina[1,2]

[1]School of Economics and Business, Telkom University, Bandung, West Java, Indonesia

[2]Center of Excellence for Key Digital Transformation Indonesia, Research Institute for Intelligent Business and Sustainable Economy, Telkom University, Bandung, West Java, Indonesia

Abstract

Electronic waste (e-waste) is becoming an increasingly significant problem, particularly due to its hazardous and toxic content. The paucity of business management research on technology-based e-waste disposal management exacerbates this issue. The novelty of this study lies in the application of design thinking not to prototype a solution, but to gain a deep understanding of stakeholder experiences and pain points, translating them into structured problem categories that support the development of an ecopreneurial information and communication technology-driven e-waste management system. Understanding the perspectives and experiences of these e-waste stakeholders helps us identify the real issues. In the empathise stage, we conducted interviews with 10 participants, who were chosen to represent the three main stakeholder categories: e-waste disposers, government institutions, and recycling institutions. In the define stage, an affinity diagram was used to analyse the gathered data and formulate the problem. We identified four primary issues with e-waste treatment: inadequate systems, a lack of socialization, low public awareness, and high processing costs. It can be concluded that developing an efficient e-waste disposal management system is a top priority. To realise the possibility of creating an ecopreneur business, further study is suggested to proceed to the ideation stage, where solutions to the four identified problems can be generated.

Keywords: Design thinking; Disposal management; Ecopreneurship; Electronic waste; Innovation management

***Corresponding author:**
Jurry Hatammimi
(jurryhatammimi@telkomuniversity.ac.id)

Citation: Hatammimi, J., & Ghina, A. (2026). Entrepreneurship in electronic waste management: A design thinking approach. *Int J Systematic Innovation*. 10(1):35-53. https://doi.org/10.6977/IJoSI.202602_10(1).0004

Received: November 20, 2025

Revised: December 9, 2025

Accepted: December 17, 2025

Published online: February 13, 2026

1. Introduction

The United Nations (UN) 2020 Global E-Waste Monitor annual report estimated that 53 million tonnes of electronic waste (e-waste) were produced in 2019 (Forti *et al.*, 2020). This large quantity of e-waste is primarily due to the difficulties users face when attempting to repair damaged electronic goods, which encourages them to purchase new ones (Rahman, 2022). Electronics makers refer to this cycle as "planned obsolescence" or purposeful obsolescence. The overall volume of e-waste is expected to reach 74 million tonnes by 2030, increasing to 120 million tonnes by 2050 (Forti *et al*., 2020). The electronics sector is a significant industry, and e-waste is currently the fastest-growing waste stream (Harikaran *et al*., 2023). Although e-waste contains a mixture of hazardous substances and valuable materials, only 17.4% of it can be collected, processed, and recycled properly (Fanthi *et al*., 2021; Lu *et al*., 2023; Wang & Anand, 2024). Ghana and China face severe air, water, and land contamination due to improper e-waste recycling (Parajuly *et al*., 2019). According to the UN, Indonesia produced approximately 1.618 million tonnes of e-waste in 2019. Specifically, Java, as the most populous island, contributes to 56% of the total amount of Indonesian e-waste (Citarum Harum, 2021).

According to *Indonesian Government Regulation No. 27* (2020), e-waste is categorised as waste that contains hazardous and toxic materials, such as lead, chromium, mercury, cadmium, and arsenic. Therefore, e-waste should not be mixed with other types of waste. According to Khalid (2022), the Indonesian government has issued at least three other regulations regarding e-waste. There have been some efforts to manage e-waste in Indonesia. For instance, the company EwasteRJ provided e-waste drop boxes at 20 points in 12 districts (Defitri, 2022). Perkins *et al.* (2014) found that only 25% of e-waste gets recycled in formal recycling plants with worker protection. The Environmental and Sanitation Department in Bandung, Indonesia, set up 10 e-waste collection drop boxes; however, some have disappeared. In addition, according to Iqbal and Dyah (2022), public awareness of current waste management practices remains low. The population still relies heavily on traditional methods, as they find it difficult to dispose of e-waste (Wanderley & Bonacin, 2019). Recently, a new approach to gaining an understanding of e-waste among students has emerged, utilising interdisciplinary design-driven studios (Wang & Anand, 2024), as well as a novel concept of incorporating e-waste components into the electronic goods design process (Lu *et al*., 2023).

Based on the amount of e-waste generated in 2020, it was estimated that Indonesia could make a profit of USD 1.8 billion if it were recycled (Puspa, 2023). The Ministry of Environment and Forestry estimates that every 1 tonne of e-waste from telecommunications equipment contains 1.44 kg of gold, silver, copper, and other high-value commodities (Dinnata, 2017). In addition, every 1 tonne of managed e-waste can also prevent the emission of 1,400 tonnes of CO_2 (Puspa, 2023). Nuryanto and Suzianti (2022) emphasised the need for e-waste collection facilities and information campaigns to raise awareness, as e-waste is often found in general waste bins (Wanderley & Bonacin, 2019). Jain *et al.* (2023) and Rive (2017) noted that e-waste is a complex problem that requires innovative solutions. Previous studies have focused on technical solutions or interface design prototypes for e-waste management; therefore, the current study focuses on mapping stakeholder pain points as the basis for an ecopreneurial business opportunity.

E-waste issues require new solutions. An innovative e-waste disposal method could be proposed. This field could have an ecopreneur enterprise. Design thinking can be used to construct this system, as Inegbedion (2022) found that it improves corporate success. Therefore, this research advances ecopreneurship entrepreneurial theory by applying design thinking to e-waste management. Understanding the perspectives and experiences of these e-waste stakeholders could help us identify their genuine concerns.

1.1. Literature Review

End-of-life electrical and electronic equipment (e-waste) is a rapidly expanding global problem. E-waste comprises both electronic equipment that is used but no longer functions and can still be repaired, and electronic equipment that is used but no longer functions and cannot be repaired (Basel Convention, 2013). Such waste generally contains valuable materials that are economically valuable when recycled. Unfortunately, most e-waste is recycled in the unregulated informal sector, posing a significant risk of toxic exposure to recyclers (Perkins *et al*., 2014). E-waste can be categorised as (Jafar, 2015):

(i) Technology and communications: Laptops, PCs, earphones, headphones, scanners, data cables, cameras, chargers, power banks, loudspeakers, etc.
(ii) Electronic entertainment: Cameras, video recorders, radios, video players, audio players, remote controls, etc.
(iii) Electronic toys and sports equipment: Toy cars, toy musical instruments, etc.
(iv) Household: Rice cookers, irons, electronic kettles, toasters, stirrers, whisks, etc.
(v) Lighting: Light bulbs and light tubes
(vi) Electrical power cables, batteries, extension cables, etc.

According to Nnorom and Osibanjo (2008), changes in government attitudes, the enactment of laws specifically addressing e-waste, control over e-waste disposal, the implementation of Extended Producer Responsibility, and the transfer of technology for proper e-waste recycling are key criteria for effective e-waste management in some emerging economies. Sundar *et al.* (2023) suggested that the government provide incentives to stakeholders involved in e-waste handling. Nowakowski and Pamuła (2020) proposed the design of an application that can identify e-waste types based on photos sent by people seeking to dispose of e-waste. There is also the idea of implementing mobile robots that identify e-waste during the collection process from households, thereby reducing the need for human labour (Shreyas Madhav *et al*., 2022).

Brindhadevi *et al.* (2023) discussed numerous present and emerging e-waste recycling systems; however, their primary focus was on the health effects of e-waste metal pollution. Yu *et al.* (2023) suggested optimising the e-waste recycling industry using legal, reasonable, and efficient methods. George and Michael (2023) proposed the idea of creating e-waste-based crafts and using a platform to sell them. Overall, e-waste studies conducted between 2008 and 2023 have shown that controlling e-waste disposal in emerging economies is hampered by several key issues.

Design thinking is a comprehensive thinking process that focuses on generating market possibilities and solutions, starting with a process of empathy for a specific human-centered need and progressing to sustainable innovation based on user demands (Brown, 2009). In this method, five stages/processes enable us to obtain innovative outcomes. These stages aim to explore user needs and determine optimal product specifications (Plattner, 2010). The five stages of design thinking are empathise, define, ideate, prototype, and test.

In the first stage, empathise, the design thinker must identify the user or intended user and then analyse the user's experience, emotions, and situation. Empathising is essential in design thinking because it ensures that the final product or service is user-centred and meets the needs of the target user. Regarding the empathy results, an empathy map is a commonly used tool to better understand the perspectives and experiences of a particular group of people. It is often employed in the fields of design and customer experience to help organisations better grasp their clients' or users' requirements and feelings. According to Bratsberg (2012), an empathy map is a human-centred tool that focuses on understanding individuals by seeing the world from the user's perspective. Four of the five stages (empathise, define, prototype, & test) are forecasters of business success (Inegbedion, 2022).

Several other studies on e-waste have applied design thinking. Nuryanto and Suzianti (2022) utilised design thinking to design and test the elements of an e-waste application, while Wanderley and Bonacin (2019) employed the Hackathon method to create a recycling bin prototype. Fanthi *et al.* (2021) used e-waste for home interior accessories. Rive (2017) highlighted the use of e-waste for installation and performance art, and Madrigal *et al.* (2024) utilised design thinking specifically for photovoltaic panels.

Other waste-related studies have also used design thinking. Specifically, Jimenez *et al.* (2021) described a sterilisation system; Rois *et al.* (2020) applied design thinking to organic waste management; Machado and Grilo (2020) focused on a general waste collection system; and Massari *et al.* (2022) used design thinking to examine food waste management.

2. Methods

As this study focused on the empathise stage of the design thinking method, it is considered a qualitative research and thus aims to present the results descriptively. According to Sekaran and Bougie (2016), descriptive research seeks to describe an interesting topic and explain the characteristics of relevant objects such as people, products, and brands. In this study, at the empathise stage, we seek to capture the characteristics of stakeholders in managing e-waste. The needs of these parties should be understood in the final phase. Based on this research strategy, this study was categorised as a case study. Case studies examine something in depth and provide a more precise explanation of the problem by examining real situations (Sekaran & Bougie, 2016).

A case study was conducted by selecting a city on Java Island, Indonesia, as the research object. Currently, this city has only six e-waste drop points (Citarum Harum, 2022). Next, we selected the parties to be interviewed using purposive sampling. Purposive sampling is based on certain criteria to support the information needed (Sekaran & Bougie, 2016). The participants were stakeholders in e-waste disposal management, comprising several parties, including e-waste disposers, government institutions, and recycling institutions. E-waste disposers were represented by a housewife, a father, a college student, a factory manager, a shopping mall manager, and a logistics manager of an educational institution. The government institutions were represented by a staff member and the manager of the Environmental and Sanitation Department. E-waste recycling organisations were represented by the director of a waste bank and the manager of a recycling company. In total, there were 10 participants.

The interview method was used to collect data. The interviews were conducted in a semi-structured manner, following verbal confirmation of informed consent from the participants. We compiled a list of interview questions for the three types of stakeholders. The questions consisted of three parts: the stage prior to e-waste disposal, the stage of e-waste disposal, and the stage after e-waste disposal.

The questions are presented below:

(i) Questions for the environmental agency

- Before disposal:

a. To what extent does the environmental agency understand electronic waste and its management?
b. Does the environmental agency have guidelines or regulations regarding the disposal, recycling, and processing of electronic waste?
c. What should be done with damaged electronic devices?
d. What kind of items are considered electronic waste?
e. How much electronic waste is generated in your work area per month?
f. What do you usually do with electronic waste in your area?
g. What are your feelings about electronic waste disposal?
h. Are there any campaigns or programs regarding electronic waste? What are they?
i. What are the obstacles facing in managing and educating about electronic waste?

- During disposal:

a. How does the environmental agency dispose of electronic waste?
b. What is the current process for people disposing of electronic waste?
c. Is there a specific time for you to dispose of electronic waste? When is that?
d. Where do you dispose of electronic waste?
e. Who should be responsible for managing electronic waste?
f. How should electronic waste be managed?

- After disposal:

a. How do you feel about the way people currently dispose of e-waste? Satisfied/dissatisfied? Explain.
b. Do you know what happens to discarded e-waste? Explain.
c. Are you aware of the dangers of improperly disposed of e-waste? Explain.
d. Does e-waste have economic value? Explain.
e. How do you measure the success of your e-waste recycling efforts?
f. What are the long-term and short-term goals of the environmental agency in managing e-waste?
g. Does the collected and processed e-waste comply with applicable environmental and safety standards?

(ii) Questions for e-waste disposers:

- Before disposal:

a. To what extent do you understand electronic waste?
b. What do you do with broken electronic devices?
c. What are the conditions of the items you dispose of as electronic waste?
d. How much electronic waste is there in your area per month?
e. What do you usually do with the electronic waste in your area?

- During disposal:

a. How do you dispose of e-waste?
b. Is there a specific time you dispose of e-waste? When is that?
c. Where do you dispose of e-waste?
d. Who should be responsible for managing e-waste?
e. How should e-waste be managed?

- After disposal:

a. How do you feel about the way you currently dispose of e-waste? Satisfied/unsatisfied? Explain.
b. Do you know what happens to the e-waste you dispose of? Explain.
c. Are you aware of the dangers of improperly disposed e-waste? Explain.
d. What do you expect to happen after disposing of e-waste?

(iii) Questions for recyclers:

- Before disposal:

a. How well do you understand electronic waste?
b. What should be done with damaged electronic devices?
c. What kind of items are considered electronic waste?
d. Are there any campaigns about electronic waste? How?
e. Are there any certifications or standards that recycling managers must meet to ensure proper recycling?

- During disposal:

a. What are the procedures for disposing of e-waste?
b. How is the community currently disposing of

e-waste?
c. Is there a specific time for people to dispose of e-waste? When is that?
d. Where do people dispose of e-waste?
e. Who should be responsible for managing e-waste?
f. How should e-waste be managed?

- After disposal:

a. Considering discarded electronic waste, how do you think it is managed? What do you think?
b. Do you know what happens to the electronic waste discarded by the public? Explain.
c. What happens to the electronic waste you manage? If it cannot be processed, what do you do?
d. What are the economic benefits of the electronic waste you manage?
e. What recycling process is considered feasible? Is it implemented by your team? What will you do if it fails?
f. What response does the recycler expect from the government and the public regarding e-waste recycling?
g. How long does it take the recycler to manage e-waste?
h. How is the success of e-waste management calculated?

Case studies often use three validity tests to verify study quality: external validity, internal validity, and construct validity (Yin, 2014). We examined the construct and external validity. Internal validity was examined only in explanatory or causal research; hence, we did not include it in our study. Three tactics can be used to increase the construct validity when conducting a case study. The first uses evidence from several sources (Yin, 2014). The data for this study were gathered from a number of sources, including physical artefacts, interviews, documentation, and archival documents. Creating a chain of evidence is the second strategy for construct validity. Most of the evidence in this study was cited and treated as a book or scientific article; therefore, it is also listed in the references. Finally, the third tactic is to have a draft report reviewed by key informants. The second test was external validity. The only way to increase external validity is through replication. This is reflected in the method used to select the sources for this study.

Yin (2014, p. 240) defines reliability as "the consistency and repeatability of research procedures used in case studies". Using a consistent approach ensures the trustworthiness of outcomes. When another researcher conducts the same case study using the same methodology as the previous researcher, they should obtain the same results and conclusions. Documentation of past study techniques is essential and can be accomplished using a case study protocol. In this study, the collected data were grouped according to their answers. The aim was to avoid errors during data collection. In qualitative research utilising interview transcripts, processes for validity (credibility) and reliability (dependability) are implemented to guarantee the trustworthiness and consistency of the findings derived from the data. These are the processes:

(i) Credibility (validity) protocol: The principal method employed in this analysis to ensure that the interpretations accurately represent the meanings expressed by interview participants is triangulation. Specifically, this study utilised source triangulation, which entails the comparison and contrast of information obtained from various, distinct sources. For example, the notion of inadequate infrastructure mentioned by household consumers was contrasted with reports from the Environmental and Sanitation Department indicating the presence of infrastructure. The triangulation indicates that although infrastructure is present in some locations, household awareness and accessibility continue to pose significant obstacles. Another example is analysing the financial perspective on e-waste: waste banks perceive e-waste as a cost/loss, while shopping malls regard e-waste disposal as a paid service. These views were then compared with the economic opportunity perspective, where factories transfer e-waste to generate value elsewhere, recycling organisations pursue repair/reuse, and the informal sector extracts gold/copper.

(ii) Dependability and confirmability (reliability) procedures: To demonstrate that the analysis is coherent and based on the data, the following processes were employed:

a. Comprehensive documentation (audit trail): The analysis is only based on the supplied transcripts, ensuring a thorough record of the original data. The background information, encompassing the interviewers' identity, date, and the interviewees' unique function, provides a robust context for the data.

b. Systematic coding and categorisation: The method of categorising data into specific themes (definition, practices, and challenges) through consistent comparison guarantees that the conclusions are produced systematically.

c. Direct quotation and citation: Each derived assertion is explicitly substantiated by citations utilising the source index. This commitment to source ensures confirmability, illustrating that the interpretations are based on the participants' own words and reducing researcher bias.

d. Thick description: Detailed accounts of the viewpoints of many roles—including students, fathers, the director of a waste bank, the manager of a factory, and the staff members of the Environmental and Sanitation Department—offer context for evaluating the applicability of the findings to analogous environments.

Lastly, according to Miles and Huberman (1994), qualitative data analysis has three stages: data reduction, data presentation, and conclusion.

3. Results

In the empathise stage, we used interviews for data collection to understand user needs in depth. The 10 stakeholders who were interviewed were classified into three categories. After all the interview audio clips were transcribed, the first step was to reduce the data. Unnecessary data were removed, and we retained only relevant and contextual data. The triangulation technique evaluated the viewpoints of home customers, institutional/corporate managers, and government/recycling organisations on certain essential themes as follows.

3.1. Definition and Comprehension of E-waste

All participants recognised e-waste as discarded or malfunctioning electronic devices; nevertheless, the understanding of its hazardous characteristics differs markedly by role, as presented in Table 1.

3.2. Practices for the Disposal and Management of E-waste

Practices vary significantly depending on the degree of institutional commitment and availability of regulated infrastructure, as shown in Table 2.

3.3. Obstacles and Systemic Deficiencies

The principal obstacles arise from insufficient public knowledge, inadequate infrastructure accessibility, and the absence of regulatory enforcement, notably the Extended Producer Responsibility policy. Details are as follows:

(i) Insufficient information and accessibility (household perspective): Consumers express confusion or apathy toward proper disposal due to a dearth of information. They observe an absence of designated e-waste drop-off sites in proximity to their residences. Their assertion indicates that the management system for home hazardous and toxic materials is suboptimal and necessitates individual effort.

(ii) Economic disincentives (waste bank/the Environmental and Sanitation Department perspective): The disposal of e-waste frequently constitutes a cost centre; in contrast to plastics

Table 1. Definition of E-waste by group

Stakeholder group	Key understanding
Household consumers	Define it as discarded or non-functional products (e.g., PCs, obsolete televisions, and damaged cables/earphones). They frequently combine minor e-waste components (batteries and tiny cables) with conventional refuse. They may be deficient in particular knowledge regarding hazardous content or its associated hazards. For example, the college student stated, "electronic waste as items that are no longer used, such as damaged goods, old computers, old televisions."
Institutional or corporate managers	Identifies e-waste as a category that necessitates specialized management due to its toxic characteristics. At the educational institution, e-waste is categorised as either non-organic (recyclable) or hazardous. The mall manager categorically classifies it as hazardous garbage, encompassing batteries and bulbs. The factory manager is aware that improper disposal constitutes hazardous waste. He said, "e-waste is trash that can be recycled by competent parties, and it is considered a hazardous waste if disposed of carelessly."
Government/Industry	Categorises e-waste as hazardous and toxic material garbage that must not be disposed of in standard household temporary disposal sites. The scope encompasses components, batteries, cables, and big appliances such as televisions, washing machines, and PCs. The waste bank director said, "E-waste is the residue from electronic goods that are no longer used."

Table 2. Disposal management of E-waste

Stakeholder group	Disposal technique
Household consumers	Primarily entails the accumulation of damaged things, the amalgamation of minor objects (such as cables, batteries, and mice) with ordinary refuse, or the transfer of substantial items to informal scavengers or hauliers. They frequently see the transfer to scavengers as a payment for disposal services. Disposal is unplanned and reactive. A housewife said, "Small items like used batteries are simply thrown into the regular waste bin." A father added, "If an item is beyond repair, large items (like old televisions or washing machines) are often given away to rag-and-bone men or the regular trash collector."
Institutional or corporate managers	Establishes a systematic internal procedure. The educational institution accumulates and retains e-waste (often for two years) before auctioning it to specialised collectors once the asset value has diminished to zero. The factory provides components to the cleaning personnel, who autonomously search for low-current electronic workshops for repair or reuse. The mall employs certified third-party suppliers for hazardous and toxic garbage collection, a fee-based service determined by volume and scheduled pickups. For example, the mall manager explained: We must operate a temporary storage site for hazardous and toxic waste and [be] registered with the Environmental and Sanitation Department. This storage requires specific infrastructure, safety procedures, and personnel. Later, the mall must then pay certified third-party vendors to collect and process the hazardous and toxic waste. The academic institutional official said, "We require a formal document to nullify the asset's value before disposal. Hence, we prioritise repair and cannibalisation (using parts from broken equipment)."
Government/Industry	Offers specialised infrastructure (drop box sites in public spaces). The Environmental and Sanitation Department provides complimentary online scheduling for the collection of large household items, excluding commercial zones. E-waste gathered by this department is dispatched to authorised transporters and processors for disposal. The manager of the Environmental and Sanitation Department said, "We established drop boxes in public areas like malls, schools, and transit stations. In Jakarta, there are 26 such locations. We also offer a free online e-waste pickup service for non-commercial residents who meet certain criteria (e.g., minimum weight)." A recycling company emphasises repair and reuse through service partners and deliberately refrains from collecting particularly hazardous goods such as batteries. The recycling company manager explained, "We accept electronics (televisions, fans) but refuse batteries and highly hazardous waste [as] they cannot safely process."

or cardboard, it lacks immediate economic value for the waste bank, necessitating a concentration only on collection and teaching. The department administers e-waste management as a public service, funded by the government.

(iii) Regulatory and policy deficiencies (government/industry perspective): Principal government agencies emphasise that adherence and enforcement are deficient. The implementation of the Extended Producer Responsibility policy, which mandates producers to reclaim their waste, is ineffective. The Environmental and Sanitation Department notes that 99% of the e-waste produced is not managed in accordance with regulations. A significant obstacle is the federal framework in which regional departments possess restricted jurisdiction over permits, which are governed by the national authority.

(iv) Informal sector risks: The informal sector significantly contributes to recycling, particularly for substantial commodities such as televisions. This frequently entails hazardous deconstruction processes to retrieve precious metals (e.g., gold, copper, and lead) without adherence to safety protocols, resulting in significant health and environmental repercussions (e.g., congenital anomalies in specific communities).

Furthermore, empathy maps were used as the framework to present the data. This tool helps understand the perspectives and experiences of the interviewed stakeholders. The following presents the empathy maps of all 10 participants:

(i) Housewife: The interview with the housewife revealed that she already had some awareness of e-waste, but only a little. She had heard that e-waste could not be recycled in the same way as other waste types, and that developed countries have specialised e-waste disposal systems. She understood that e-waste must be disposed of specifically; however, she still disposed of small amounts of e-waste in the general waste bin. She acknowledged her limited awareness and actions in sorting waste due to a lack of information. She suggested that it is necessary to award points to individuals who properly handle e-waste. The highlights of the housewife's interview are presented in Figure 1.

(ii) Father: Based on the interview, the father had less knowledge than the housewife regarding e-waste. He still assumed that e-waste had some value that could benefit the waste collector or the junkman. He was unaware of the dangers associated with e-waste. The highlights of the father's interviews are presented in Figure 2.

(iii) College student: The college student observed that it was easy to produce e-waste. If she experienced a minor malfunction in an electronic item, she considered it e-waste. She had little knowledge about e-waste; for instance, she believed that e-waste could be disposed of in general rubbish containers. She thought that e-waste would be collected by waste collectors/junkmen. The highlights of the college students' interviews are presented in Figure 3.

(iv) Factory manager: The factory manager seemed to treat e-waste as a source of company income. He had a policy of offering e-waste to other parties, such as electronic repair shops, so that some components could be sold and reused. However, he realised the dangers of e-waste and knew that it could not be disposed of carelessly. The highlights of the factory manager's interviews are presented in Figure 4.

(v) Shopping mall manager: Based on the interview, the manager was the interviewee with the most awareness in the e-waste disposal category. He was the only one who was convinced that e-waste created costs for businesses. He had already treated e-waste properly, including hiring a vendor to transport his mall's e-waste. The highlights of the shopping mall manager's interview are presented in Figure 5.

(vi) Logistics manager of an educational institution: The logistics manager lacked knowledge about e-waste. She had created a comprehensive procedure for waste management, but still mixed e-waste (lamps/bulbs) with other glass waste. The

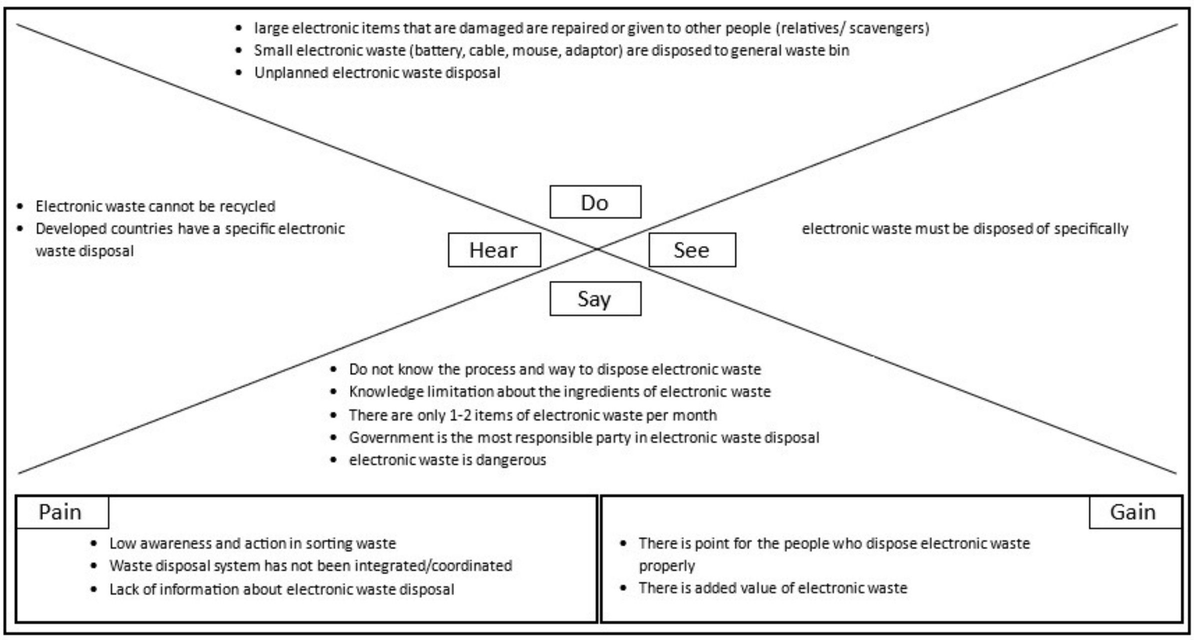


Figure 1. The empathy map of the housewife

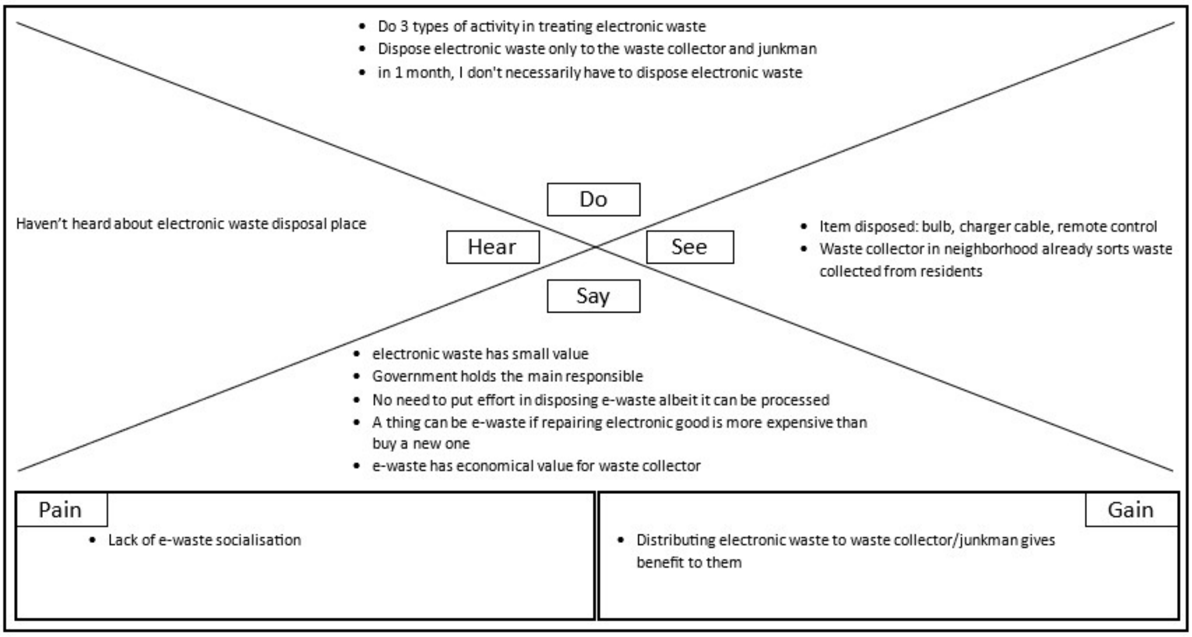


Figure 2. The empathy map of the father

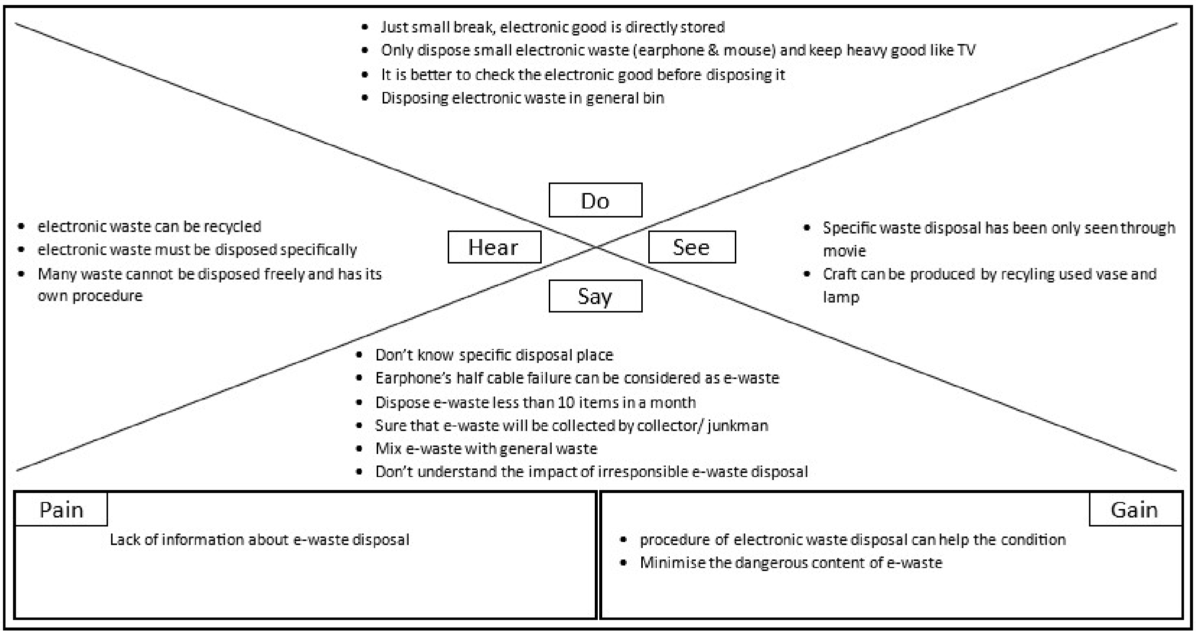


Figure 3. The empathy map of the college student

highlights of the logistics manager's interview are presented in Figure 6.

(vii) Staff member of the Environmental and Sanitation Department: The staff member mentioned that the Environmental and Sanitation Department has several e-waste management programmes and extensive knowledge of how to handle e-waste. However, there is no integrated data on e-waste processes. The highlights of the staff member's interview are presented in Figure 7.

(viii) The manager of the Environmental and Sanitation Department: The Environmental and Sanitation

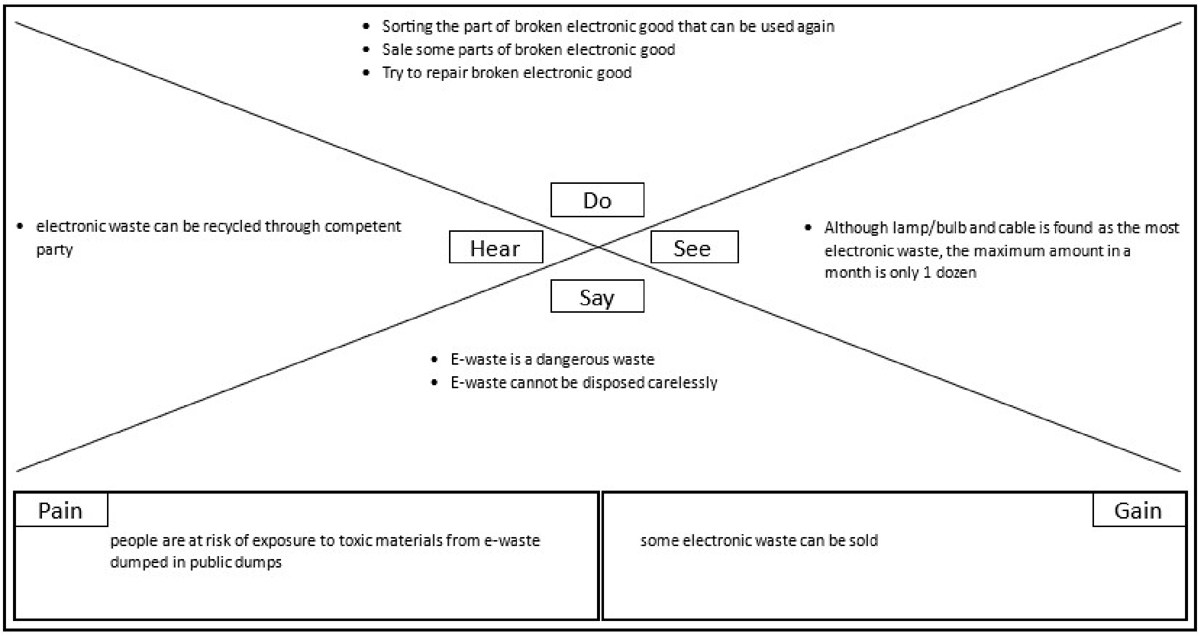


Figure 4. The empathy map of the factory manager

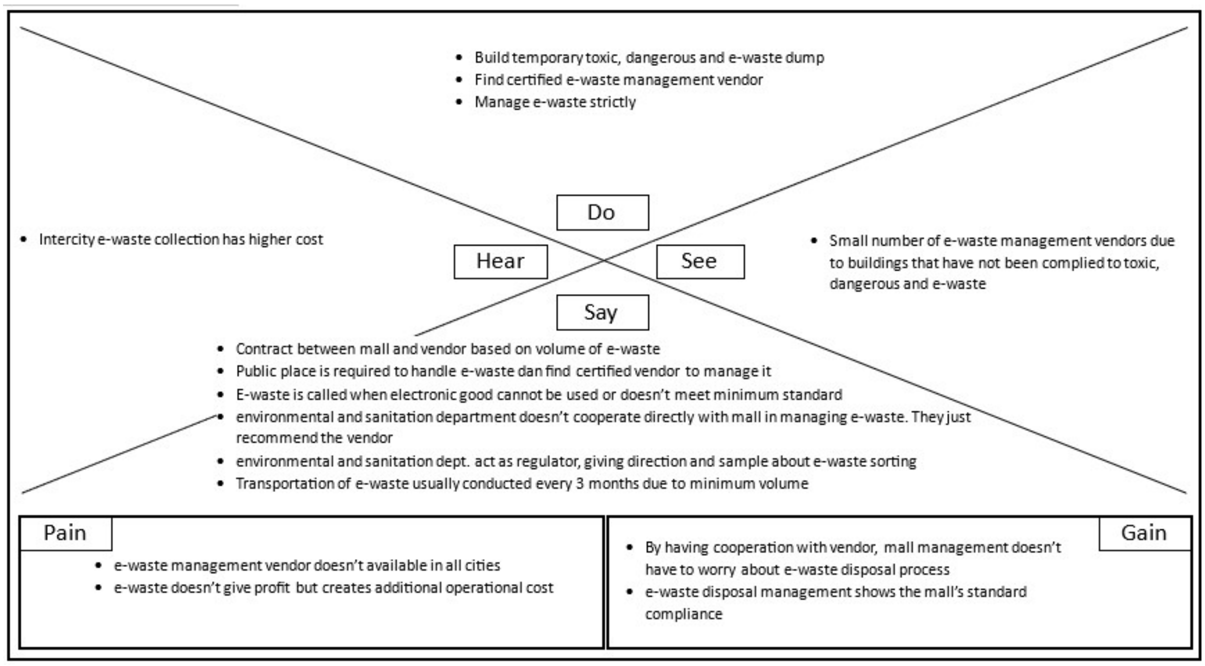


Figure 5. The empathy map of the shopping mall manager

Department of Bandung city has a more advanced e-waste treatment than other cities. However, it still depends on the private sector for the disposal of e-waste. The manager acknowledged the lack of a waste sorting process at the household level. The highlights of the interviews with the manager of the Environmental and Sanitation Department are presented in Figure 8.

(ix) Waste bank director: As the director of a waste bank, the participant confessed that the

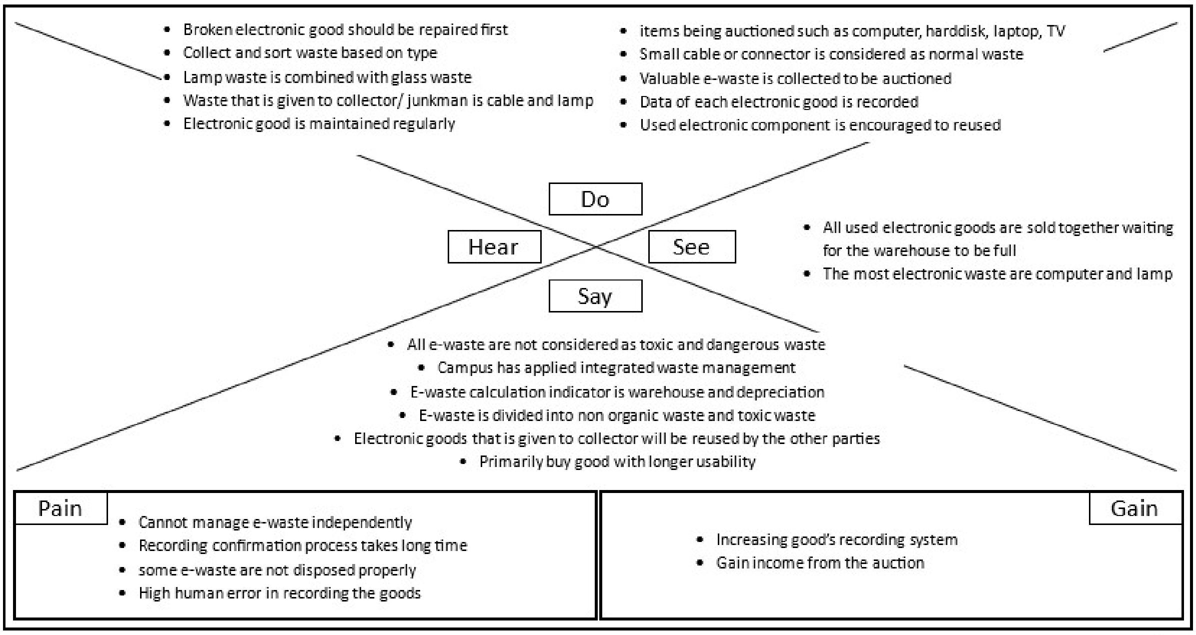


Figure 6. The empathy map of the logistics manager of an educational institution

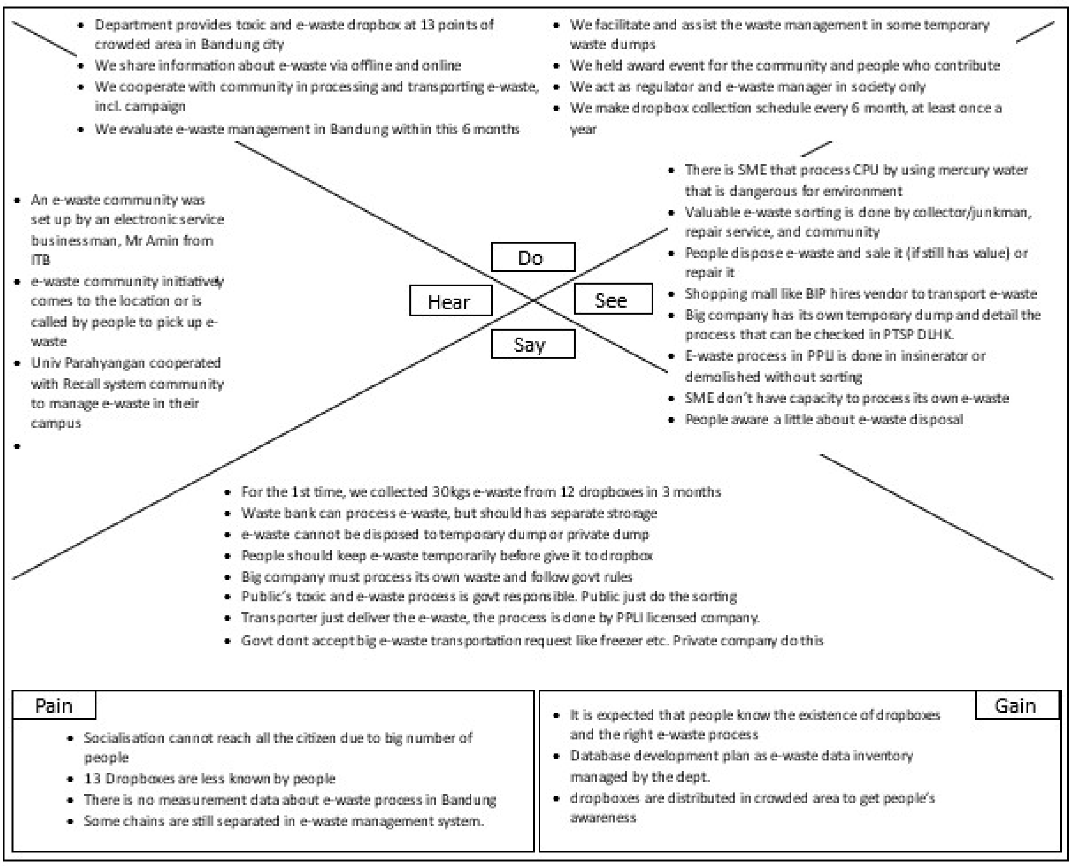


Figure 7. The empathy map of the staff member of the Environmental and Sanitation Department. Abbreviations: BIP: Bandung Indah Plaza, DLHK: Dinas Lingkungan Hidup dan Kebersihan, ITB: Institut Teknologi Bandung, PPLI: Prasadha Pamunah Limbah Industri, PTSP: Pelayanan Terpadu Satu Pintu; SME: Small- and medium-sized enterprise.

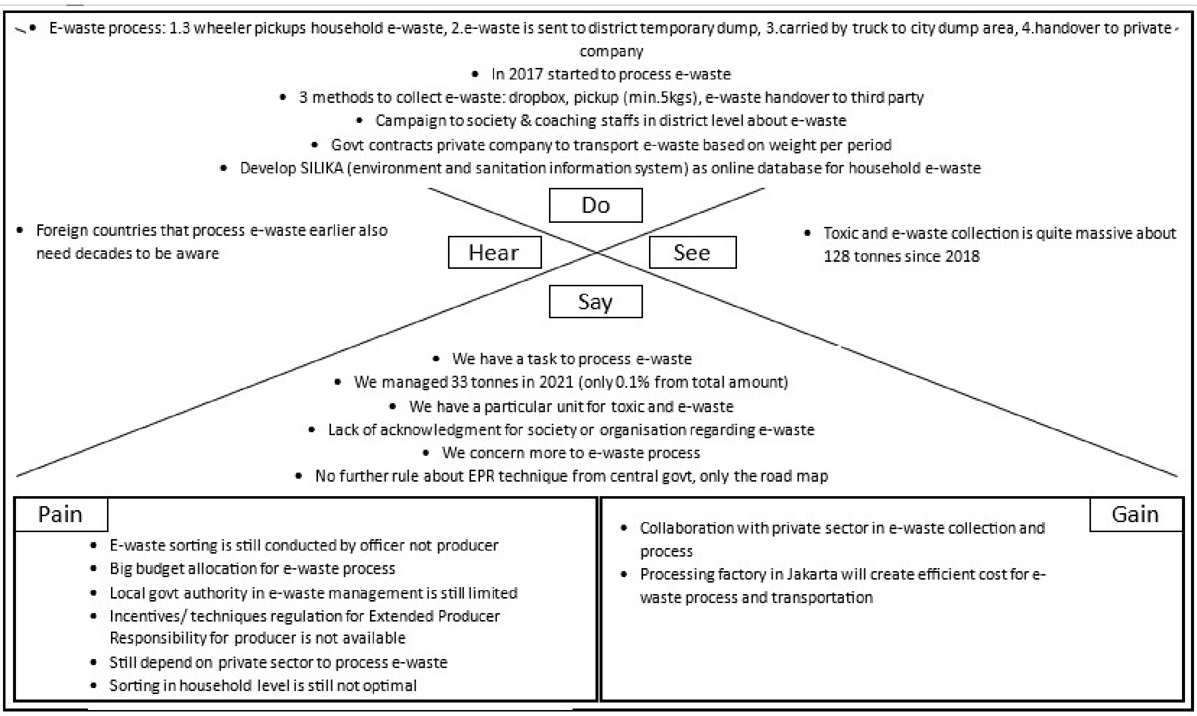


Figure 8. The empathy map of the manager of the Environmental and Sanitation Department. Abbreviation: EPR: Extended Producer Responsibility

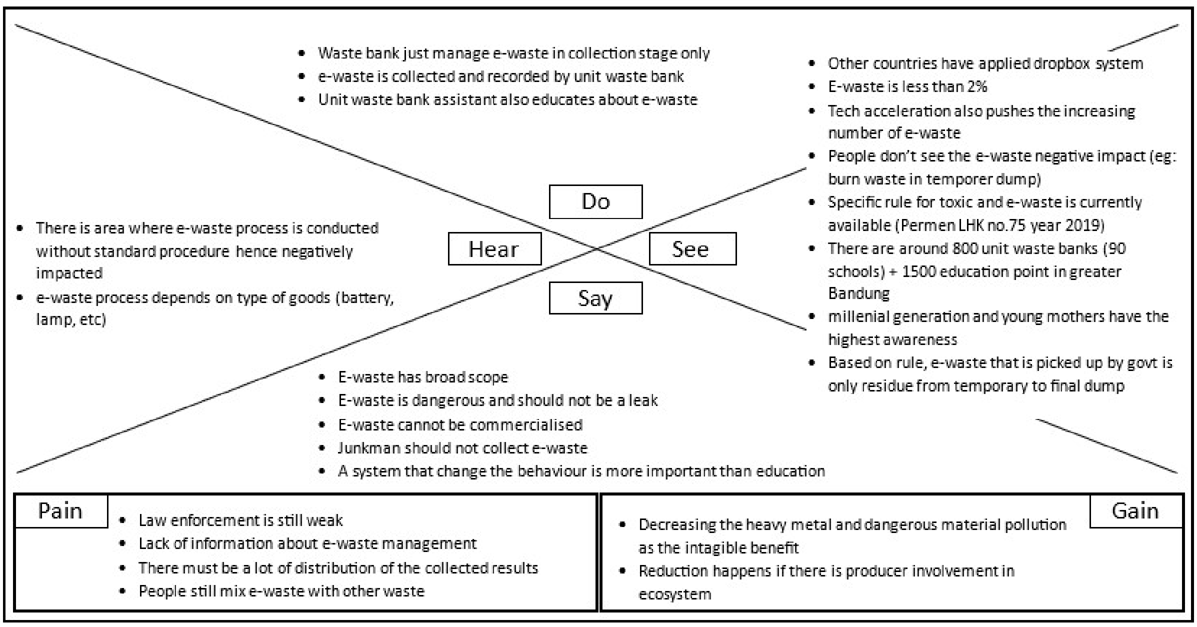


Figure 9. The empathy map of the waste bank director

bank did not wield considerable authority in e-waste disposal management. E-waste must be transferred to a certified e-waste processing facility. She understood the lack of knowledge and legal enforcement experience in the city. The highlights of the director's interview are presented in Figure 9.

(x) Recycling company manager: As the leader of an active recycling company, the manager often finds it challenging to determine the next steps for the

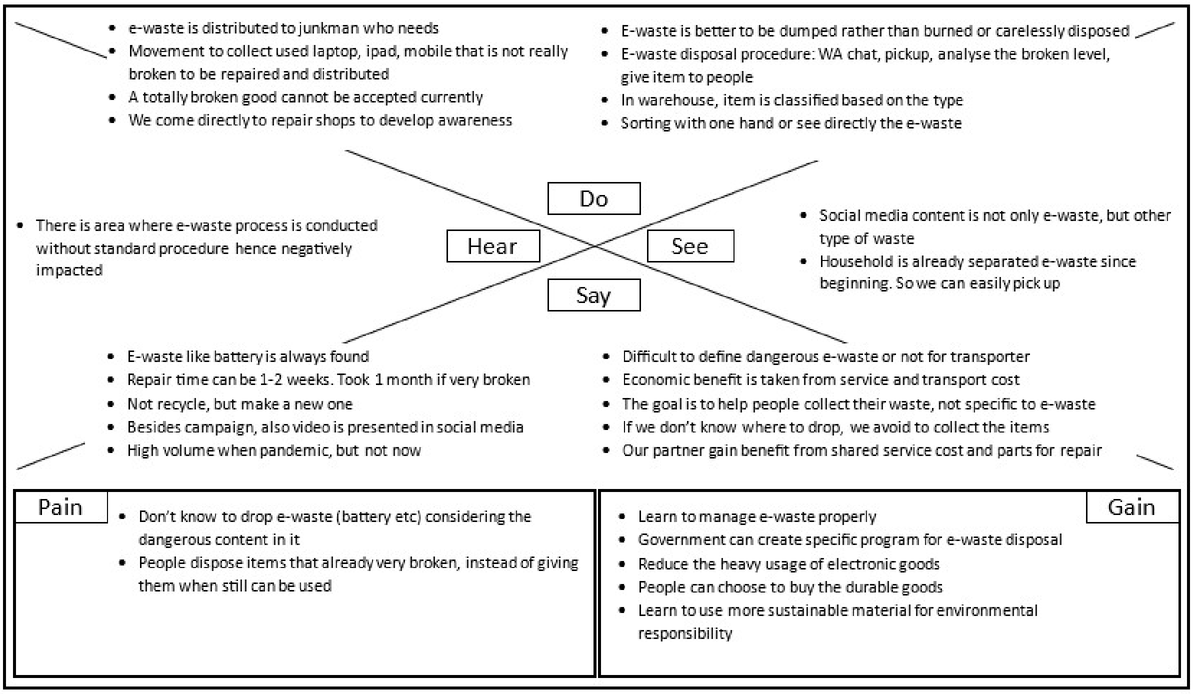


Figure 10. The empathy map of the recycling company manager

collected e-waste. Although their goal is to assist people with proper waste disposal, they still lack information about the suitable drop-off locations for e-waste. The highlights of the recycling company manager's interview are presented in Figure 10.

4. Discussion

This study used a formal opportunity recognition method, systematically transforming complex, "wicked problems" such as e-waste into actionable opportunities (Jain *et al.*, 2023; Rive, 2017; Wanderley & Bonacin, 2019). After developing the empathy maps based on the 10 interviews, the design thinking process progressed to the "define" step. We identified the challenges experienced by participants by grouping their pain points. Consequently, we focused on the pain section of the empathy maps, compiling a list of difficulties experienced by participants through the identification and categorisation of specific problems. An affinity diagram was used to organise and analyse the data obtained during the empathise stage, as shown in Table 3.

As shown in Table 3, the 25 pain statements were divided into four categories. The category "System is not good" contains 11 statements. As this category has the most pain statements, it is considered the priority category (marked with number 1). The category "Less socialisation/information" consists of seven pain statements. As this category has the second-most statements, it is considered the second-priority category. The category "Low awareness of the community" consists of five pain statements. This category is considered the third-priority category. Finally, the category "Expensive processing cost" consists of two pain statements. This category is regarded as the fourth-priority category. These are the four main challenges in e-waste disposal management. The first three categories align with the management process issues and public knowledge identified by Nuryanto and Suzianti (2022). However, they focused directly on the design elements of the application and tested the interface prototype. By contrast, this study focuses on assessing the challenges faced by e-waste stakeholders and generating ideas to address these problems.

The "System is not good" category is the primary e-waste problem, confirming findings by Jain *et al.* (2023), Nnorom and Osibanjo (2008), and Parajuly *et al.* (2019). The global lack of effective e-waste management, including insufficient e-waste processing centres, is particularly pronounced in developing countries. Recognising this structural failing as the main pain point drives future ecopreneurial activity toward large-scale infrastructure solutions rather than superficial tweaks, revealing where investment and innovation are required most (Nuryanto & Suzianti, 2022). To develop an improved system for e-waste management, inspiration can be drawn from the works by

Table 3. Affinity diagram from the pain section in empathy maps

Category	Housewife	Father	College student	Factory manager	Shopping mall manager	Logistic manager	Staff at the government office	Manager at the government office	Recycling company manager	Waste bank director
Low awareness of community (3)	Low awareness and action to sort e-waste	-	-	-	-	Some e-waste is not disposed of properly	-	E-waste sorting is still carried out by officers and not by waste producers	Electronic goods that are still functional are not being utilised by those who need them	People still mix e-waste with other types of waste
System is not good (1)	The waste management system is not yet integrated (coordination)	-	-	-	Not all cities have vendors who manage e-waste	- Not yet able to manage e-waste independently - High level of human error in recording goods - Long recording confirmation process	-	- The authority of local governments in managing B3 waste is still limited - Still rely on third parties as management institutions - There are currently no incentive mechanisms or technical regulatory systems for producers under EPR	There is uncertainty about where to distribute e-waste containing hazardous materials	- The collected results need to be widely distributed - Enforcement of regulations is still weak
Less socialisation/ information (2)	Information on e-waste disposal is still lacking	Lack of socialisation of e-waste	Lack of information about e-waste disposal	E-waste thrown into public dumps will endanger people	-	-	Socialisation has not reached all communities due to their large number	-	People do not know much about e-waste disposal	Lack of information about e-waste management
Expensive processing costs (4)	-	-	-	-	E-waste does not provide income but instead adds to operational costs	-	-	The need for a large budget allocation for processing services	-	-

Note: Numbers in brackets under "category" refer to the priority level.
Abbreviations: E-waste: electronic waste; EPR: Extended Producer Responsibility.

Nowakowski and Pamuła (2020), Shreyas Madhav *et al.* (2022), and Wanderley and Bonacin (2019), who utilised the latest technology in e-waste management. This aligns with our intention to incorporate technology into e-waste management. Technology-based educational and logistical tools, such as mobile applications and smart recycling bins, can address the issues of "less socialisation/information" and "low awareness of the community" as confirmed by external research (Nuryanto & Suzianti, 2022; Wanderley & Bonacin, 2019). Moreover, the "Low awareness of community" issue has been acknowledged by Iqbal and Dyah (2022), Khalid (2022), and research conducted by the United Nations Institute for Training and Research (Widodo, 2023).

The issue of "expensive processing cost" in e-waste management confirms the findings by George and Michael (2023) and Yu *et al.* (2023), which highlight that the e-waste recycling industry is still inefficient. Moreover, a lack of returns and profits is mentioned as a barrier to e-waste management in the United Kingdom (Sundar *et al.*, 2023). These high processing costs can be compared with the potential economic advantages of e-waste, as calculated by doctoral candidates specialising in e-waste recycling (Dinnata, 2017; Puspa, 2023). Recently, the recycling process has been enhanced by a tool that enables electronic designers to incorporate recycled components into their designs for electronic appliances (Lu *et al.*, 2023). In addition to its economic advantages, e-waste recycling has the potential to reduce carbon emissions. Realising this potential and raising awareness within broader communities could help address the problem of "less socialisation/information." To tackle this, we applied the interdisciplinary learning approach proposed by Wang and Anand (2024). In summary, this study confirms the findings of previous studies from 2008 to 2024, demonstrating that the disposal of e-waste in developing countries continues to face several challenges.

The contribution of this research lies in applying design thinking specifically to e-waste disposal management from an ecopreneurial perspective. Design thinking was employed throughout this research, validating its effectiveness in addressing complex sustainability concerns related to human behavior. Design thinking excels at solving unknown challenges (Design Foundation, 2018; Fontichiaro, 2015; Schallmo, 2018). This study also confirms previous research showing the utility of design thinking in e-waste contexts, such as designing application interfaces or conceptualising material reuse, solidifying design thinking as a framework for managing technology, human behaviour, and disposal (Fanthi *et al.*, 2021; Madrigal *et al.*, 2024; Nuryanto & Suzianti, 2022; Rive, 2017; Wanderley & Bonacin, 2019).

The novelty of this study lies in utilising design thinking exclusively during the "empathise" and "define" phases to delineate and classify stakeholders' pain points. Although the results of this study are descriptive and do not introduce a framework, model, or actionable system for practitioners, they focus on mapping stakeholder pain points as the basis for an ecopreneurial business opportunity, differing from other studies that employ design thinking primarily to prototype and test user interface designs. For example, earlier research used design thinking to rapidly design and test user interface prototypes for e-waste applications (Nuryanto & Suzianti, 2022) or to create physical prototypes such as smart recycling bins (Wanderley & Bonacin, 2019). In contrast, this study evaluated the actual challenges encountered by stakeholders prior to the formulation of solutions.

Moreover, this study aims to enhance entrepreneurship theory within the ecopreneurship domain by addressing the existing research gap on effective, technology-driven e-waste disposal management from a business management perspective (Inegbedion, 2022). The principal finding identifies four fundamental issues that constitute the foundation for developing a future ecopreneurial, information and communication technology (ICT)-driven e-waste management system.

Design thinking is particularly important in developing economies, where management often has trouble with basic methods, weak enforcement, and low public awareness (Nnorom & Osibanjo, 2008; Parajuly *et al.*, 2019). This study establishes the fundamental framework necessary to address this intricate, globally acknowledged issue by affirming the principal challenges (such as the deficient system and limited awareness) within this context. This methodological framing is relatively unique compared to typical e-waste studies that focus on technical solutions or environmental impacts (Nowakowski & Pamuła, 2020; Shreyas Madhav *et al.*, 2022; Wanderley & Bonacin, 2019).

This study develops a systematic framework for ICT-based e-waste management solutions, advancing ecopreneurship. Ecopreneurship integrates sustainability and profit (Parajuly *et al.*, 2019; Rive, 2017). The findings identify market failures and outline the criteria that a viable entrepreneurial firm must address:

(i) Addressing market failure (the system gap): The "system is not effective" category suggests that e-waste is a rapidly growing, unmanaged stream. This gap presents a huge market opportunity for sustainable businesses (Forti *et al.*, 2020). Only 17.4% of e-waste is collected and recycled, leaving a massive opportunity for entrepreneurs (Harikaran *et al.*, 2023).

(ii) Monetising environmental value (economic gain): E-waste's revenue potential boosts ecopreneurship's financial side, dispelling "expensive processing costs." It has been estimated that recycling e-waste might earn Indonesia a total of USD 1.8 billion by 2020 (Puspa, 2023). E-waste from telecommunications contains valuable elements such as gold, silver, and copper, and managing one tonne of it can prevent 1,400 tonnes of CO_2 emissions (Dinnata, 2017). Ecopreneurial efforts are motivated by financial incentives and environmental benefits.

The limitation of this study is its coverage. This study only focused on the early stages of design thinking: empathise and define. Therefore, future research is advised to progress to the "ideate stage," utilising the four primary issues identified as a framework. Considering the objective is an ICT-driven system, the solution created in the "prototype stage" would presumably be digital and logistical in nature. Then, the prototype will be evaluated in the "test stage" with representatives from the three primary stakeholder categories: disposers, government, and recyclers. Another limitation is in its location. The study focused on a singular metropolitan context, implying that the recognised issues and subsequent challenges—such as the deficient system, insufficient social interaction, and minimal public awareness—are explicitly linked to the local regulatory, logistical, and social milieu of that city. Therefore, it is also suggested to study a broader context to gain different perspectives.

5. Conclusion

This study classified 25 pain statements into four core challenges to lay the groundwork for an ecopreneurial ICT-driven e-waste management system:

(i) System is not good (Priority 1);
(ii) Less socialisation/information (Priority 2);
(iii) Low awareness of the community (Priority 3); and
(iv) Expensive processing costs (Priority 4).

The application of design thinking in ecopreneurship-focused e-waste disposal management is the main contribution of this research. The uniqueness lies in using design thinking primarily through the "empathise" and "define" phases to identify and classify stakeholder pain issues. This deliberate approach transforms the complex, "wicked problem" of e-waste into business potential. While earlier design thinking research often focused on designing and testing user interface prototypes for e-waste applications, this study evaluates the real challenges faced by diverse stakeholders before proposing technical solutions. Rather than developing solutions directly, this study provides a diagnostic basis—the affinity map and four primary problem categories—upon which effective solutions can be built.

It is recommended that ecopreneurs invest in large-scale infrastructure rather than superficial modifications to capture this unregulated waste stream. Startups should actively capitalise on e-waste's economic potential to counter the perception of "expensive processing costs." Recycling is financially motivated by recovering valuable materials, such as gold, silver, and copper, and highlighting the environmental value—managing one tonne of e-waste can reduce 1,400 tonnes of CO_2 emissions. E-waste recycling could generate a total of USD 1.8 billion for Indonesia by 2030, underscoring its financial potential. Furthermore, startups can develop technological, educational, and logistical tools to address issues of "less socialisation/information" and "low awareness of community." Mobile applications or smart recycling bins that reward appropriate disposal can bridge knowledge gaps and influence customer behaviour. Given the current system's lack of integrated information, government departments should improve e-waste data integration. Additionally, waste management institutions should reduce human error and cumbersome confirmation processes that delay the formal disposal system.

Acknowledgments

None.

Funding

Financial support to perform the research was received from the Telkom University under the Innovation Technology Product Research Scheme grant, Period I Fiscal Year 2023, number 266/PNLT3/PPM/2023.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualisation: Jurry Hatammimi
Formal analysis: All authors
Investigation: Jurry Hatammimi
Methodology: All authors
Writing–original draft: All authors
Writing–review & editing: Jurry Hatammimi

Ethics approval and consent to participate

The study was approved by the Institutional Review Board (or Ethics Committee) of The Social Humanities Research Ethics Committee, Universitas Islam Bandung (protocol code 008/B.004/KEPSOSHUM/VI/2025; date of approval: April 9, 2025). Verbal consent was obtained from each subject to participate in this study.

Consent for publication

Verbal consent was obtained from each of the subjects to publish their data and/or images by anonymising their identities.

Availability of data

Data derived from the interviews and transcripts are available upon reasonable request to the corresponding author.

References

Basel Convention. (2013). *Development of Technical Guidelines on e-waste*. Available from: http://www.basel.int/Implementation/Ewaste/TechnicalGuidelines/tabid/2377/Default.aspx [Last accessed on 2023 Jan 5].

Bratsberg, H. M. (2012). *Empathy maps of the foursight preferences (No. 176; Creative Studies Graduate Student Master's Projects)*. Buffalo State College. Available from: https://digitalcommons.buffalostate.edu/cgi/viewcontent.cgi?article=1180&context=creativeprojects [Last accessed 2023 June 24].

Brindhadevi, K., Barceló, D., Chi, N. T. L., & Rene, E. R. (2023). E-waste management, treatment options and the impact of heavy metal extraction from e-waste on human health: Scenario in Vietnam and other countries. *Environmental Research*, 217, 114926.

https://doi.org/10.1016/j.envres.2022.114926

Gonen, E. (2009). Change by Design: How design thinking transforms organisations and inspires innovation. Markets, Globalization & Development Review. *MUT Journal of Business Administration*, 4(2), 8.

https://doi.org/10.23860/MGDR-2019-04-02-08

Citarum Harum. (2021). *Bijak Dalam Mengelola Sampah Elektronik (Wise in Managing Electronic Waste)*. Available from: https://citarumharum.jabarprov.go.id/bijak-dalam-mengelola-sampah-elektronik/ [Last accessed 2023 April 18].

Citarum Harum. (2022). *Siapkan Titik Pembuangan Sampah Elektronik (Preparing E-Waste Drop Point)*. Available from: https://citarumharum.jabarprov.go.id/kota-bandung-siapkan-titik-pembuangan-sampah-elektronik/ [Last accessed 2023 April 18].

Defitri, M. (2022). *Mengenal Komunitas E-Waste RJ yang Kelola Sampah Elektronik (To know E-Waste RJ Community that manage e-waste)*. Waste4Change. Available from: https://waste4change.com/blog/mengenal-komunitas-e-waste-rj-yang-kelola-sampah-elektronik/ [Last accessed 2023 April 20].

Dinnata, R. (2017). *Pemerintah Dorong Industri Pengolahan Sampah Elektronik (Govt push e-waste processing factory)*. Bisnis.com. Available from: https://ekonomi.bisnis.com/read/20171002/257/695100/pemerintah-dorong-industri-pengolahan-sampah-elektronik [Last accessed 2023 April 21].

Fanthi, R., Fajarwati, A. A. S., & Chadijah, S. (2021). E-waste for interior accessories: An exploration of material recycling. *IOP Conference Series: Earth and Environmental Science*, 794(1), 012070.

https://doi.org/10.1088/1755-1315/794/1/012070

Forti, V., Balde, C. P., Kuehr, R., & Bel, G. (2020). The Global E-waste Monitor 2020: Quantities, flows and the circular economy potential. Washington DC: United Nations University.

George, S., & Michael, A. (2023). E-Waste Management System (EMS): An Online Platform to Sell Crafts Made From Tech Waste. In Rawani, A., Sahu, M., Chakarabarti, S., & Singh, A. (Eds.). *Sustainable Approaches and Strategies for E-Waste Management and Utilization*. Hershey PA: IGI Global. pp. 122–139.

https://doi.org/10.4018/978-1-6684-7573-7

Harikaran, M., Boopathi, S., Gokulakannan, S., & Poonguzhali, M. (2023). Study on the source of e-waste management and disposal methods. In Rawani, A., Sahu, M., Chakarabarti, S., & Singh, A. (Eds.). *Sustainable Approaches and Strategies for E-Waste Management and Utilization*. Hershey PA: IGI Global. pp. 39–60.

https://doi.org/10.4018/978-1-6684-7573-7.ch003

Inegbedion, H. (2022). Entrepreneurial design thinking and business success: Empirical evidence from Nigeria. *Problems and Perspectives in Management,* 20(2), 186–198.

https://doi.org/10.21511/ppm.20(2).2022.16

Iqbal, M., & Dyah, T. (2022). *Menggugah Rasa Resah Pada Masalah Sampah (Encourage concern about waste). Validnews*. Available from: https://validnews.id/kultura/menggugah-rasa-resah-pada-masalah-sampah [Last accessed 2023 May 3].

Jafar, R. (2015). E-Waste–Sampah Elektronik. Jakarta: Ewaste RJ.

Jain, M., Kumar, D., Chaudhary, J., Kumar, S., Sharma, S., & Singh Verma, A. (2023). Review on E-waste management and its impact on the environment and society. *Waste Management Bulletin,* 1(3), 34–44.

https://doi.org/10.1016/j.wmb.2023.06.004

Jimenez, I.A.C., Mauro, S., Napoli, D., *et al.* (2021). Design thinking as a framework for the design of a sustainable waste sterilisation system: the case of Piedmont region, Italy.

Electronics, 10(21), 2665.

https://doi.org/10.3390/electronics10212665

Khalid, H. (2022). *Timbunan Sampah Elektronik dan Pengelolaannya di Indonesia (Electronic Waste Piles and Their Management in Indonesia)*. Environment-Indonesia. Available from: https://environment-indonesia.com/timbunan-sampah-elektronik-dan-pengelolaanya-di-indonesia/ [Last accessed 2024 Jan 12].

Lu, J., Desta, B., Wu, K. D., Nith, R., Passananti, J. E., & Lopes, P. (2023). ecoeda: Recycling e-waste during electronics design. In: Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, October 29-November 1, 2023. San Francisco CA USA.

https://doi.org/10.1145/3586183.3606745

Machado, F., & Grilo, A. (2020). How can design thinking and lean startup improve waste collection systems. In: Proceedings of the International Conference on Industrial Engineering and Operations Management, June 22-24, 2026. Portugal. pp. 655-666.

Madrigal, A. N., Iyer-Raniga, U., & Yang, R. (2024). Exploring design thinking processes in circular economy strategies for PV waste management in Australia. In: Proceedings of the IOP Conference Series: Earth and Environmental Science, June 12-14, 2024; Online. 1363(1), 12051.

https://doi.org/10.1088/1755-1315/1363/1/012051

Massari, S., Principato, L., Antonelli, M., & Pratesi, C. A. (2022). Learning from and designing after pandemics. CEASE: A design thinking approach to maintaining food consumer behaviour and achieving zero waste. *Socio-Economic Planning Sciences*, 82, 101143.

https://doi.org/10.1016/j.seps.2021.101143

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oak: Sage.

Nnorom, I. C., & Osibanjo, O. (2008). Overview of electronic waste (e-waste) management practices and legislations, and their poor applications in the developing countries. *Resources, Conservation and Recycling,* 52(6), 843–858.

https://doi.org/10.1016/j.resconrec.2008.01.004

Nowakowski, P., & Pamuła, T. (2020). Application of deep learning object classifier to improve e-waste collection planning. *Waste Management*, 109, 1–9.

https://doi.org/10.1016/j.wasman.2020.04.041

Nuryanto A.D, & Suzianti, A. (2022). *Pengembangan Desain Antarmuka Pengumpulan Sampah Elektronik Berbasis Kebutuhan Dengan Pendekatan Design Thinking (Development of Need-Based Electronic Waste Collection User Interface Design With Design Thinking Approach)*. Bachelor Thesis. Engineering Faculty of Universitas Indonesia. Available from: https://lib.ui.ac.id/detail?id=20518225&lokasi=lokal [Last accessed on]

Parajuly, K., Kuehr, R., Awasthi, A. K., *et al.* (2019). *Future e-waste scenarios*. StEP (Bonn), UNU ViE-SCYCLE (Bonn) & UNEP IETC (Osaka).

Perkins, D. N., Drisse, M.-N. B., Nxele, T., & Sly, P. D. (2014). E-waste: A global hazard. *Annals of Global Health*, 80(4), 286–295.

https://doi.org/10.1016/j.aogh.2014.10.001

Plattner, H. (2010). *An Introduction to Design Thinking Process Guide*. Standford: The Institute of Design Standford.

Peraturan Pemerintah No. 27 Tahun 2020 tentang pengelolaan sampah spesifik (Government Regulation of the Republic of Indonesia Number 27 of 2020 concerning the Management of Specific Waste). (2020) (testimony of President of Republic of Indonesia). Available from: https://peraturan.bpk.go.id/Details/138876/pp-no-27-tahun-2020 [Last accessed on]

Puspa, A. (2023). *KLHK Prediksi Timbulan Sampah Elektronik Capai 12 Ribu Ton Per Hari pada 2030 (The Ministry of Environment and Forestry Predicts Electronic Waste Generation Will Reach 12,000 Tons Per Day by 2030)*. Media Indonesia. Available from: https://mediaindonesia.com/humaniora/558430/klhk-prediksi-timbulan-sampah-elektronik-capai-12-ribu-ton-per-hari-pada-2030 [Last accessed 2024 Jan 15].

Rahman, F. (2022). *Sampah Elektronik dan Hak Untuk Memperbaiki (Electronic Waste and Right to Repair)*. Pusat Studi Lingkungan Hidup Universitas Gadjah Muda. Available from: https://pslh.ugm.ac.id/sampah-elektronik-dan-hak-untuk-memperbaiki-right-to-repair/ [Last accessed 2024 Jan 15].

Rive, P. (2017). E Wastrels and Eco-Disasters: Speculative Design, Innovation and Global e-Waste. In Loué, C., Slimane, S. B. (Eds), In: Proceedings of the 12th European Conference on Innovation and Entrepreneurship, September 21-22, 2017; Paris, France. pp. 548–555.

Rois, M., Mubarak, A., & Suzianti, A. (2020). Designing Solution for Organic Waste Management System with Design Thinking Approach (Case Study in Depok). In: Proceedings of the IOP Conference Series: Earth and Environmental Science, October 21-22, 2025; online.

Sekaran, U., & Bougie, R. (2016). *Research Methods for Business A Skill Building Approach*, 7th ed. Chichester: John Wiley & Sons Ltd.

Shreyas Madhav, A. V, Rajaraman, R., Harini, S., & Kiliroor, C. C. (2022). Application of artificial intelligence to enhance collection of E-waste: A potential solution for household WEEE collection and segregation in India. *Waste Management & Research*, 40(7), 1047–1053.

https://doi.org/10.1177/0734242X211052846

Sundar, D., Mathiyazhagan, K., Agarwal, V., Janardhanan, M., & Appolloni, A. (2023). From linear to a circular economy in the e-waste management sector: Experience from the transition barriers in the United Kingdom. *Business Strategy* *and the Environment*, 32(7), 4282–4298.

https://doi.org/10.1002/bse.3365

Wanderley, A. R. M. C., & Bonacin, R. (2019). Designing mobile and iot solutions for sustainable smart cities: Studies with electronic waste disposal. In: Streitz, N., Konomi, S. (Eds), *Distributed, Ambient and Pervasive Interactions.* Cham: Springer International Publishing. pp. 212-226.

Wang, C., & Anand, S. (2024). Creative Design-Driven Interdisciplinary Studio Course for Electronic Waste Management. *The International Journal of Interdisciplinary Educational Studies*, 19(2), 69-80.

https://doi.org/10.18848/2327-011X/CGP/v19i02/69-80

Widodo, R. (2023). *Biasnya Anggapan Pada Limbah Elektronik (biased assumptions on electronic waste).* Validnews. Available from: https://validnews.id/kultura/biasnya-anggapan-pada-limbah-elektronik [Last accessed 2024 Jan 12].

Yin, R. K. (2014). *Case Study Research: Design and Methods*, 5th ed. Thousand Oak: Sage Publication.

Yu, Z., Gao, C., Yang, C., & Zhang, L. (2023). Insight into quantities, flows, and recycling technology of E-waste in China for resource sustainable society. *Journal of Cleaner Production*, 393, 136222.

https://doi.org/10.1016/j.jclepro.2023.136222

ACCSCIENCE PUBLISHING

International Journal of Systematic Innovation

ARTICLE

Blockchain vs. generative artificial intelligence in India: A comparative study of adoption drivers, barriers, and diffusion trajectories

Siddhartha Nigam*, **O. P. Wali**

Indian Institute of Foreign Trade, New Delhi, Delhi, India

Abstract

Blockchain and generative artificial intelligence (GenAI) are two contemporary emerging technologies that have exhibited different adoption trajectories since their inception. Blockchain technology traces its origins to 2008, when it was first conceptualized, whereas GenAI is a more recent development that entered the mainstream with the introduction of ChatGPT by OpenAI. India, as a developing economy, has consistently been at the forefront of technological innovations; however, the adoption patterns for these innovations have been notably different. Using secondary data retrieved from peer-reviewed research and systematic reviews, along with industry and market intelligence reports, this research revealed that blockchain, as a technology, adopts a bottom-up approach driven by financial inclusion imperatives and is inherently decentralized by design. GenAI, on the other hand, adopts a top-down approach, fueled by enterprise-driven adoption and rapid scaling across various sectors. Our findings suggest that the difference in their diffusion approaches is attributed to the persistent regulatory uncertainty and infrastructure constraints faced by blockchain, whereas GenAI has benefited from clearer policy support and lower entry barriers. This paper provides a framework-based, side-by-side comparison of two high-impact technologies in a single national context, linking micro-level adoption mechanisms to macro-level diffusion outcomes. These nuances could have significant implications for policymaking and recalibrating India's position in the global landscape.

Keywords: Blockchain adoption; Digital transformation; Emerging technologies; Generative artificial intelligence; India; Regulatory frameworks; Technology diffusion

***Corresponding author:**
Siddhartha Nigam
(nigam.siddhartha@gmail.com)

Citation: Nigam, S. & Wali, O. P. (2026). Blockchain vs. generative artificial intelligence in India: A comparative study of adoption drivers, barriers, and diffusion trajectories. *Int J Systematic Innovation*. 10(1):54-63. https://doi.org/10.6977/IJoSI.202602_10(1).0005

Received: November 23, 2025

Revised: December 28, 2025

Accepted: January 9, 2026

Published online: February 13, 2026

1. Introduction

Blockchain, a distributed ledger system that enables trustless peer-to-peer transactions (Crosby, 2016; Pilkington, 2016), has witnessed significant grassroots adoption in India despite regulatory uncertainty. In the cryptocurrency domain itself, India ranked first globally in the 2024 Chainalysis Global Crypto Adoption Index (Chainalysis Team, 2024), with over 35 million trading accounts and 12% of global blockchain developers concentrated domestically. Overall, the Indian blockchain market is expected to grow from US$657 million (2024) to US$61.5 billion (2033) with a compound annual growth rate (CAGR) of 66% (Imarc group, 2024a). On the other hand, generative artificial intelligence (GenAI)—which utilizes advanced neural frameworks to produce human-like content across different forms of communication (Cao et al., 2023; Goodfellow et al., 2020)—has also witnessed an exponential enterprise adoption, with the Indian GenAI market expected to grow from US$1.3 billion (2024) to US$5.4 billion (2033) with a CAGR of 15% (Imarc group, 2024c).

India, as the world's fastest-growing major economy with a digitally native demographic, presents a unique case study for examining emerging technology adoption patterns. While both technologies promise disruptive innovation, their adoption trajectories diverge significantly in speed and directionality. This raises some fundamental research questions:

(i) What are the factors underlying the differential adoption rates?
(ii) How do regulatory frameworks and infrastructure requirements shape technology diffusion patterns?
(iii) What implications do these patterns then hold for India's technological competitiveness?

Despite growing literature on blockchain adoption and the rapid emergence of GenAI studies, most of the studies examined these technologies in isolation and often used inconsistent dimensions and metrics, making cross-technology comparisons difficult—particularly in a national (in this case, Indian) context. Moreover, only a few studies have connected micro-level adoption mechanisms (e.g., perceived usefulness, trialability, legitimacy) to macro-level diffusion patterns, explaining why adoption trajectories diverge. Hence, this study aims to address this gap by applying a single, theory-grounded analytical framework to systematically compare the adoption of blockchain and GenAI in India.

This paper addresses these questions through a comparative analysis, contributing to the technology adoption literature by examining two concurrent yet divergent adoption patterns for emerging technologies within the same geographical context: India.

2. Theoretical framework and literature review

Numerous theories of technology adoption have been proposed, providing foundational insights into the diffusion mechanisms of emerging technologies. Rogers' (1987) diffusion of innovation theory identifies perceived attributes—relative advantage, compatibility, complexity, trialability, and observability—as determinants of adoption rates. The technology acceptance model emphasizes perceived usefulness and ease of use (Davis, 1989), whereas the institutional theory highlights regulatory environments, normative pressures, and mimetic isomorphism in organizational adoption (DiMaggio & Powell, 1983). Recent blockchain adoption studies have extended these frameworks, identifying trust mechanisms, interoperability, scalability, and regulatory certainty as critical factors (Clohessy et al., 2020; Taherdoost, 2022; Queiroz & Wamba, 2019).

2.1. Blockchain adoption

Blockchain adoption research reveals distinct patterns across various sectors, depending on the specific use case. Kshetri (2018) demonstrates the efficacy of blockchain in enhancing supply chain transparency, whereas Casino et al. (2019) have conducted a systematic analysis of blockchain applications across healthcare, finance, and governance. In emerging economies, blockchain adoption is driven by financial inclusion initiatives, lack of or limited institutional trust, and leapfrogging opportunities (Kshetri, 2018; Ghode et al., 2020). However, blockchain adoption faces persistent barriers, such as scalability challenges (Eyal et al., 2016; Zheng et al., 2018), energy consumption concerns (Beck et al., 2016), regulatory ambiguity (Kiviat, 2015; Wright & De Filippi, 2015), and technical complexity in implementation (Mthimkhulu & Jokonya, 2022).

Indian blockchain adoption is an interesting landscape. It has support at the grassroots level via cryptocurrency enthusiasm, but interestingly, that coexists with regulatory hostility, in the form of a 30% capital gains tax and central bank skepticism equating cryptocurrencies to "speculative gambling" (Chainalysis Team, 2024). Despite these headwinds, India's blockchain market reached US$657 million (2024) and is projected to achieve US$61.5 billion by 2033 (65.6% CAGR) (Imarc group, 2024a). Government initiatives include the Vishvasya National Blockchain Framework, operationalizing Blockchain-as-a-Service (BaaS) infrastructure across three data centers, and state-level implementations in Telangana (land records), Jharkhand (agricultural supply chains), and Goa (public registries).

2.2. Generative artificial intelligence adoption

Generative AI literature focuses on enhancing productivity (Brynjolfsson et al., 2023; Noy & Zhang, 2023), creative augmentation (Epstein et al., 2023; Haase & Hanel, 2023), and organizational transformations (Budhwar et al., 2023; Korzynski et al., 2023). Unlike blockchain's infrastructure-heavy requirements, GenAI exhibits lower entry barriers through application programming interface-accessible models (Fui-Hoon et al., 2023; Su & Yang, 2023), enabling rapid organizational adoption. However, concerns regarding algorithmic bias (McGee, 2023), hallucinations (Alkaissi & McFarlane, 2023), intellectual property rights (Hacker et al., 2023), and labor displacement (Ponce, 2023) have an impact on its adoption trajectory.

India's GenAI adoption demonstrates unique characteristics of its own. Some examples that illustrate this are the development of a multilingual model (Hanooman supports 98 languages, including 12 Indian languages), sovereign AI initiatives (such as myShakti chatbot), and sector-specific applications across information technology (IT), healthcare, banking, and e-commerce. The Indian economy accounts for 6.1% of global GenAI revenue (2024), with enterprise adoption rates (73%) exceeding the United States (45%) and the United Kingdom (29%). A major difference here, compared to the case of blockchain adoption, has been the role of the government. Instead of a default skepticism for blockchain, the Indian government has been a big supporter of GenAI through the IndiaAI Mission (US$1.25 billion allocation), National Association of Software and Service Companies (NASSCOM)'s AI Adoption Index framework, and state-level AI policies, which created a conducive institutional environment, notably absent in the context of blockchain.

2.3. Literature comparison

Across the two literatures, a consistent pattern emerges: blockchain adoption is frequently ecosystemdependent (requiring coordinated participation, governance, and standards), while GenAI adoption is often modular and individually trialable, enabling faster diffusion. Institutional legitimacy (DiMaggio & Powell, 1983) appears to operate differently: blockchain legitimacy is strongly influenced by regulatory interpretation and association with cryptocurrency markets (Kiviat, 2015), whereas GenAI legitimacy is increasingly shaped by responsible AI governance and sector-specific controls (Hacker et al., 2023).

3. Methodology

This comparative study employed mixed-methods analysis, integrating quantitative market data with qualitative policy analysis. Data sources include: (i) market reports from the Imarc group, Grand View Research, BlueWeave Consulting, and KPMG; (ii) adoption indices from Chainalysis, NASSCOM, and Statista; (iii) government policy documents, including Vishvasya framework, IndiaAI Mission guidelines, and Reserve Bank of India circulars; and (iv) academic literature encompassing blockchain and GenAI adoption research papers.

The analytical framework comprised four dimensions: (i) market trajectory analysis, which compares growth rates, market valuations, and sector penetration; (ii) regulatory environment assessments to evaluate policy clarity, institutional support, and compliance frameworks; (iii) adoption barrier identification categorized using the technical, economic, organizational, and social impediments; and (iv) use-case maturity to analyze deployment breadth and impact across sectors. Temporal scope spanned 2023–2033, capturing the current state and near-term projections. The geographic focus centered on India, with selective global comparisons provided for added context.

To operationalize this comparison, Figure 1 synthesizes insights from the technology acceptance model/diffusion of innovations theory, institutional theory, and a general-purpose technology (GPT) framing into four comparative dimensions, also marked as the core adoption factors (market trajectory, regulatory environment, infrastructure/entry barriers, and use-case maturity/risk) and the mediating adoption mechanisms indicated throughout the rest of the paper. Both of these, in combination, then drive the respective technology adoption trajectory in some scenarios, underpinning complementary synergies to adoption.

4. Comparative analysis: Adoption trajectories

4.1. Market growth dynamics

The blockchain and GenAI markets exhibit contrasting growth profiles, with nuances observed across their underlying sectors. As seen from Figure 2, India's blockchain market reached US$657 million (2024) with an expected projection to US$61.5 billion by 2033, representing 65.6% CAGR (Imarc group, 2024a; Imarc, 2024b). Blockchain fintech markets were valued at US$101 million (2024), growing to US$2 billion (2033) at a 39% CAGR. Conversely, GenAI demonstrated a US$1.3 billion (2024) baseline with projections of US$5.4 billion (2033), indicating an average 15% CAGR and varying further depending on sector inclusion.

This differential is characteristic of each technology's profile and respective maturation stages: blockchain remains early-stage with infrastructure buildout, pilot implementations, and regulatory navigation, while GenAI leverages existing cloud infrastructure and established machine learning pipelines, enabling rapid scaling. Blockchain adoption follows S-curve dynamics, characterized by prolonged early-adopter phases, whereas GenAI exhibits exponential growth, a characteristic of platform technologies with network effects (Arthur, 1989). However, in the long run, once the early-adopter phase has passed, it appears that blockchain will experience

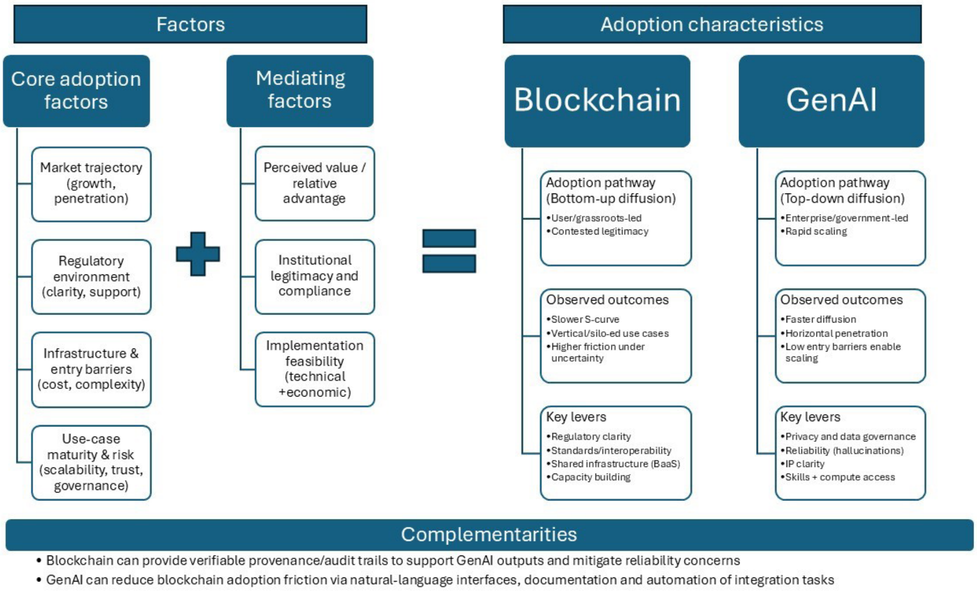


Figure 1. Conceptual framework summary for comparing adoption trajectories of blockchain vs. generative artificial intelligence (GenAI) in India

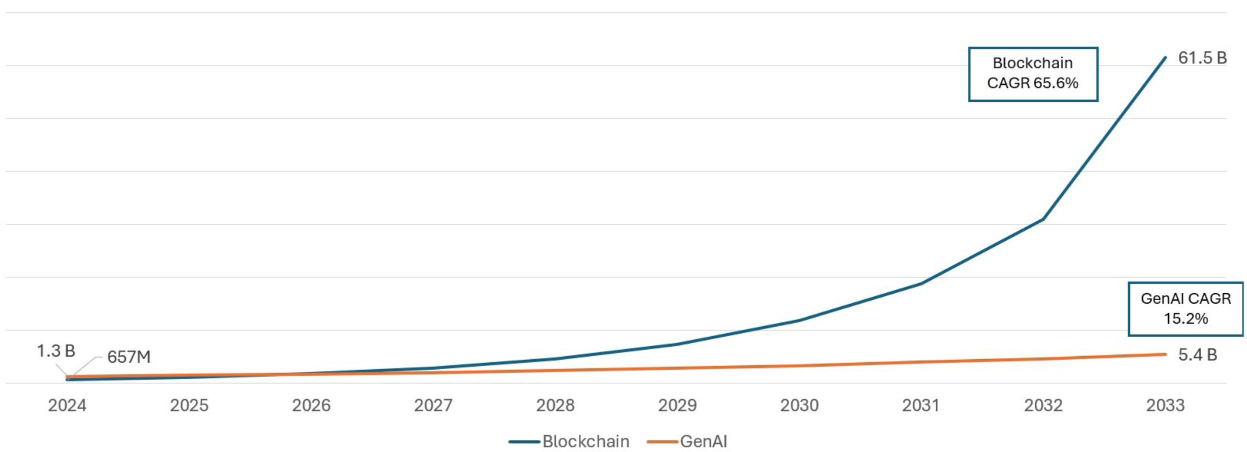


Figure 2. Blockchain vs. generative artificial intelligence (GenAI) growth projections (in United States Dollars). Data adapted from the Imarc group (2024a, 2024c).
Abbreviation: CAGR: Compound annual growth rate.

accelerated growth and overtake GenAI in adoption.

4.2. Sectoral penetration patterns

Blockchain adoption is concentrated in specific high-value use cases, including financial services (cross-border payments, DeFi protocols), supply chain management (pharmaceuticals tracking, agricultural provenance), state land registry systems (such as those in Telangana and Goa), and trade finance (as seen in the Reserve Bank of India Innovation Hub pilot). Adoption remains vertically silo-ed with limited cross-sector interoperability, reflecting Kshetri's (2018) observation that blockchain value propositions are use-case specific rather than universally applicable.

In contrast to this, GenAI demonstrates broader horizontal penetration: IT services (code generation, documentation), healthcare (diagnostic support, administrative automation), banking (customer service, fraud detection), e-commerce (personalization, recommendation engines), media (content generation), and education (adaptive learning platforms). This diffusion pattern closely aligns with the characteristics of GPT (Bresnahan & Trajtenberg, 1995), where foundational innovations enable diverse applications across sectors with minimal modification or customization.

4.3. Regulatory environment comparison

Regulatory certainty is also a critical differentiator. Blockchain faces sustained ambiguity and skepticism from stakeholders in India, as evidenced by ongoing discussions about cryptocurrency prohibition since 2018, the introduction of a 30% capital gains tax in 2022, the imposition of a 1% tax deducted at source on transactions, and actions by the Financial Intelligence Unit to shut down offshore exchanges. At the same time, some positive signals have emerged, particularly government-led initiatives in blockchain infrastructure development, such as Vishvasya. This regulatory dichotomy—embracing blockchain technology on one hand while suppressing cryptocurrency applications—creates adoption friction, particularly for private-sector initiatives, as they are unsure of creating future-proof applications. There is little confidence or certainty regarding the legality of such applications if they are developed.

Generative AI benefits from supportive policy architecture from the beginning, including IndiaAI Mission (2024) allocating US$1.25 billion for compute infrastructure, datasets, and application development, NASSCOM AI Adoption Index providing sector benchmarks, state governments (Karnataka, Telangana, Maharashtra) offering AI policy frameworks, and the absence of prohibitive regulations beyond current data protection and IT Act provisions. This permissive environment, combined with global AI governance uncertainty, positions India favorably for experimental deployment and rapid iteration in the field of GenAI, relative to blockchain.

4.4. Infrastructure requirements

Blockchain, as a technology, requires significant infrastructure investments, including distributed node networks, consensus mechanism implementations, cryptographic key management systems, interoperability protocols, and substantial energy resources (depending on the algorithms used, such as proof-of-work chains). Furthermore, public blockchain participation requires continuous uptime, bandwidth provisioning, and more stringent security controls. These requirements create high entry barriers, particularly for small to medium-sized enterprises and government agencies with legacy systems (Ghode et al., 2020; Mthimkhulu & Jokonya, 2022).

Generative AI instead leverages existing cloud infrastructure through application programming interface-accessible models (OpenAI, Google Gemini, Anthropic Claude), reducing upfront capital expenditure requirements. Organizations can implement GenAI applications with minimal infrastructure investment, relying on pay-per-use consumption models. India-specific initiatives (Yotta's Shakti Cloud, Airtel–Google Cloud partnership) further democratize access to GenAI through localized compute resources and multilingual model availability. These infrastructure advantages accelerate adoption not only among individual users but also across enterprises seeking to deploy these solutions at scale.

4.5. Adoption drivers and barriers

Blockchain drivers include: (i) financial inclusion addressing 190 million unbanked individuals, (ii) remittance efficiency for US$100+ billion annual inflows, (iii) supply chain transparency combating counterfeiting in pharmaceuticals and agriculture, (iv) land registry digitization reducing fraud in property transactions, (v) DeFi protocols enabling permissionless financial services, and (vi) developer ecosystem growth (12% global share) supporting talent density.

However, several can become barriers to blockchain adoption, such as (i) regulatory uncertainty deterring institutional investment, (ii) scalability limitations (Bitcoin: 7 transactions/second vs. Visa: 24,000+), (iii) energy consumption concerns contradicting sustainability commitments, (iv) user experience complexity requiring technical literacy, (v) interoperability challenges across blockchain platforms, and (vi) limited public awareness beyond cryptocurrency speculation.

Generative AI drivers include: (i) productivity enhancement delivering a 3.7× return per dollar invested, (ii) content creation democratization through natural language interfaces, (iii) multilingual capability serving diverse linguistic demographics, (iv) cloud-native deployment reducing implementation timelines, (v) government support through IndiaAI Mission and sector policies, and (vi) global platform availability (ChatGPT, Gemini) providing immediate access.

Generative AI adoption also faces barriers, such as (i) data privacy concerns regarding sensitive information processing, (ii) algorithmic hallucinations compromising reliability in critical applications, (iii) talent shortages despite a large global AI workforce (approximately 50% based in the United States), (iv) energy-intensive training of large language models, (v) intellectual property ambiguities surrounding generated content, and (vi) quality control challenges requiring human oversight.

Overall, blockchain exhibits higher friction to adoption, driven by regulatory ambiguity, integration complexity, and a slower time-to-value. GenAI, on the other hand, has a generally faster time-to-value ratio, easier trialability, and a more modular approach to adoption; however, it is deterred by non-trivial governance or privacy concerns.

Figure 3 visualizes the adoption friction profile for both these technologies alongside these dimensions.

5. Discussion

5.1. Bottom-up vs. top-down adoption

Blockchain adoption in India exhibits bottom-up characteristics driven by individual users seeking financial autonomy, peer-to-peer transaction efficiency, and speculative investment opportunities. Grassroots adoption (Chainalysis Team, 2024) precedes institutional legitimation, creating tension with top-down regulatory frameworks. This inversion of typical technology diffusion—where institutional adoption precedes consumer uptake—reflects blockchain's disintermediation value proposition, which inherently challenges existing institutional power structures (Wright & De Filippi, 2015).

GenAI follows a traditional top-down diffusion pattern: large enterprises deploy GenAI for operational efficiency, followed by the adoption of GenAI by small and medium-sized enterprises through cloud platforms, and eventually, consumer applications. Enterprise adoption rates (73% in India) drive market expansion, with institutional actors shaping the evolution of use cases and best practices. This pattern enables regulatory co-evolution, where policymakers observe enterprise implementations before establishing governance frameworks, reducing adoption friction.

This aligns with diffusion theory (Davis, 1989), which posits that trialability and low perceived complexity accelerate early adoption, whereas dependencies, such as multi-actor coordination and those related to the ecosystem, slow down diffusion even when the potential value from adoption is high. This also entails legitimacy foundations, where technologies that are associated with higher uncertainty in regulatory matters often require stronger support to scale up.

Figure 3. Adoption friction profile (higher score = more favorable)
Abbreviations: GenAI: Generative artificial intelligence; Infra: Infrastructure.

5.2. Technology maturity and use-case specificity

Blockchain's adoption trajectory reflects an early-stage technology with limited proven use cases beyond cryptocurrency. Despite its theoretical applicability across trust-intensive domains, practical implementations face challenges, including smart contract vulnerabilities (Casino et al., 2019), oracle problems connecting blockchain to real-world data (Zheng et al., 2018), and unclear value propositions compared to centralized databases for many applications. This maturity gap extends adoption timelines as organizations await definitive evidence of transformative value.

Generative AI demonstrates rapid maturity acceleration through foundation models, enabling diverse applications without fundamental architectural redesign. Single models (GPT-4, Gemini) support text generation, code writing, data analysis, image creation, and conversational interfaces, reducing deployment complexity. This generality contrasts with blockchain's application-specific implementations, which facilitate faster adoption cycles and iterative improvements based on user feedback.

These results are consistent with prior blockchain research, which emphasizes the "pilot-to-production" gap driven by scalability, interoperability, and governance constraints. In contrast, GenAI studies focus on modular development through standardized tools and platforms.

5.3. Global positioning and strategic implications

India's dual technology positioning presents strategic opportunities and risks. Blockchain leadership through grassroots adoption and developer concentration positions India as a potential hub for Web3 innovation, yet regulatory hostility toward cryptocurrencies undermines this advantage. Countries embracing clearer regulatory frameworks (United Arab Emirates, Singapore, Switzerland) may capture institutional blockchain development despite lower grassroots adoption.

The adoption of GenAI positions India among global leaders, thanks to government support and large talent pools that create competitive advantages. However, dependence on Western foundation models (OpenAI, Google, Anthropic) raises concerns about sovereignty. Indigenous model development (Hanooman, myShakti) addresses this vulnerability, but it requires sustained investment in computing infrastructure and research capabilities. The IndiaAI Mission's US$1.25 billion allocation signals governmental recognition of its strategic importance, yet it pales in comparison to the United States' US$109 billion in private investment (as of 2024) and China's AI funding.

From a strategic perspective, the findings support the diffusion of GenAI, given that it is easy to embed and diffuse horizontally across sectors, whereas blockchain tends to thrive in specific verticals where coordination can be enforced.

5.4. Socioeconomic implications

Blockchain adoption is concentrated among educated, urban, and technologically literate populations, with limited penetration in rural areas that lack digital infrastructure and financial literacy. This pattern risks exacerbating digital divides unless accompanied by targeted education and investment in infrastructure. DeFi protocols, in particular, exclude populations that are unable to navigate complex interfaces or lack cryptocurrency on-ramps.

Generative AI demonstrates broader accessibility through natural language interfaces and support for vernacular languages, making it more accessible to a wider audience. Multilingual models (such as Hanooman's 98 languages) enable rural populations to interact with AI systems in native languages, potentially democratizing access to information, education, and services. However, algorithmic biases encoded in training data may perpetuate existing inequalities, requiring careful governance to ensure equitable outcomes.

6. Conclusion

This comparative analysis reveals fundamentally divergent adoption trajectories for blockchain and GenAI technologies in India. Blockchain follows a contested pathway characterized by grassroots enthusiasm, regulatory ambiguity, infrastructure challenges, and concentrated use-case deployment, resulting in measured growth despite its leadership in global adoption. GenAI demonstrates accelerated diffusion enabled by supportive policy frameworks, low entry barriers, broad sectoral applicability, and established cloud infrastructure, positioning it for rapid scaling across the Indian economy.

These differential trajectories reflect underlying technological characteristics: blockchain's trust infrastructure proposition requires fundamental reimagining of transactional architectures, while GenAI augments existing workflows through intelligent automation. Blockchain disrupts intermediaries, generating regulatory resistance; GenAI enhances organizational capabilities, attracting institutional support. Blockchain demands specialized technical expertise and infrastructure investment; GenAI provides accessible interfaces with pay-per-use models.

For policymakers, these findings suggest the need for nuanced technology governance, recognizing distinct adoption dynamics. Blockchain requires regulatory clarity, striking a balance between encouraging innovation and protecting consumers, moving beyond binary prohibition–acceptance frameworks toward risk-based regulation, and acknowledging the diverse applications of blockchain beyond cryptocurrency speculation. GenAI necessitates proactive governance that addresses algorithmic accountability, data privacy, intellectual property, and labor market disruption, while maintaining an innovation-conducive permissiveness.

For organizational strategists, the analysis highlights strategic imperatives. Blockchain adoption requires patience, pilot-scale experimentation, and contingency planning for regulatory evolution; GenAI demands rapid experimentation, ethical frameworks, and talent development to capture first-mover advantages in a fast-evolving landscape. Organizations pursuing both technologies should recognize the complementarities—blockchain provides verifiable data provenance that addresses GenAI hallucination concerns, GenAI enables natural language blockchain interfaces, and reduces adoption barriers.

In terms of limitations, this study relied on secondary sources and qualitative synthesis; hence, it did not estimate causal effects or statistically test the results. The analysis focused on India, which may limit generalizability to other emerging economies with different regulatory regimes and digital infrastructure. Finally, both technologies are evolving rapidly; therefore, some findings—particularly those related to regulation, model capability, and enterprise adoption practices—may shift as new policies and products emerge.

Future work could validate and refine the framework using primary data (e.g., surveys/interviews), sector-level adoption datasets, or longitudinal case studies. Some potential examples can include: (i) longitudinal adoption studies tracking blockchain and GenAI diffusion across organizational types and geographic regions, (ii) hybrid architectures combining blockchain verification with GenAI capabilities, (iii) regulatory effectiveness analysis comparing jurisdictions with varying governance approaches, (iv) socioeconomic impact assessments measuring technology effects on employment, inequality, and access to services, and (v) comparative studies examining adoption trajectories in other emerging economies.

India's position at the intersection of grassroots blockchain adoption and enterprise GenAI leadership presents unique opportunities to shape global technology governance, develop indigenous capabilities, and leverage digital transformation for inclusive economic growth. Realizing this potential requires coordinated action across government, industry, and civil society, balancing innovation imperatives with equity concerns and sovereignty objectives in an increasingly technology-mediated world.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: All authors
Formal analysis: All authors
Investigation: All authors
Methodology: All authors
Writing–original draft: All authors
Writing–review & editing: All authors

Availability of data

Data will be made available upon request to the corresponding author.

References

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 15(2), e35179.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, 99(394), 116-131.

Beck, R., Stenum Czepluch, J., Lollike, N., & Malone, S. (2016). Blockchain—The gateway to trust-free cryptographic transactions. In: Proceedings of the Twenty-Fourth European Conference on Information Systems (ECIS 2016) (Research Paper 153). June 12-15, 2016; Istanbul, Turkey. Association for Information Systems. Available from: https://aisel.aisnet.org/ecis2016_rp/153 [Last accessed on 2025 Nov 1].

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies: Engines of growth? *Journal of Econometrics*, 65(1), 83-108.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work (NBER Working Paper No. 31161)*. Cambridge, MA, USA: National Bureau of Economic Research.

https://doi.org/10.3386/w31161

Budhwar, P., Chowdhury, S., Wood, G., et al. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal,* 33(3), 606–659.

https://doi.org/10.1111/1748-8583.12524

Cao, Y., Li, S., Liu, Y., et al. (2024). A survey of AI-generated content (AIGC). *ACM Computing Surveys,* 57(5), 1–38.

https://doi.org/10.1145/3704262

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, 36, 55-81.

Chainalysis Team. (2024). *The 2024 global adoption index: Central & Southern Asia and Oceania (CSAO) region leads the world in terms of global cryptocurrency adoption*. Chainalysis. Available from: https://www.chainalysis.com/blog/2024-global-crypto-adoption-index/ [Last accessed on 2025 Nov 1].

Clohessy, T., Acton, T., & Rogers, N. (2020). Antecedents of blockchain adoption: An integrative framework. *Strategic Change*, 29(5), 501-515.

Crosby, M., Pattanayak, P., Verma, S., & Kalyanaraman, V. (2016). Blockchain technology: Beyond bitcoin. *Applied Innovation*, 2(6-10), 71.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review,* 48(2), 147–160.

https://doi.org/10.2307/2095101

Epstein, Z., Hertzmann, A., Herman, L., et al. (2023). Art and the science of generative AI: A deeper dive. *Science,* 380(6650), 1110–1111.

https://doi.org/10.1126/science.adh4451

Eyal, I., Gencer, A. E., Sirer, E. G., & van Renesse, R. (2016). Bitcoin-NG: A scalable blockchain protocol. In: Proceedings of the 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI '16). March 18, 2016; Santa Clara, CA, USA. USENIX Association. pp. 45–59. Available from: https://www.usenix.org/conference/nsdi16/technical-sessions/presentation/eyal [Last accessed on 2025 Nov 1].

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277-304.

Ghode D, Yadav V, Jain R, Soni G. (2020). Adoption of blockchain in supply chain: An analysis of influencing factors. *Journal of Enterprise Information Management*, 33(3), 437–456.

https://doi.org/10.1108/JEIM-07-2019-0186

Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2020). Generative adversarial networks. *Communications of the ACM,* 63(11), 139-144.

https://doi.org/10.1145/3422622

Haase, J., & Hanel, P. H. (2023). Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity. *Journal of Creativity*, 33(3), 100066.

Hacker, P., Engel, A., & Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. In: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23). June 12-15, 2023; Chicago, IL, USA. Association for Computing Machinery. pp. 1112–1123.

https://doi.org/10.1145/3593013.3594067

Imarc Group. (2024a). *India blockchain market size, share, trends and forecast by component, provider, type, deployment mode, organization size, vertical, and region, 2025–2033.* Available from https://www.imarcgroup.com/india-blockchain-market [Last accessed on 2025 Dec 28].

Imarc Group. (2024b). *India fintech blockchain market size, share, trends and forecast by industry, application, end user, and region, 2025–2033.* Available from https://www.imarcgroup.com/india-fintech-blockchain-market [Last accessed on 2025 Dec 28].

Imarc Group. (2024c). *India generative AI market size, share, trends and forecast by component, technology, application, model, customers, end use, and region, 2025–2033.* Available from https://www.imarcgroup.com/india-generative-ai-market [Last accessed on 2025 Dec 28].

Kiviat, T. I. (2015). Beyond bitcoin: Issues in regulating blockchain transactions. *Duke Law Journal*, 65, 569-608.

Korzynski, P., Mazurek, G., Altmann, A., et al. (2023). Generative artificial intelligence as a new context for management theories: Analysis of ChatGPT. *Central European Management Journal*, 31(1), 3-13.

Kshetri, N. (2018). Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management*, 39, 80-89.

McGee, R. W. (2023). *Is ChatGPT biased against conservatives? An empirical study* (SSRN Scholarly Paper No. 4359405). Amsterdam: SSRN.

https://doi.org/10.2139/ssrn.4359405

Mthimkhulu, A., & Jokonya, O. (2022). Exploring the factors affecting the adoption of blockchain technology in the supply chain and logistic industry. *Journal of Transport and Supply Chain Management,* 16, 750.

https://doi.org/10.4102/jtscm.v16i0.750

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187-192.

Pilkington, M. (2016). Blockchain technology: Principles and applications. In Olleros, F. X. & Zhegu, M. (Eds.) *Research Handbook on Digital Transformations*. Cheltenham: Edward Elgar Publishing. pp. 225-253.

Ponce Del Castillo, A. (2024). Generative AI, generating precariousness for workers? *AI & Society,* 39(5), 2601–2602.

https://doi.org/10.1007/s00146-023-01719-9

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.

Rogers, E. M. (1987). Diffusion of innovations: An overview. In Anderson, J. G. & Jay, S. J. (Eds.) *Use and impact of computers in clinical medicine*. Luxembourg: Springer. pp. 113–131.

https://doi.org/10.1007/978-1-4613-8674-2_9

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355-366.

Taherdoost, H. (2022). A critical review of blockchain acceptance models—Blockchain technology adoption frameworks and applications. *Computers*, 11(2), 24.

Wright, A., & De Filippi, P. (2015). *Decentralized blockchain technology and the rise of lex cryptographia*. Amsterdam: SSRN.

https://doi.org/10.2139/ssrn.2580664

Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, 14(4), 352-375.

INSTRUCTIONS TO AUTHORS

Submission of papers

The International Journal of Systematic Innovation is a refereed journal publishing original papers four times a year in all areas of SI. Papers for publication should be submitted online to the IJoSI website (http://www.ijosi.org) In order to preserve the anonymity of authorship, authors shall prepare two files (in MS Word format or PDF) for each submission. The first file is the electronic copy of the paper without author's (authors') name(s) and affiliation(s). The second file contains the author's (authors') name(s), affiliation(s), and email address(es) on a single page. Since the Journal is blind refereed, authors should not include any reference to themselves, their affiliations or their sponsorships in the body of the paper or on Figs and computer outputs. Credits and acknowledgement can be given in the final accepted version of the paper.

Editorial policy

Submission of a paper implies that it has neither been published previously nor submitted for publication elsewhere. After the paper has been accepted, the corresponding author will be responsible for page formatting, page proof and signing off for printing on behalf of other co-authors. The corresponding author will receive one hardcopy issue in which the paper is published free of charge.

Manuscript preparation

The following points should be observed when preparing a manuscript besides being consistent in style, spelling, and the use of abbreviations. Authors are encouraged to download manuscript template from the IJoSI website, http://www.ijosi.org.

1. *Language.* Paper should be written in English except in some special issues where Chinese may be accepTable Each paper should contain an abstract not exceeding 200 words. In addition, three to five keywords should be provided.
2. *Manuscripts.* Paper should be typed, single-column, double-spaced, on standard white paper margins: top = 25mm, bottom = 30mm, side = 20mm. (The format of the final paper prints will have the similar format except that double-column and single space will be used.)
3. *Title and Author.* The title should be concise, informative, and it should appear on top of the first page of the paper in capital letters. Author information should not appear on the title page; it should be provided on a separate information sheet that contains the title, the author's (authors') name(s), affiliation(s), e-mail address(es).
4. *Headings.* Section headings as well as headings for subsections should start front the left-hand margin.
5. *Mathematical Expressions.* All mathematical expressions should be typed using Equation Editor of MS Word. Numbers in parenthesis shall be provided for equations or other mathematical expressions that are referred to in the paper and be aligned to the right margin of the page.
6. *Tables and Figs.* Once a paper is accepted, the corresponding author should promptly supply original copies of all drawings and/or tables. They must be clear for printing. All should come with proper numbering, titles, and descriptive captions. Fig (or table) numbering and its subsequent caption must be below the Fig (or table) itself and as typed as the text.
7. *References.* Display only those references cited in the text. References should be listed and sequenced alphabetically by the surname of the first author at the end of the paper. For example:

Altshuller, G. (1998). *40 Principles: TRIZ Keys to Technical Innovation*, Technical Innovation Center.
Sheu, D. & Lee, H. (2011). A Proposed Process for Systematic Innovation, International Journal of Production Research, Vol. 49, No. 3, 2011, 847-868.

The International Journal of Systematic Innovation
Journal Order Form

Organization Or Individual Name	
Postal address for delivery	
Person to contact	Name: e-mail: Position: School/Company:
Order Information	**I would like to order ___ copy(ies) of the** *International Journal of Systematic Innovation***:** **Period Start: 1st/ 2nd half ____, Year:___(Starting 2010)** **Period End : 1st/ 2nd half ____, Year:** **Price:** **Institutions: US $150 (yearly) / NT 4,500 (In Taiwan only)** **Individuals: US $50 (yearly) / NT 1500 (In Taiwan only)** (Local postage included. International postage extra) **E-mail to**: IJoSI@systematic-innovation.org or fax: +886-3-572-3210 Air mail desired □ (If checked, we will quote the additional cost for your consent)
Total amount due	**US$**

Payment Methods:

1. **Credit Card (Fill up the following information and e-mail/ facsimile this form to The Journal office indicated below)**
2. **Bank transfer**
3. **Account:** The Society of Systematic Innovation
4. **Bank Name:** Mega International Commercial BANK
5. **Account No:** 020-53-144-930
6. **SWIFT Code:** ICBCTWTP020
7. **Bank code**：017-0206
8. **Bank Address:** No. 1, Xin'an Rd., East Dist., Hsinchu City 300, Taiwan (R.O.C.)

VISA / Master/ JCB/ AMERICAN Cardholder Authorization for Journal Order

Card Holder Information

Card Holder Name	(as it appears on card)		
Full Name (Last, First Middle)			
Expiration Date	/ (month / year)	Card Type	□ VISA □ MASTER □ JCB
Card Number	□□□□-□□□□-□□□□-□□□□	Security Code	□□□ 
Amount Authorized		Special Messages	
Full Address (Incl. Street, City, State, Country and Postal code)			

Please Sign your name here _________________ (same as the signature on your card)

The Society of Systematic Innovation

6 F, #352, Sec. 2, Guanfu Rd,
Hsinchu, Taiwan, 30071, R.O.C.