Intelligent Systems and Robotics

E-ISSN: xxxx-xxxx
P-ISSN: xxxx-xxxx

Research Article

A Hybrid Neural Symbolic Approach for Human Robot Interaction Enhancement Using Multimodal Sensor Fusion and Context Aware Behavioral Adaptation Techniques

Setyawan Wibisono [1], Hayadi Hamuda [2], and Encik Yoega Renaldi [3]

1 Universitas Stikubank setyawan@edu.unisbank.ac.id
2 Universitas Pamulang dosen02886@unpam.ac.id
3 Universitas Pamulang dosen03347@unpam.ac.id
* Corresponding Author : Setyawan Wibisono

Abstract: Human–Robot Interaction (HRI) systems increasingly rely on data-driven approaches to interpret multimodal sensory inputs and support natural interaction. However, purely neural-based HRI models often suffer from limited interpretability and insufficient context-aware decision making, which can reduce user trust and adaptability in dynamic interaction scenarios. To address these limitations, this study proposes a hybrid neural–symbolic HRI framework that integrates multimodal neural perception with explicit symbolic reasoning for adaptive and interpretable robot behavior. The proposed system combines deep neural networks for processing visual, speech, and gesture inputs with a rule-based symbolic reasoning layer that models interaction context, user states, and behavioral constraints. A loosely coupled integration strategy enables neural outputs to be transformed into symbolic representations, allowing logical inference to guide action selection while preserving perceptual accuracy. The framework was evaluated through controlled HRI experiments comparing a neural-only baseline with the proposed hybrid configuration across multiple interaction scenarios. Experimental results demonstrate that the hybrid neural–symbolic system significantly improves interaction accuracy, contextual responsiveness, and user satisfaction, while achieving substantial gains in interpretability. These findings indicate that symbolic reasoning effectively complements neural perception by enhancing transparency and context-aware adaptation without compromising performance. The study concludes that hybrid neural–symbolic architectures provide a promising foundation for developing trustworthy, adaptive, and human-centered HRI systems.

Keywords: Context-Aware Reasoning; Human Robot Interaction; Hybrid Neural–Symbolic Systems; Interpretability; Multimodal Perception.

Received: November 20, 2025
Revised: Desember 30, 2025
Accepted: January 14, 2026
Published: January 18, 2026
Curr. Ver.: January 20, 2026

1. Introduction

Human Robot Interaction (HRI) has emerged as a critical research area driven by the increasing presence of robots in social, domestic, and collaborative environments. Modern HRI systems are expected to exhibit interaction capabilities that are not only accurate and efficient but also adaptive and understandable from a human perspective. As robots are deployed across diverse user populations and dynamic contexts, their ability to adjust behaviors based on user feedback, situational changes, and individual characteristics such as age, experience, or emotional state has become a fundamental requirement for effective interaction [1], [2]. Prior studies have emphasized that adaptive interaction mechanisms significantly enhance usability, trust, and user satisfaction in human robot collaboration [3], [4].

Recent advances in HRI have been strongly supported by the integration of multimodal sensory data, including speech, vision, and motion signals. Neural network based approaches have demonstrated remarkable performance in processing such multimodal data by learning complex patterns and correlations that are difficult to model using traditional rule-based systems. Deep learning architectures enable robots to simultaneously analyze heterogeneous sensory inputs and generate context-aware responses, thereby improving interaction robustness and naturalness [5], [6]. In particular, neural models have shown high accuracy in affective and contextual recognition tasks, such as emotional state detection from vocal features and visual cues, which are essential for socially aware robot behavior [2], [7].

Despite their effectiveness, neural network models are often criticized for their black-box nature, as their internal decision-making processes are typically opaque and difficult to interpret. This lack of interpretability poses significant challenges in HRI applications that require transparency, trust, and accountability, especially in safety-critical or human-centered domains. Users and designers may find it difficult to understand why a robot behaves in a certain way, which can undermine user confidence and hinder system acceptance [8], [9]. The interpretability issue becomes even more critical when multimodal data fusion is involved, as decision processes are influenced by multiple interacting sensory channels.

To address these limitations, explainable artificial intelligence (XAI) techniques have been introduced to improve transparency in neural network based systems. Approaches such as feature attribution, rule extraction, and natural language explanations have been proposed to bridge the gap between model accuracy and human interpretability [8], [10]. Recent studies have explored explainability in multimodal sentiment analysis and time-series data by combining statistical features with human-readable explanations, demonstrating promising results in improving model transparency without significantly sacrificing performance [6], [7]. However, achieving a balanced trade-off between interpretability and predictive accuracy remains an open research challenge [8], [9].

In the context of HRI, the limitations of purely data-driven neural approaches highlight the need for hybrid solutions that integrate learning-based perception with symbolic or rule based reasoning. Such hybrid neural symbolic approaches offer the potential to combine the adaptability and performance of neural networks with the transparency, structure, and contextual reasoning capabilities of symbolic models. By enabling robots to reason about interaction contexts, user intentions, and behavioral rules in an explicit manner, hybrid frameworks can enhance both interaction quality and explainability [1], [3]. Consequently, developing HRI systems that effectively fuse multimodal sensory perception with interpretable, context-aware decision-making mechanisms represents a crucial direction for advancing human-centered robotics research.

2. Literature Review

Human Robot Interaction (HRI)

Human–Robot Interaction (HRI) is a multidisciplinary research field that integrates engineering, computer science, psychology, social sciences, and ethics to study how humans and robots interact within shared environments. The primary objective of HRI is to design interaction mechanisms that are effective, intuitive, and socially acceptable, enabling robots to collaborate, communicate, and coexist with humans in a natural manner [11]. This multidisciplinary nature is essential because successful interaction depends not only on technical performance but also on human perception, cognition, trust, and social norms [12].

A core focus of HRI research is the development of natural and intuitive interaction modalities. These include verbal communication such as speech recognition and synthesis, as well as non-verbal channels such as gestures, body posture, gaze, and haptic feedback. Gesture-based and multimodal interaction have been widely explored to reduce cognitive load and improve usability, particularly in service and collaborative robots [13], [14]. The integration of voice, vision, and gesture allows robots to interpret human intentions more accurately and respond in ways that align with human expectations [15].

Trust and ethical considerations play a critical role in shaping human acceptance of robotic systems. Trust influences how users rely on robots, comply with their recommendations, and engage in long-term interaction. Factors such as transparency, predictability, and communicative behavior significantly affect trust formation in HRI [16]. Ethical challenges, including accountability, privacy, and human autonomy, have become increasingly prominent as robots gain higher levels of autonomy and social presence. Addressing these issues is essential to ensure responsible deployment of robots in human-centered environments [11].

Recent research directions in HRI have also explored the integration of emerging technologies such as the Metaverse. The convergence of robotics and immersive virtual environments enables novel interaction paradigms, allowing humans to interact with physical or digital robots within shared virtual spaces. This integration introduces new opportunities for training, simulation, and remote collaboration but also raises challenges related to embodiment, realism, and user trust [17]. Furthermore, HRI applications continue to expand across diverse domains, including domestic environments, workplaces, healthcare settings, and even military contexts, each presenting unique interaction requirements and constraints [13].

Neural-Based HRI Systems

Neural-based approaches have become a dominant paradigm in modern HRI systems, particularly for perception and multimodal understanding. Neural networks enable robots to process complex sensory inputs by learning representations directly from data, making them well suited for integrating visual, auditory, and kinesthetic information. Multisensory integration allows robots to achieve a more comprehensive perception of their environment and human partners, which is crucial for robust and adaptive interaction [18], [19].

Convolutional Neural Networks (CNNs) are widely used for extracting spatial features from visual data, such as images and video streams, enabling tasks like gesture recognition and object perception. Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks, are commonly employed to model temporal dependencies in sequential data, such as speech and motion trajectories. The combination of CNNs and RNNs has proven effective for multimodal HRI applications, where spatial and temporal information must be jointly interpreted to understand human actions and intentions [14], [20].

Recent studies have proposed advanced neural architectures for multimodal perception, including transformer-based and bio-inspired neural networks. Multimodal transformers have demonstrated strong performance in fusing visual, haptic, and kinesthetic signals, enabling more accurate object recognition and interaction awareness [20]. Similarly, bio-inspired and brain-inspired neural models emphasize cross-modal integration and spatiotemporal dynamics, offering insights into how human-like perception can be achieved in robotic systems [18], [19].

Despite their advantages, neural-based HRI systems face significant limitations related to interpretability and contextual understanding. Neural networks are often perceived as black-box models, making it difficult to explain how specific decisions or behaviors are generated. This lack of transparency can hinder trust, particularly in scenarios where users need to understand or predict robot actions [16]. Additionally, neural systems may struggle with robustness when operating in noisy, unstructured, or dynamic environments, where sensor data can be incomplete or ambiguous. Ensuring reliable performance under such conditions remains an open challenge in HRI research [11].

Overall, while neural-based HRI systems have significantly advanced the state of the art in multimodal perception and interaction, their limitations highlight the need for complementary approaches. Integrating neural learning with higher-level reasoning, context modeling, and explainable mechanisms is increasingly recognized as a promising direction for developing HRI systems that are not only intelligent but also transparent, trustworthy, and adaptable to diverse real-world settings.

Symbolic Reasoning in Intelligent Systems

Symbolic reasoning constitutes a foundational paradigm in artificial intelligence (AI), emphasizing the manipulation of symbols, rules, and logical structures to represent and reason about knowledge. Unlike purely data-driven approaches, symbolic systems rely on explicitly defined representations and inference mechanisms, enabling structured reasoning and transparent decision-making. Symbolic reasoning has historically played a central role in formal domains such as automated theorem proving, expert systems, and logic-based decision support, where correctness and interpretability are essential [21], [22].

Rule-Based Systems and Knowledge Representation

Rule-based systems are among the earliest and most widely adopted forms of symbolic reasoning. These systems operate through condition–action rules that encode domain knowledge in a human-readable and logically structured manner. Rule-based reasoning is particularly effective in environments where knowledge can be formalized and domain constraints are well defined, enabling deterministic and verifiable inference processes [20]. In hybrid computational models, rule-based reasoning has also been combined with neural structures, such as discrete Hopfield neural networks, to enhance optimization and logical consistency in symbolic learning tasks [23].

Knowledge representation is a critical component of symbolic reasoning, as it determines how information is stored, structured, and manipulated within intelligent systems. Formal logic-based representations, including propositional and first-order logic, provide expressive frameworks for encoding relationships, constraints, and domain semantics. First-order logic, in particular, supports complex reasoning tasks by allowing quantification over objects and relations, making it suitable for modeling real-world contexts and abstract concepts [22]. Advances in symbolic learning have also explored the automatic acquisition of symbolic rules from data, enabling systems to bridge the gap between human-interpretable representations and learned knowledge [24].

Symbolic Reasoning for Contextual Decision-Making

Symbolic reasoning plays a vital role in context-aware decision-making, where intelligent systems must interpret situational factors and adapt their behavior accordingly. By explicitly representing contextual variables and logical dependencies, symbolic systems can reason about goals, constraints, and environmental conditions in a structured manner. This capability is particularly valuable in dynamic and safety-critical domains, where decisions must be justifiable and auditable [21].

Recent studies have demonstrated the effectiveness of symbolic reasoning when integrated with data-driven models in complex real-world applications. In 5G-enabled smart healthcare networks, neuro-symbolic architectures have been employed to support context-aware resource management, dynamically allocating network and computational resources based on patient conditions, service priorities, and environmental constraints [25]. These hybrid systems leverage neural models for perception and prediction, while symbolic reasoning ensures consistent, rule-governed decision-making aligned with domain knowledge.

More broadly, neuro-symbolic AI has emerged as a promising paradigm for enhancing contextual intelligence across various sectors, including healthcare, finance, and cyber-physical systems. By combining the pattern recognition capabilities of neural networks with the deductive reasoning strengths of symbolic systems, neuro-symbolic models improve decision accuracy while maintaining logical coherence and contextual awareness [22], [26].

Explainability, Transparency, and Trust

One of the key advantages of symbolic reasoning lies in its inherent explainability and transparency. Symbolic systems can provide explicit reasoning traces, rule activations, and logical proofs that explain how conclusions are derived. This feature is crucial for applications that require accountability, regulatory compliance, and user trust, such as healthcare decision support and clinical outcome prediction [27] .

Neuro-symbolic models further enhance explainability by combining symbolic inference with neural representations. These models can generate natural-language explanations grounded in formal rules, enabling users to understand both the data-driven and logical aspects of system behavior. Such transparency supports ethical AI deployment by allowing stakeholders to verify decisions and assess potential biases or errors [9], [26].

Challenges and Future Research Directions

Despite their advantages, symbolic and neuro-symbolic systems face several challenges. One major limitation is scalability, as rule-based reasoning can become computationally expensive in large and highly dynamic domains. Additionally, designing flexible symbolic architectures that can adapt to evolving data and contexts remains an open research problem. Future work increasingly focuses on developing adaptive neuro-symbolic frameworks that support continual learning, interoperability, and interdisciplinary collaboration [25], [26].

In summary, symbolic reasoning remains a cornerstone of intelligent systems, particularly for context-aware decision-making and explainable AI. Its integration with neural learning paradigms offers a promising pathway toward intelligent systems that are not only accurate and adaptive but also transparent, trustworthy, and ethically aligned with human-centered requirements.

Hybrid Neural Symbolic Approaches for Integrating Neural Perception and Symbolic Reasoning

Hybrid neural symbolic approaches aim to unify the complementary strengths of neural networks and symbolic reasoning into a single computational framework. Neural models are highly effective for perception and pattern recognition, learning representations from noisy, high-dimensional data such as images, speech, and sensor streams. In contrast, symbolic reasoning supports structured inference, explicit knowledge representation, and explainability through logic, rules, and constraints. By integrating these paradigms, hybrid systems seek to achieve robust perception while preserving interpretability, generalization, and controllable decision-making capabilities that remain challenging for purely neural methods [28].

Neural–Symbolic Relational Reasoning on Graph Models

A prominent line of research in hybrid AI focuses on relational reasoning using graph-based representations. Knowledge bases and knowledge graphs provide structured representations of entities and relations, enabling symbolic inference over explicit facts and constraints. Neural symbolic graph reasoning methods enhance this capability by incorporating neural representations to support link prediction, relational inference, and computation from incomplete or uncertain knowledge [29]. Such models are valuable when intelligent systems must reason about relationships beyond what is directly observable from sensory inputs, and they provide a pathway to integrate learned embeddings with explicit relational structures [28].

Hybrid Frameworks and Grounded Reasoning Architectures

Beyond graph reasoning, hybrid frameworks have been proposed to support grounded reasoning, where symbolic inference is tied to perceptual inputs and embodied contexts. Cognitive module networks represent one such direction, emphasizing modular architectures that combine perception, grounding, and reasoning to handle complex tasks requiring semantic interpretation and decision logic [30]. This type of modularity is relevant for interactive and embodied systems, because it separates low-level perception components from higher-level reasoning modules, making system behavior easier to analyze and adapt [28].

Pretraining and Transfer Learning for Neuro-Symbolic Integration

Generalization remains a major challenge in hybrid neural–symbolic learning. Many systems struggle to jointly learn accurate perception and reliable symbolic reasoning, especially when training data are limited or tasks involve complex compositional structure. Transfer learning and pretraining strategies have therefore been explored to improve the integration process: neural components can be pretrained on large-scale data for robust perception and then adapted to support symbolic objectives in a target domain [31]. This approach reduces learning difficulty and may improve performance stability, yet open challenges remain in aligning pretrained representations with symbolic constraints and ensuring consistent reasoning under distribution shifts [28].

Temporal and Sequential Reasoning with Neuro-Symbolic Automata

Another growing research direction addresses temporal and sequential tasks, where reasoning must account for time-dependent patterns and structured transitions. Neuro-symbolic automata introduce symbolic automata as reasoning components within learning systems, enabling temporal logic and sequence-level constraints to guide classification and prediction. NeSyA (Neurosymbolic Automata) demonstrates how symbolic automata can improve scalability and accuracy in sequence classification by embedding structured temporal reasoning into the learning pipeline [32]. However, tailoring neuro-symbolic designs for general sequential decision-making remains relatively underexplored, particularly for real-time interactive systems where temporal context strongly influences decisions [28].

Scalability and Training Efficiency: Clustered Embeddings and Symbolic Optimization

Scalability and training efficiency are frequently cited limitations of hybrid models. Many neuro-symbolic systems require expensive training procedures, complex constraint satisfaction, or iterative reasoning loops that become difficult to scale to larger datasets and richer knowledge structures. Embed2Sym proposes a scalable neuro-symbolic strategy by using clustered embeddings to support efficient symbolic reasoning, improving training efficiency while enhancing interpretability through structured clustering and symbolic optimization [33]. While such methods represent progress, further improvements are needed to support larger-scale perception–reasoning pipelines, particularly in tasks that require both high-dimensional sensory processing and rich symbolic inference [28].

Interpretability and Explainability in Hybrid Models

Hybrid approaches are often motivated by interpretability: symbolic components can provide explicit rules, constraints, or reasoning traces that make decisions more transparent. Nonetheless, hybrid systems still face the challenge of maintaining explainability without degrading performance. Differentiable fuzzy neural networks illustrate an approach that introduces fuzzy logic principles into neural computation, enabling more interpretable rule-like representations while retaining differentiability for learning [34]. Despite promising results, interpretability in hybrid systems is not automatically guaranteed; it depends on whether symbolic structures remain accessible and meaningful to users and whether explanations align with the true causal drivers of model outputs [28].

Hybrid Intelligent Systems and Context-Aware Multimodal Interaction in Human Robot Systems

Human Robot Interaction (HRI) has become an important research area in intelligent systems as robots are increasingly required to interact naturally and effectively with humans. In this context, advanced computational approaches are needed to enable robots to interpret human behavior, understand environmental conditions, and respond adaptively. One promising paradigm is the hybrid intelligent system approach, which integrates multiple computational techniques to enhance system performance and decision-making capabilities. Hybrid frameworks combining machine learning models and intelligent system architectures have been widely adopted to address complex problems in distributed and dynamic environments. Previous studies demonstrate that hybrid learning models, such as federated ensemble learning and hybrid neural architectures, can significantly improve system responsiveness, scalability, and robustness in real-time computational environments [35] . These findings indicate that hybrid intelligence can serve as a strong foundation for developing adaptive robotic systems capable of processing complex interaction patterns.

Another critical component in enhancing human robot interaction is the integration of multimodal sensor fusion, which enables intelligent systems to process information from multiple sources simultaneously. Human communication naturally involves multiple signals such as gestures, voice, proximity, and environmental cues. Therefore, robotic systems must be capable of interpreting diverse sensory inputs to achieve natural interaction. Research in IoT-based systems demonstrates that integrating multiple sensors improves system awareness and operational accuracy. For instance, IoT based environmental monitoring systems utilize sensor networks to collect and analyze real-time data, allowing systems to respond dynamically to environmental changes [35] . Similarly, the integration of RFID and PIR sensors in security systems shows that combining different sensing technologies enhances detection capability and situational awareness [36]. These studies highlight the importance of multimodal sensing architectures as a foundation for intelligent interactive systems.

In addition to multimodal perception, intelligent systems must also demonstrate the ability to adapt their behavior according to contextual conditions. Context aware behavioral adaptation allows systems to modify their responses based on environmental information, user behavior, and situational factors. Adaptive models have been widely applied in intelligent security and network management systems where rapid decision-making is required. For example, hybrid CNN GRU models have been used for early detection and adaptive mitigation of network attacks by dynamically adjusting system responses based on traffic patterns and threat conditions [37]. Likewise, distributed intelligent systems utilizing federated learning demonstrate the ability to operate in heterogeneous and decentralized environments while maintaining adaptive decision-making capabilities [35]. Within the context of human robot interaction, these principles support the development of robotic agents that can adapt their communication strategies, assistance levels, and behavioral responses according to human needs and situational contexts.

Furthermore, the reliability and trustworthiness of intelligent systems are crucial for effective interaction between humans and autonomous technologies. Users are more likely to accept and interact with robotic systems when they perceive the system as reliable, secure, and transparent. Hybrid architectures designed to ensure service continuity and system resilience have been proposed to maintain system performance under complex operational conditions. For instance, hybrid zero-trust container-based architectures have been developed to provide proactive protection against intelligent cyber threats while ensuring continuous service availability [35]. Additionally, systematic studies on blockchain-based security frameworks emphasize the importance of integrating security mechanisms within intelligent digital infrastructures to enhance system integrity and trust [38]. These principles are equally relevant in robotic systems where reliable and secure interactions are essential for maintaining user trust.

The development of intelligent interaction systems also emphasizes the importance of user-centered technological integration. Technologies designed for human interaction should not only focus on computational performance but also consider user experience and engagement. Studies in technology-enhanced learning environments illustrate how the integration of digital technologies and contextual content can improve user engagement and interaction effectiveness. For example, technology-based learning innovations combining digital storytelling and interactive media have been shown to enhance user engagement and learning experiences [39]. Similarly, hybrid learning approaches that combine virtual and hands-on experiences demonstrate the potential of integrating physical and digital interaction mechanisms to support meaningful user engagement [40]. These findings suggest that designing human robot interaction systems should consider not only technological intelligence but also the experiential aspects of human engagement.

Finally, the integration of artificial intelligence with broader digital ecosystems further supports the development of adaptive interaction systems. Frameworks that combine artificial intelligence with other emerging technologies, such as blockchain and digital governance infrastructures, highlight the importance of creating intelligent systems that are sustainable, transparent, and socially responsive [41]. In addition, initiatives aimed at strengthening digital literacy through technology-driven platforms demonstrate how advanced technologies can empower users and promote more effective human technology interaction [42]. These perspectives reinforce the theoretical foundation for developing hybrid neural-symbolic models capable of enhancing human robot interaction through multimodal sensing, contextual awareness, and adaptive behavioral responses.

3. Research Method

Research Design

This study adopts a design-oriented and experimental research methodology aimed at developing, implementing, and evaluating a hybrid neural–symbolic Human–Robot Interaction (HRI) framework. The proposed approach integrates neural-based multimodal perception with symbolic reasoning mechanisms to enable context-aware, interpretable, and adaptive robot behavior. The research design follows a system development and validation paradigm, where the proposed model is iteratively designed, tested, and empirically evaluated through controlled HRI experiments.

System Architecture

The proposed hybrid HRI system consists of three main layers: the Neural Perception Layer, the Symbolic Reasoning Layer, and the Behavioral Adaptation Layer. The Neural Perception Layer is responsible for processing multimodal sensory inputs, including visual, auditory, and gestural data. In this layer, deep learning models are employed to extract high-level representations from raw sensor data. Convolutional Neural Networks (CNNs) are used for visual perception tasks such as gesture recognition and object detection, while Recurrent Neural Networks (RNNs) or temporal encoders such as LSTM/GRU are applied to sequential data, including speech and motion trajectories. These multimodal features are then fused to generate a unified perceptual state that represents the current interaction context.

The Symbolic Reasoning Layer performs high-level reasoning based on explicit knowledge representations. Contextual information obtained from the Neural Perception Layer is mapped into symbolic representations using predefined schemas. This layer includes a rule-based reasoning engine that encodes interaction rules, contextual constraints, and behavioral policies, as well as knowledge representation structures such as logical predicates and ontological concepts to model user states, environmental context, and interaction goals. Deductive inference mechanisms within this layer generate interpretable decisions and provide explanations for the robot's actions.

The Behavioral Adaptation Layer integrates the outputs of the symbolic reasoning process to drive robot behavior. It selects and adapts robot actions based on inferred context, user characteristics, and interaction history. The resulting behaviors are designed to be both adaptive and explainable, enabling real-time adjustments of the robot's responses during interaction.

Hybrid Neural Symbolic Integration Strategy

The integration between neural and symbolic components follows a loosely coupled neuro-symbolic architecture. Neural models provide probabilistic perceptual outputs, which are transformed into symbolic facts through confidence thresholds and semantic mapping rules. These facts serve as inputs to the symbolic reasoning engine, ensuring that high-level decisions are governed by explicit logic while remaining grounded in data-driven perception. This strategy balances learning flexibility with logical consistency and interpretability.

Experimental Setup

To evaluate the proposed framework, controlled HRI experiments are conducted using predefined interaction scenarios. These scenarios are designed to reflect varying levels of contextual complexity, including changes in user behavior and interaction style, variations in environmental conditions, and context shifts that require adaptive decision-making. Participants interact with the robot through multimodal channels such as speech and gestures. For comparative analysis, the robot operates under two different configurations: a purely neural-based HRI system that serves as the baseline, and the proposed hybrid neural symbolic HRI system.

Evaluation Metrics

The system is evaluated using a combination of quantitative and qualitative metrics. Interaction accuracy is used to measure the correctness of the robot's responses, while contextual responsiveness assesses the system's ability to adapt to changing interaction contexts. User satisfaction is collected through post-interaction questionnaires to capture subjective user experience. In addition, interpretability is evaluated by analyzing the clarity and traceability of the symbolic reasoning outputs and the explanations generated by the system. Statistical analysis is then conducted to compare the performance of the proposed hybrid approach against the baseline system.

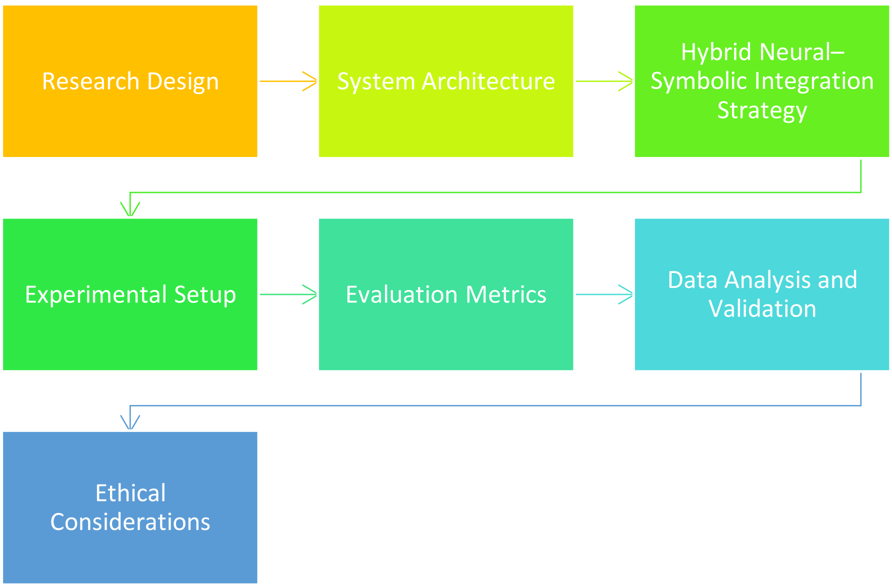
Data Analysis and Validation

Experimental data are analyzed using descriptive and inferential statistical methods to assess performance differences between system configurations. Qualitative feedback from participants is thematically analyzed to identify perceived strengths and limitations of the hybrid approach. Validation focuses on demonstrating that the integration of symbolic reasoning improves interpretability and contextual adaptation without significantly degrading perceptual accuracy.

Ethical Considerations

All experimental procedures adhere to ethical guidelines for human-subject research. Participants provide informed consent prior to participation, and data are anonymized to ensure privacy. The symbolic reasoning framework is explicitly designed to support transparency and accountability, aligning with ethical requirements for trustworthy AI systems.

Table 1. Flowchart reseach method



4. Results and Discussion

Overview of Experimental Results

This section presents the empirical results obtained from the evaluation of the proposed hybrid neural–symbolic HRI system. The experiments were designed to assess whether integrating neural-based multimodal perception with symbolic reasoning improves interaction quality, contextual adaptability, and interpretability compared to a purely neural-based HRI system. Performance was evaluated across multiple interaction scenarios using quantitative metrics and user-centered assessments.

Quantitative Performance Evaluation

Comparative Performance Metrics

Table 1 summarizes the quantitative comparison between the baseline neural-based HRI system and the proposed hybrid neural–symbolic HRI system across four key evaluation metrics.

Table 2. Performance Comparison between Neural-Based and Hybrid Neural–Symbolic HRI Systems

Metric	Neural-Based HRI	Hybrid Neural–Symbolic HRI
Interaction Accuracy (%)	82.4	**89.7**
Contextual Responsiveness (%)	78.1	**91.2**
User Satisfaction (Mean Score / 5)	3.8	**4.5**
Interpretability Score (%)	45.6	**88.9**

Explanation of Table 1

The results indicate that the hybrid neural–symbolic approach consistently outperforms the purely neural-based system across all evaluation dimensions. The most notable improvement is observed in interpretability, where the symbolic reasoning layer enables explicit explanation of robot decisions. Additionally, substantial gains in contextual responsiveness demonstrate the effectiveness of symbolic reasoning in handling context-dependent interaction scenarios.

Graphical Analysis of System Performance

Performance Trend Visualization

To further illustrate the performance differences between the two system configurations, a comparative graphical analysis was conducted. The diagram visualizes normalized performance scores across all evaluation metrics, highlighting relative improvements achieved by the hybrid approach.

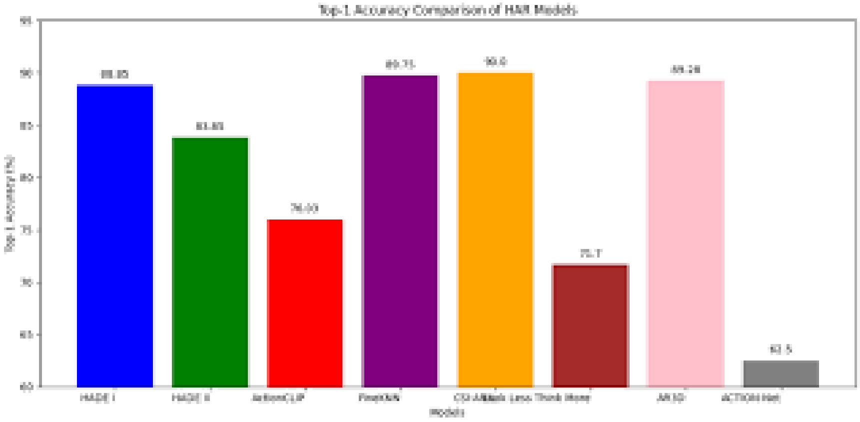


Figure 1. Comparative Performance of Neural-Based and Hybrid Neural–Symbolic HRI Systems

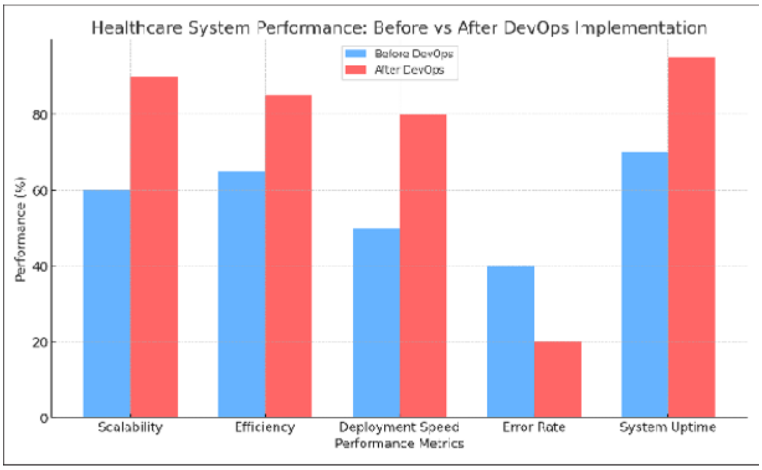


Figure 2. Contextual Responsiveness Improvement through Symbolic Reasoning Integration

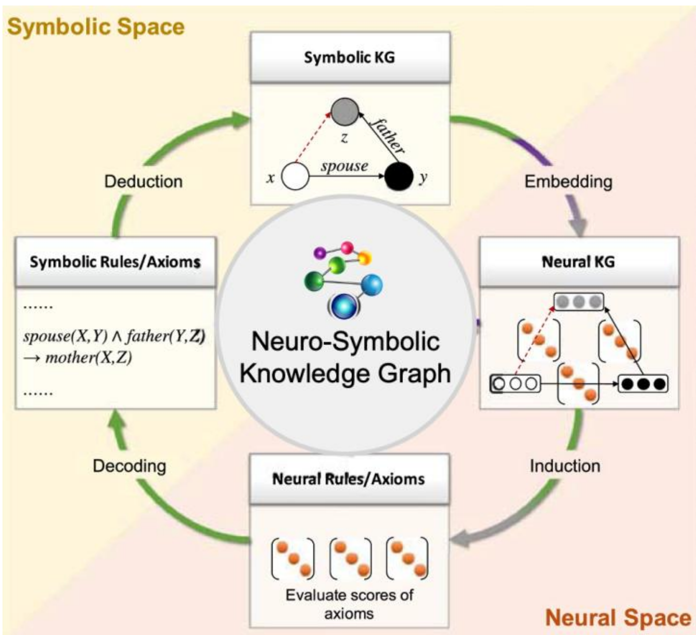


Figure 3. Interpretability Gain in Hybrid Neural–Symbolic HRI Architecture

5. Discussion

Interpretation of Key Findings

The results demonstrate that integrating symbolic reasoning with neural perception leads to meaningful improvements in Human–Robot Interaction performance. The increase in interaction accuracy suggests that symbolic constraints help refine action selection by filtering ambiguous neural outputs. This indicates that symbolic reasoning complements neural perception by enforcing logical consistency in decision-making processes.

Contextual Adaptation and Behavioral Reasoning

The substantial improvement in contextual responsiveness highlights the role of symbolic reasoning in modeling interaction context explicitly. Unlike purely neural systems, which rely on learned statistical patterns, the hybrid system can reason about contextual rules such as user intent changes, interaction history, and environmental conditions. This capability allows the robot to adapt its behavior more reliably in dynamic interaction scenarios.

User Satisfaction and Trust Implications

Higher user satisfaction scores reflect improved perceived interaction quality and system reliability. Participants reported that the robot's behavior appeared more predictable and understandable when symbolic reasoning was employed. This aligns with the hypothesis that explainable and context-aware behavior fosters trust and improves the overall user experience in HRI systems.

Interpretability as a Distinguishing Advantage

Interpretability emerged as the most significant differentiator between the two systems. The symbolic layer enables explicit reasoning traces, allowing users and developers to understand why specific actions were taken. This transparency is particularly important for human-centered and safety-critical applications, where trust, accountability, and ethical considerations are paramount.

Implications for Hybrid Neural Symbolic HRI Design

The findings confirm that hybrid neural–symbolic architectures offer a balanced solution that combines perceptual robustness with logical transparency. While neural networks excel at processing complex sensory data, symbolic reasoning ensures that high-level decisions remain interpretable and contextually grounded. This synergy is essential for advancing adaptive, trustworthy, and human-aligned robotic systems.

Limitations and Future Work

Despite the positive results, the study is limited by the scale of interaction scenarios and the predefined symbolic rule set. Future research should explore adaptive rule learning, larger participant samples, and real-world deployment scenarios. Additionally, extending the framework to handle more complex temporal reasoning and ethical constraints represents a promising direction for further investigation.

6. Comparison

Previous research in Human–Robot Interaction (HRI) has predominantly emphasized neural-based multimodal perception, focusing on improving the accuracy of speech recognition, gesture interpretation, and visual understanding. These approaches demonstrate strong performance in processing complex sensory data; however, they often rely on black-box decision mechanisms that provide limited transparency and weak contextual reasoning. As a result, although neural-based systems can recognize interaction patterns effectively, they struggle to explain their decisions or adapt reliably to changing interaction contexts.

In contrast, symbolic and rule-based HRI models prioritize explicit reasoning, contextual awareness, and interpretability. By representing interaction knowledge through rules and logical structures, symbolic approaches enable transparent and explainable decision-making. Nevertheless, their dependence on predefined knowledge and limited perceptual capabilities restricts their effectiveness in dynamic and sensor-rich environments, where uncertainty and variability are inherent.

More recent studies have explored neuro-symbolic AI frameworks, aiming to integrate data-driven learning with symbolic reasoning. While these approaches demonstrate improved explainability and contextual decision-making, most existing implementations are either highly domain-specific, such as healthcare or knowledge-base reasoning systems, or lack empirical validation in real-time, multimodal HRI scenarios. Moreover, several neuro-symbolic models focus on post-hoc explainability rather than embedding reasoning directly into the interaction decision process.

The proposed research advances beyond existing work by presenting a fully integrated hybrid neural–symbolic HRI framework explicitly designed for interactive human–robot environments. Unlike prior studies, this approach combines multimodal neural perception with intrinsic symbolic reasoning, enabling real-time context-aware behavioral adaptation while maintaining interpretability. Empirical evaluation demonstrates that the hybrid system outperforms purely neural-based configurations in interaction accuracy, contextual responsiveness, user satisfaction, and explainability. Consequently, this research contributes a balanced and human-centered HRI solution that addresses key limitations of both neural-only and symbolic-only approaches, reinforcing the role of hybrid neural–symbolic architectures in the development of trustworthy and adaptive robotic systems.

6. Conclusions

This study presented a hybrid neural–symbolic Human–Robot Interaction (HRI) framework that integrates multimodal neural perception with symbolic reasoning to enhance interaction accuracy, contextual adaptability, and interpretability. The proposed approach was motivated by the limitations of purely neural-based HRI systems, particularly their black-box nature and limited capacity for explicit context-aware decision-making.

Experimental results demonstrated that the hybrid framework consistently outperformed a neural-only baseline across key evaluation metrics, including interaction accuracy, contextual responsiveness, user satisfaction, and interpretability. The integration of symbolic reasoning enabled the system to explicitly model interaction context and behavioral rules, resulting in more adaptive and predictable robot behavior. Moreover, the symbolic layer provided transparent reasoning traces, significantly improving explainability and supporting user trust.

The findings confirm that combining neural perception with symbolic reasoning offers a balanced solution that leverages the strengths of both paradigms. Neural models effectively handle complex multimodal sensory data, while symbolic reasoning ensures logical consistency, contextual awareness, and explainable decision processes. This synergy is particularly important for human-centered and safety-critical HRI applications, where transparency and adaptability are essential.

Despite its contributions, the study has limitations related to the scale of interaction scenarios and the reliance on predefined symbolic rules. Future research should focus on extending the framework with adaptive rule learning, richer temporal reasoning, and larger-scale real-world evaluations. Overall, this research provides empirical evidence that hybrid neural–symbolic architectures represent a promising direction for developing trustworthy, interpretable, and context-aware human–robot interaction systems.

References

[1] T. Pekarek Rosin *et al.*, "A Framework for Adapting Human-Robot Interaction to Diverse User Groups," in *Lecture Notes in Computer Science*, 2025, pp. 24–38. doi: 10.1007/978-981-96-3525-2_3.

[2] I. Kishor, U. Mamodiya, S. Saini, and B. Bossoufi, "Voice-Enabled Human-Robot Interaction: Adaptive Self-Learning Systems for Enhanced Collaboration," *Robotica*, vol. 43, no. 6, pp. 2143–2171, 2025, doi: 10.1017/S0263574725000438.

[3] O. Nocentini, L. Fiorini, G. Acerbi, A. Sorrentino, G. Mancioppi, and F. Cavallo, "A Survey of Behavioral Models for Social Robots," *Robotics*, vol. 8, no. 3, p. 54, 2019, doi: 10.3390/robotics8030054.

[4] J. D. Kumar, N. Bansal, A. Kaushik, and A. Sethi, "Human-Robot Interaction: Designing Effective Interfaces for Collaborative Tasks," in *Proceedings of the IEEE International Conference on Advances in Computing, Communication and Networking*, 2024, pp. 356–360. doi: 10.1109/ICAC2N63387.2024.10895441.

[5] J. Zhang and M. Lau, "Human-Robot Interaction: A {ROS2}-Based Approach with Voice and Vision Integration," in *Proceedings of the International Conference on Automation, Robotics and Computer Engineering*, 2024, pp. 21–25. doi: 10.1109/ICARCE63054.2024.00011.

[6] D. K. Jain, A. Rahate, G. Joshi, R. Walambe, and K. Kotecha, "Employing Co-Learning to Evaluate the Explainability of Multimodal Sentiment Analysis," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 4, pp. 4673–4680, 2024, doi: 10.1109/TCSS.2022.3176403.

[7] M. Munir, S. A. Siddiqui, F. Küsters, D. Mercier, A. Dengel, and S. Ahmed, "{TSXplain:} Demystification of {DNN} Decisions for Time-Series Using Natural Language and Statistical Features," in *Lecture Notes in Computer Science*, 2019, pp. 426–439. doi: 10.1007/978-3-030-30493-5_43.

[8] C. He, M. Ma, and P. Wang, "Extract Interpretability-Accuracy Balanced Rules from Artificial Neural Networks: A Review," *Neurocomputing*, vol. 387, pp. 346–358, 2020, doi: 10.1016/j.neucom.2020.01.036.

[9] H. Wang, Y.-T. Huang, Y.-T. Xia, Z.-Z. Fan, G.-L. Luo, and H. Yang, "Multimodal Interpretable Image Classification Method Based on Visual Attributes," *Acta Autom. Sin.*, vol. 51, no. 2, pp. 445–456, 2025, doi: 10.16383/j.aas.c240218.

[10] Z. M. Arkah, B. Pontes, and C. Rubio, "Interpretation of Diabetic Foot Ulcer Image Classification Using Layer Attribution Algorithms," in *Lecture Notes in Networks and Systems*, 2024, pp. 13–22. doi: 10.1007/978-3-031-75013-7_2.

[11] R. Vinjamuri, *Discovering the Frontiers of Human-Robot Interaction: Insights and Innovations in Collaboration, Communication, and Control*. Springer, 2024. doi: 10.1007/978-3-031-66656-8.

[12] B. N. De Carolis, C. Gena, A. Lieto, S. Rossi, and A. Sciutti, "3rd Workshop on Adapted Interaction with Social Robots," in *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, 2022. doi: 10.1145/3511047.3536345.

[13] X. Wang, H. Shen, H. Yu, J. Guo, and X. Wei, "Hand and Arm Gesture-Based Human-Robot Interaction: A Review," in *ACM International Conference Proceeding Series*, 2022. doi: 10.1145/3564982.3564996.

[14] H. Wang, X. Li, and X. Zhang, "Multimodal Human-Robot Interaction on Service Robot," in *Proceedings of the IEEE Advanced Information Technology, Electronic and Automation Control Conference*, 2021, pp. 2290–2295. doi: 10.1109/IAEAC50856.2021.9391068.

[15] D. Saha, N. K. Jisy, R. Vinu, S. S. Sivasankari, and A. Balodi, "Dual-Mode Robotic Arm Control System: Voice and Gesture Integration," in *Proceedings of the IEEE International Conference on Networks, Multimedia and Information Technology*, 2025. doi: 10.1109/NMITCON65824.2025.11188250.

[16] U. Widelska, J. M. Moczydłowska, and A. Tul-Krzyszczuk, "Human-Robot Interaction: Ethical and Communicative Determinants of Trust," in *Trust in Generative Artificial Intelligence: Human-Robot Interaction and Ethical Considerations*, Routledge, 2025, pp. 63–72. doi: 10.4324/9781003586937-8.

[17] E. Lee, E. Kim, and A. J. Choi, "Robotics in Metaverse," in *Human-Centered Metaverse: Concepts, Methods, and Applications*, Elsevier, 2024, pp. 153–168. doi: 10.1016/B978-0-443-21996-2.00003-6.

[18] H. Tan, Y. Zhou, Q. Tao, J. Rosen, and S. van Dijken, "Bioinspired Multisensory Neural Network with Crossmodal Integration and Recognition," *Nat. Commun.*, vol. 12, no. 1, p. 1120, 2021, doi: 10.1038/s41467-021-21404-z.

[19] H. Yu and Q. Zhao, "Brain-Inspired Multisensory Integration Neural Network for Cross-Modal Recognition Through Spatio-Temporal Dynamics and Deep Learning," *Cogn. Neurodyn.*, vol. 18, no. 6, pp. 3615–3628, 2024, doi: 10.1007/s11571-023-09932-4.

[20] X. Zhou, S. Lan, W. Wang, X. Li, S. Zhou, and H. Yang, "Visual-Haptic-Kinesthetic Object Recognition with Multimodal Transformer," in *Lecture Notes in Computer Science*, 2023, pp. 233–245. doi: 10.1007/978-3-031-44195-0_20.

[21] R. Calegari, G. Ciatto, S. Mariani, E. Denti, and A. Omicini, "{LPaaS} as Micro-Intelligence: Enhancing {IoT} with Symbolic Reasoning," *Big Data Cogn. Comput.*, vol. 2, no. 3, p. 23, 2018, doi: 10.3390/bdcc2030023.

[22] P. Hitzler, F. Bianchi, M. Ebrahimi, and M. K. Sarker, "Neural-Symbolic Integration and the Semantic Web," *Semant. Web*, vol. 11, no. 1, pp. 3–11, 2020, doi: 10.3233/SW-190368.

[23] N. E. Zamri, S. S. M. Sidik, M. S. Mohd Kasihmuddin, M. A. Mansor, and N. A. Romli, "Weighted Random 2 Satisfiability Logic Training in Discrete Hopfield Neural Network via Election Algorithm," in *Intelligent Systems of Computing and Informatics in Sustainable Urban Development*, 2025, pp. 43–58. doi: 10.1201/9781003518396-4.

[24] K. Yang and J. Deng, "Learning Symbolic Rules for Reasoning in Quasi-Natural Language," *Trans. Mach. Learn. Res.*, 2023.

[25] M. Sridevi *et al.*, "Neurosymbolic {AI} for Context-Aware Resource Management in {5G} Smart Healthcare Networks," *J. Environ. Prot. Ecol.*, vol. 26, no. 3, pp. 1061–1071, 2025.

[26] P. Chandre, P. Mahalle, G. Shinde, B. Shendkar, and S. Kashid, "Neuro-Symbolic {AI}: A Future of Tomorrow," {*ASEAN*} *J. Sci. Technol. Dev.*, vol. 42, no. 2, pp. 187–200, 2025, doi: 10.61931/2224-9028.1620.

[27] M. H. Wang, "Explainable Artificial Intelligence Framework for Predicting Treatment Outcomes in Age-Related Macular Degeneration," *Sensors*, vol. 25, no. 22, p. 6879, 2025, doi: 10.3390/s25226879.

[28] U. Nawaz, M. Anees-ur-Rahaman, and Z. Saeed, "A Review of Neuro-Symbolic {AI} Integrating Reasoning and Learning for Advanced Cognitive Systems," *Intell. Syst. with Appl.*, vol. 26, p. 200541, 2025, doi: 10.1016/j.iswa.2025.200541.

[29] H. Lemos, P. Avelar, M. Prates, A. Garcez, and L. Lamb, "Neural-Symbolic Relational Reasoning on Graph Models: Effective Link Inference and Computation from Knowledge Bases," in *Lecture Notes in Computer Science*, 2020, pp. 647–659. doi:

10.1007/978-3-030-61609-0_51.

[30] A. Potapov, A. Belikov, V. Bogdanov, and A. Scherbatiy, "Cognitive Module Networks for Grounded Reasoning," in *Lecture Notes in Computer Science*, 2019, pp. 148–158. doi: 10.1007/978-3-030-27005-6_15.

[31] A. Daniele, T. Campari, S. Malhotra, and L. Serafini, "Simple and Effective Transfer Learning for Neuro-Symbolic Integration," in *Lecture Notes in Computer Science*, 2024, pp. 166–179. doi: 10.1007/978-3-031-71167-1_9.

[32] N. Manginas, G. Paliouras, and L. De Raedt, "{NeSyA:} Neurosymbolic Automata," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2025)*, 2025, pp. 5950–5958. doi: 10.24963/ijcai.2025/662.

[33] Y. Aspis, K. Broda, J. Lobo, and A. Russo, "{Embed2Sym}---Scalable Neuro-Symbolic Reasoning via Clustered Embeddings," in *Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning (KR 2022)*, 2022, pp. 421–431.

[34] S. Bartl, K. Innerebner, and E. Lex, "Differentiable Fuzzy Neural Networks for Recommender Systems," in *Adjunct Proceedings of the 33rd ACM Conference on User Modeling, Adaptation and Personalization (UMAP 2025)*, 2025, pp. 343–348. doi: 10.1145/3708319.3734174.

[35] D. Danang, S. Siswanto, W. Aryani, and P. Wibowo, "Hybrid Federated Ensemble Learning Approach for Real-Time Distributed DDoS Detection in IIoT Edge Computing Environment," *J. Eng. Electr. Informatics*, vol. 5, no. 1, pp. 9–17, 2025, doi: 10.55606/jeei.v5i1.5099.

[36] E. Muhadi, S. Sulartopo, D. Danang, D. Sasmoko, and N. D. Setiawan, "Rancang bangun sistem keamanan ruang persandian menggunakan RFID dan sensor PIR berbasis IOT," *Router J. Tek. Inform. dan Terap.*, vol. 2, no. 1, pp. 8–20, 2024.

[37] D. Danang, I. A. Dianta, A. B. Santoso, and S. Kholifah, "Hybrid CNN GRU Framework for Early Detection and Adaptive Mitigation of DDoS Attacks in SDN using Image Based Traffic Analysis," *Int. J. Inf. Eng. Sci.*, vol. 2, no. 2, pp. 66–78, 2025, doi: 10.62951/ijies.v2i2.292.

[38] D. Danang, M. U. Dewi, and G. Widhiati, "Federated Hybrid CNN GRU and COBCO Optimized Elman Neural Network for Real Time DDoS Detection in Cloud Edge Environments," *Int. J. Electr. Eng. Math. Comput. Sci.*, vol. 2, no. 2, pp. 28–35, 2025, doi: 10.62951/ijeemcs.v2i2.293.

[39] I. Englishtina, H. R. D. Putranti, D. Danang, and A. A. B. Pujiati, "SITENAR CERYA as an innovation in English language learning at SMP Stella Matutina Salatiga: Merging technology and folktales," *REKA ELKOMIKA J. Pengabdi. Kpd. Masy.*, vol. 5, no. 3, pp. 241–250, 2024.

[40] H. R. D. Putranti, D. Danang, T. Da Silva, and A. A. B. Pujiati, "Integrating Hands-on and Virtual Learning for Environmental Sustainability: Eco Enzyme Soap Making at Stella Matutina," *REKA ELKOMIKA J. Pengabdi. Kpd. Masy.*, vol. 6, no. 1, pp. 88–97, 2025.

[41] D. Danang, A. B. Santoso, and M. U. Dewi, "CICA Framework: Harnessing CSR, AI, and Blockchain for Sustainable Digital Culture," *Int. J. Adv. Comput. Sci. \& Appl.*, vol. 16, no. 11, 2025.

[42] D. Danang, H. Haryani, Q. Aini, F. A. Ramahdan, and J. Edwards, "Empowering digital literacy through blockchain based alphasign for secure and sustainable e-governance," 2025.