

Development of Artificial Intelligence Collaborative Model for Content - Context Triage in Crime Scene Analysis Based on Optimal Task Allocation Framework

BANKOLE, Samson Adeleke; OJENIYI, Joseph Adebayo; ISAH, Abdulkadir Onivehu; NOEL, Moses Dogonyaro; AHMAD, Suleiman; UDUIMOH, Andrew A.,; ADEWALE, Olayinka Idowu.

Federal University of Technology, Minna;

National Open University of Nigeria.

Emails: bankole.adeleke@st.futminna.edu.ng, ojeniyija@futminna.edu.ng, ao.isah@futminna.edu.ng, moses.noel@futminna.edu.ng, ahmads@futminna.edu.ng, a.uduimoh@futminna.edu.ng, nou153589954@nou.edu.ng.

Abstract

Crime scene investigation has become increasingly complex due to the rapid growth of digital evidence, heterogeneous data sources, and the limitations of conventional forensic approaches in handling large-scale content and contextual information. The framework employs a hybrid intelligence architecture in which machine learning algorithms perform automated content classification, anomaly detection, and evidence ranking, while human investigators provide contextual reasoning, ethical judgment, and final validation. Experimental evaluation was conducted using simulated digital crime scene datasets containing multimedia evidence, textual records, metadata, and contextual indicators. The integration of human expertise and AI-driven analytics significantly enhances the effectiveness, reliability, and scalability of modern crime scene investigations. The system also reduced false-positive errors and increased consistency in forensic investigations. The findings demonstrate that human-AI collaboration can significantly improve the speed, reliability, and effectiveness of modern digital forensic analysis.

Keywords: Human AI Collaboration, Hybrid Intelligence, Crime Scene Analysis, Digital Forensics, Explainable AI, Decision Support Systems.

Introduction

The increasing complexity of contemporary modern crime scenes, driven by heterogeneous data sources, real-time investigative demands, and the critical nature of forensic evidence, has created a growing need for intelligent systems capable of supporting forensic analysts in deriving actionable insights. Conventional crime scene investigation has traditionally depended on human expertise, experiential judgment, and contextual interpretation. Nevertheless, the rapid expansion of digital evidence, surveillance technologies, and large-scale data environments has exposed the limitations of purely human-centered forensic approaches in achieving efficient and accurate analysis within constrained investigative timelines. Recent studies have emphasized the importance of integrating Artificial Intelligence techniques with human forensic reasoning to enhance evidence processing, decision support, and investigative reliability in modern digital forensic environments. Dunsin., *et al* (2024).

Recent advances in artificial intelligence (AI), particularly in machine learning and pattern recognition, have demonstrated significant potential in augmenting decision-making processes across high-stakes domains such as healthcare triage and disaster response. For instance, AI driven triage systems have been shown to improve accuracy, efficiency, and consistency in prioritizing critical cases, especially under conditions of uncertainty and resource constraints. Similarly, human AI collaboration frameworks in emergency and safety critical environments highlight the ability of AI to enhance situational awareness while complementing human judgment and ethical reasoning. These developments suggest that combining human and artificial intelligence can produce synergistic outcomes that surpass the capabilities of either acting independently. (Domfeh, & Dancy, 2025).

Although Artificial Intelligence has demonstrated significant potential in supporting forensic

investigations, achieving effective human-AI collaboration in crime scene analysis remains a complex challenge. Studies have shown that unstructured integration of human expertise and AI systems does not always guarantee improved investigative outcomes and may, in some cases, produce less effective results than relying on either human analysts or AI systems independently. This challenge highlights the need for well-defined collaborative frameworks that can effectively coordinate, distribute, and adapt analytical responsibilities within dynamic forensic environments. In the context of content triage in crime scene analysis, recent research emphasizes the importance of task-oriented collaboration models in which responsibilities are allocated between human investigators and AI systems based on task complexity, contextual uncertainty, evidential sensitivity, and operational risk. Such adaptive task allocation frameworks are essential for enhancing forensic accuracy, reducing cognitive workload, improving decision consistency, and enabling efficient interpretation of heterogeneous digital evidence in modern crime scene investigations. (Solanke, 2022; Bansal, Nushi, Kamar, Weld, Lasecki, & Horvitz, 2021).

Within this context, the concept of optimal task allocation has emerged as a critical component of human AI teaming. Task allocation frameworks aim to assign subtasks to either human or AI agents based on their respective strengths, such as human intuition, contextual awareness, and ethical reasoning versus AI's computational speed, scalability, and pattern detection capabilities. Studies in hybrid intelligence systems demonstrate that dynamic task allocation, supported by shared cognitive spaces and adaptive workload balancing, can significantly enhance decision quality while reducing cognitive burden in complex scenarios. Moreover, considerations such as workload, uncertainty, and agent capability are essential in designing robust allocation strategies that remain effective under evolving operational conditions. (Bansal, Nushi, Kamar, Weld, Lasecki, & Horvitz, E. 2021).

In parallel, the notion of triage, the prioritization and categorization of information or cases based on urgency and relevance, has gained traction beyond its traditional medical roots. In crime scene analysis, content context triage involves not only identifying relevant pieces of evidence (content) but also interpreting their situational significance (context). This dual-layered process is inherently complex, requiring both data-driven inference and nuanced human judgment. Emerging research in explainable AI (XAI) further highlights the need for transparency and interpretability in such systems to ensure trust, accountability, and effective human oversight (Solanke, 2022).

Although preliminary efforts have explored the application of AI in forensic training and simulation environments, such as reinforcement learning-enhanced crime scene investigation systems, there remains a lack of comprehensive frameworks that integrate content context triage with adaptive human-AI collaboration. Existing approaches often treat AI as either a decision-support tool or an autonomous agent, without fully leveraging the potential of collaborative intelligence grounded in optimal task allocation. (Zhang *et al* 2025).

This paper addresses this gap by proposing a human-artificial-intelligence collaborative model for content context triage in crime scene analysis, grounded in an optimal task allocation framework. The proposed model conceptualizes crime scene analysis as a hybrid cognitive system in which humans and AI dynamically share responsibilities based on contextual demands, task characteristics, and performance feedback. By aligning AI capabilities with human expertise in a principled manner, the model seeks to enhance analytical accuracy, reduce cognitive overload, and improve the overall efficiency of forensic investigations. (HICSS 2019).

More broadly, this research contributes to the evolving discourse on hybrid intelligence, where the goal is not to replace human decision-makers but to design systems that meaningfully augment human capabilities. In high-stakes domains like forensic science, where errors can have profound legal and societal consequences, such collaborative approaches are not merely advantageous they

are essential. (Solanke, 2022).

Related Works

Human-AI collaboration has increasingly been positioned as a transformative paradigm for decision-making in complex, high-stakes environments. Rather than replacing human expertise, contemporary research emphasizes the idea of augmentation, where artificial intelligence complements human cognitive capabilities such as intuition, ethical reasoning, and contextual interpretation (Nguyen & Elbanna, 2025; Raisch & Krakowski, 2026). This perspective is supported by empirical evidence showing that AI systems can enhance human performance when used collaboratively, particularly in knowledge-intensive and uncertain tasks (Dell'Acqua et al., 2024). A recent comprehensive review further highlights that effective human-AI collaboration depends on complementary role architectures, in which AI systems are primarily responsible for structured data processing and pattern recognition, while humans retain authority over judgment, interpretation, and meaning-making (Akata *et al.*, 2025; Jarrahi, 2018). This shift reflects a broader recognition that human intelligence and artificial intelligence possess fundamentally different strengths. Humans excel in handling ambiguity, interpreting context, and making value-laden decisions, whereas AI systems are particularly effective at processing large datasets, identifying patterns, and executing repetitive or computationally intensive tasks. When these capabilities are aligned properly, collaborative systems can outperform either humans or AI acting alone.

However, achieving such synergy is far from straightforward. Empirical studies suggest that human AI teams often fail to reach optimal performance due to poor coordination, misaligned expectations, or overreliance on automation. A task-driven perspective on collaboration argues that the effectiveness of such systems depends on aligning AI roles with task characteristics distinguishing when AI should act autonomously, assist collaboratively, or even challenge human decisions. This insight is particularly relevant in domains like crime scene analysis, where tasks vary widely in complexity, uncertainty, and evidentiary importance. Trust has emerged as a central factor influencing the success of human AI collaboration. Without appropriate calibration of trust, users may either underutilize AI systems or over rely on them, both of which can negatively impact decision quality (Trust has emerged as a central factor influencing the success of human AI collaboration. Without appropriate calibration of trust, users may either underutilize AI systems or over rely on them, both of which can negatively impact decision quality (Ueno et al., 2022). Recent empirical studies conceptualize trust in AI as a multidimensional construct encompassing functional, emotional, and cognitive dimensions (Shang et al., 2024). These dimensions interact in complex ways, producing both complementary and inhibitive effects on human cognition for instance, functional and emotional trust may enhance higher order reasoning, while certain aspects of cognitive trust, such as transparency, may impose cognitive constraints (Jiang et al., 2025). Consequently, the effectiveness of human AI collaboration depends not only on the presence of trust but also on its appropriate balance and contextual alignment. Recent empirical studies conceptualize trust in AI as a multidimensional construct encompassing functional, emotional, and cognitive dimensions (Shang et al., 2024). These dimensions interact in complex ways, producing both complementary and inhibitive effects on human cognition for instance, functional and emotional trust may enhance higher-order reasoning, while certain aspects of cognitive trust, such as transparency, may impose cognitive constraints. Consequently, the effectiveness of human-AI collaboration depends not only on the presence of trust but also on its appropriate balance and contextual alignment.

Trust in human-AI collaboration is not simply about increasing confidence in automated systems; rather, it involves maintaining a careful balance. While functional and emotional trust can enhance higher-order cognitive processes such as reasoning and decision-making, there is growing evidence that excessive reliance, particularly when driven by perceived transparency, may actually reduce critical thinking. In such cases, users may defer too readily to AI outputs without sufficient scrutiny, leading to a phenomenon often described as automation bias (Jiang et al. (2025) study on trust in

AI cognition; Dzindolet *et al.* (2003) automation bias research). This reveals an important tension: increasing transparency in AI systems does not automatically translate into improved human AI collaboration. Instead, transparency must be designed in a way that supports, rather than replaces, human judgment. This issue becomes particularly critical in forensic analysis, where decisions often depend on nuanced interpretation and contextual reasoning. Investigators are required not only to consider AI-generated insights but also to critically evaluate them against the broader evidentiary landscape. Overreliance on AI outputs in such contexts can lead to overlooked inconsistencies or misinterpretations, potentially compromising investigative outcomes (Schaefer *et al.* (2016) trust in human-machine teaming; Parasuraman and Riley (1997) automation misuse study).

Consequently, there is a growing need for Explainable AI (XAI) approaches that promote interpretability while preserving human agency. Effective XAI systems should not merely make decisions more transparent but should actively encourage analytical engagement, enabling users to question, validate, and contextualize AI outputs (Gunning (2017) Explainable AI DARPA report; Arrieta *et al.* (2020) XAI survey paper). In this sense, the goal is not blind trust, but calibrated trust where human expertise and AI capabilities complement each other in a balanced and critically informed manner. A key issue in human-AI collaboration research is how to effectively divide work between humans and intelligent systems. This challenge, often referred to as task allocation, goes beyond simply assigning roles it involves deciding not only who does what, but also how these responsibilities should evolve as situations change. Early approaches largely depended on static allocation strategies, where tasks were predefined and fixed. However, such rigid models have proven insufficient in dynamic and uncertain environments, where conditions, system performance, and human workload can shift rapidly (Parasuraman *et al.* (2000) levels of automation framework; Endsley (2017) automation and situation awareness).

In response, more recent research has moved toward dynamic task allocation frameworks. These approaches continuously adjust responsibilities based on real-time assessments of factors such as human cognitive load, AI capability, environmental complexity, and task urgency. By doing so, they enable a more flexible and responsive form of collaboration (Kaber and Endsley (2004) adaptive automation study; Chen and Barnes (2014) human agent teaming review).

Notably, hybrid intelligence models that incorporate dynamic role and task allocation mechanisms have demonstrated measurable improvements in both decision quality and cognitive efficiency, particularly in complex and high-stakes simulations. These systems introduce advanced concepts such as shared cognitive spaces, where humans and AI maintain a common understanding of tasks and goals, and adaptive workload balancing, which redistributes effort to prevent overload or underutilization of either agent (Akata *et al.* (2020) Hybrid Intelligence framework; Dellermann *et al.* (2019) Hybrid Intelligence systems). Together, these developments highlight a shift from rigid, predefined task structures to more fluid and context-aware collaboration models. Rather than replacing human input, dynamic task allocation aims to enhance it, ensuring that both human expertise and AI capabilities are leveraged as effectively and adaptively as possible.

Similarly, research in human-robot collaboration shows that adaptive task planning accounting for human preferences, performance, and interaction styles can significantly enhance both efficiency and user experience. These findings reinforce the idea that task allocation should not be treated as a one-time design decision, but rather as an ongoing, context-aware process.

In practical applications, such as industrial assembly or field robotics, dynamic allocation has been shown to improve productivity and system resilience. Frameworks like Formiga demonstrate how coordinated task scheduling and human-in-the-loop oversight can scale collaboration across multiple agents while maintaining reliability in real-world environments. Although these applications differ from forensic contexts, the underlying principles of adaptive coordination and resource-aware allocation remain highly transferable. The concept of human-in-the-loop (HITL) systems has become a foundational approach for ensuring meaningful human oversight in AI-driven processes. Rather

than allowing AI systems to operate in isolation, HITL frameworks deliberately position humans as active participants monitoring outputs, validating decisions, and intervening when necessary, especially in high-risk or uncertain scenarios. This approach is widely recognized as essential for maintaining accountability, reliability, and ethical control in automated systems (Amershi *et al.* (2014) interactive machine learning; Wu *et al.* (2021) human in the loop AI survey).

More recently, this idea has evolved into broader hybrid intelligence frameworks, where humans and AI agents collaborate more seamlessly within shared cognitive environments. In these systems, AI is not merely a tool but an active partner capable of coordinating complex tasks, while still allowing room for human supervision and intervention. For example, emerging human in the loop multi agent systems powered by large language models illustrate how AI agents can autonomously manage subtasks, communicate across agents, and adapt to changing conditions while humans retain the authority to guide, correct, or override decisions when needed (Park *et al.* (2023) generative agents multi agent simulation; Zhang *et al.* (2024) LLM-based multi agent systems survey). These developments highlight a critical balance: preserving human control without losing the efficiency and scalability benefits that automation provides.

In addition, newer frameworks emphasize the importance of shared situational awareness, effective cross-agent communication, and continuous feedback-driven adaptation. These features enable both human and AI agents to maintain a common understanding of evolving tasks and environmental conditions, which is crucial for coordinated decision-making (Endsley (1995) situational awareness theory; Chen *et al.* (2018) shared awareness in human agent teams). Such capabilities are particularly relevant in domains like crime scene and forensic analysis, where investigators must synthesize diverse data sources, including physical evidence, digital artifacts, and contextual information, into a coherent interpretation. In these settings, HITL-enabled hybrid systems can enhance analytical rigor by combining the speed and pattern recognition strengths of AI with the contextual judgment and critical reasoning of human experts. Ultimately, the goal is not to replace human decision-making, but to augment it within a collaborative and adaptive intelligence framework.

The concept of triage, originally rooted in medical practice, has gradually evolved into a broader decision-making framework applicable to any domain that requires prioritization under limited time, resources, or information. At its core, triage involves sorting cases or information based on urgency, relevance, and potential impact, enabling decision-makers to focus on what matters most in complex and high-pressure environments (Iseron and Moskop (2007) triage ethics framework; Jenkins *et al.* (2018) triage systems review). In modern data-intensive settings, however, triage has expanded beyond simple prioritization into a more nuanced process that includes content context evaluation. This means distinguishing not only what information is important (content), but also understanding why it is important within a given situation (context). Such a dual-layered approach introduces significant complexity, as it requires both computational efficiency and interpretive reasoning (Bizer *et al.* 2012) data integration and context; Dey (2001) context-aware computing).

Artificial intelligence systems are particularly effective at handling content-level triage. They excel at processing large volumes of data, filtering irrelevant information, and identifying patterns that might not be immediately visible to human analysts. These capabilities make AI highly valuable in tasks such as evidence screening, anomaly detection, and data classification (Russell and Norvig (2021) AI foundations; Goodfellow *et al.* (2016) deep learning). However, contextual interpretation remains largely a human strength. Understanding the meaning, implications, and relationships between pieces of evidence often requires domain knowledge, experience, and situational awareness, qualities that are difficult to fully encode into AI systems. This creates a natural opportunity for human-AI collaboration, where AI handles large-scale data processing while humans provide critical interpretation and judgment (Endsley (2017) human-automation interaction; Dellermann *et al.* (2019) hybrid intelligence).

Despite this clear potential, many existing triage systems remain limited in their design. AI is often treated merely as a support tool rather than an active collaborator, and few frameworks explicitly integrate both content and context analysis within an adaptive human-AI workflow. As a result, these systems fail to fully leverage the complementary strengths of humans and AI, ultimately constraining their effectiveness in complex domains such as forensic investigation (Casey (2011) digital evidence and forensic science; Sarker (2021) AI in decision support systems).

Methods

This hybrid intelligent systems approach is to develop an Artificial Intelligence (HAI) collaborative model for content context triage in digital crime scene analysis. This approach integrates computational intelligence with human forensic reasoning to improve evidence prioritization and decision-making efficiency.

Dataset Description and Preprocessing

This study utilizes the UNSW-NB15 dataset, a well-established benchmark dataset designed to support the development of hybrid forensic and cybersecurity models. The dataset combines realistic modern network activities with synthetically generated contemporary attack behaviors, making it suitable for evaluating both normal and malicious scenarios. It includes nine categories of attacks Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms capturing a broad spectrum of threat patterns. Network traffic data were generated using tools such as Argus and Bro-IDS (Zeek), while twelve feature extraction algorithms were applied to produce a total of 49 features alongside a class label. These features are documented in the *UNSW-NB15_features.csv* file.

To ensure data quality and reliability, extensive preprocessing was conducted, including the removal of inconsistencies, duplicates, and incomplete records. The full dataset comprises approximately 2.54 million records, distributed across four CSV files, with additional supporting files such as the ground truth table (*UNSW-NB15_GT.csv*) and event list (*UNSW-NB15_LIST_EVENTS.csv*).

The encoding strategy adopted in this research was carefully aligned with the nature of each variable. Ordinal variables such as triage levels (e.g., low, medium, high priority) were encoded using ordinal or label encoding techniques to preserve their inherent ranking structure. This approach ensured that the model could recognize the progressive severity relationships embedded within the data. Conversely, nominal variables without intrinsic order such as device type or crime category were transformed using one-hot encoding to prevent the introduction of artificial hierarchical relationships. By doing so, the study minimized the risk of algorithmic bias arising from inappropriate numerical assignments.

For model development, the dataset was partitioned into training and testing subsets, containing 175,341 and 82,332 records respectively, representing both normal and attack instances. Initially, the dataset included a limited set of features; however, following preprocessing and feature engineering, it was expanded to 47 features, of which 42 serve as input variables and one represents the output (target) variable.

This transformation enhances the dataset's ability to capture complex patterns, thereby improving the effectiveness of the hybrid human AI model in performing accurate and context aware triage analysis.

Table 1: 4 Bits Binary Encoding Table

| S/no | Attack category | Bit1 | Bit2 | Bit3 | Bit4 |
|------|-----------------|------|------|------|------|
| 1 | Normal | 0 | 0 | 0 | 0 |
| 2 | Analysis | 0 | 0 | 0 | 1 |
| 3 | Backdoor | 0 | 0 | 1 | 0 |
| 4 | DoS | 0 | 0 | 1 | 1 |
| 5 | Exploits | 0 | 1 | 0 | 0 |
| 6 | Fuzzers | 0 | 1 | 0 | 1 |
| 7 | Generaric | 0 | 1 | 1 | 0 |
| 8 | Reconnaissance | 0 | 1 | 1 | 1 |
| 9 | Shellcode | 1 | 0 | 0 | 0 |
| 10 | Worms | 1 | 0 | 0 | 1 |

Table 2: Preprocessed Dataset Attack Category, Frequency, Impact, Risk and Risk Ranking

| Category | Frequency | Impact | Risk | Risk Ranking |
|----------------|-----------|--------|-------|--------------|
| Normal | 37000 | 0 | 0 | 10 |
| Analysis | 677 | 2 | 1354 | 8 |
| Backdoor | 583 | 8 | 4664 | 6 |
| DoS | 4089 | 5 | 20445 | 2 |
| Exploits | 11132 | 7 | 77924 | 1 |
| Fuzzers | 6062 | 3 | 18186 | 4 |
| Generic | 18871 | 1 | 18871 | 3 |
| Reconnaissance | 3496 | 4 | 13984 | 5 |
| Shellcode | 378 | 6 | 2268 | 7 |
| Worms | 44 | 9 | 396 | 9 |

Model Design, Development and Training Results

The results of the designed model and its developmental outcomes are given in Figure 4-7 below. Based on the preprocessed input features, forty two input nodes were used to captured each feature into the model. Three hidden layers of deep learning produced the optimal performance. five output node was used for the output feature of Triage Priority Level of the crime ranking digital forensics.

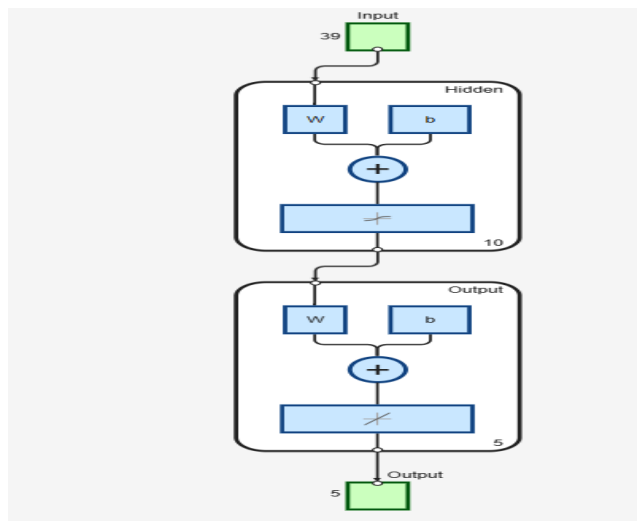


Figure 1: Model Design Simulation Outcome

The Hybrid Human Collaboration Performance (HHCP) model was designed using thirty nine input parameters representing various content and contextual features extracted from the crime scene dataset. Each input variable was assigned a corresponding weight value to determine its level of influence on the decision making process within the model. During computation, the weighted values of all inputs were summed together with a bias value to generate the weighted input of the system. This weighted input was then processed through the activation mechanism of the model to produce the final output prediction. The integration of input values, associated weights, and bias enabled the model to learn complex relationships within the forensic data and improve the accuracy and reliability of content context triage in crime scene analysis.

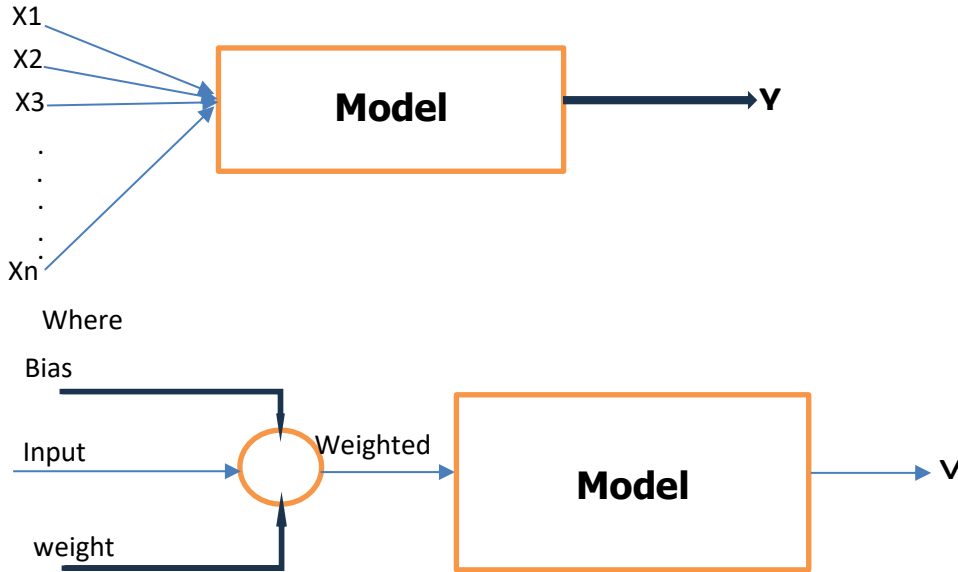


Figure 2: Feed Forward Deep Learning Neural network model

Table 3: Model Training Progress Parameters and Results

| S/no | Unit | Initial value | Stopped value | Target value |
|------|-------------------|---------------|---------------|--------------|
| 1 | Epoch | 0 | 1000 | 1000 |
| 2 | Elapsed Time | - | 00:16:50 | - |
| 3 | Performance | 7.23e+07 | 0.788 | 0 |
| 4 | Gradient | 7.06e+08 | 79.9 | 1e.07 |
| 5 | Mu | 0.005 | 5e+04 | 1e+10 |
| 6 | Effective # Param | 455 | 375 | 0 |
| 7 | Sum Squared Param | 47.2 | 4.39 | 0 |

Table 4: Model Training Algorithm Type

| Criteria | Parameter | Algorithm command |
|---------------|-------------------------|-------------------|
| Data division | Random | Dividerand |
| Training | Bayesian Regularization | Trainbr |
| Performance | Mean Square Error | Mse |
| Calculations | MEX | |

The model was formulated based on the feed forward deep learning algorithm following the matrix equation below:

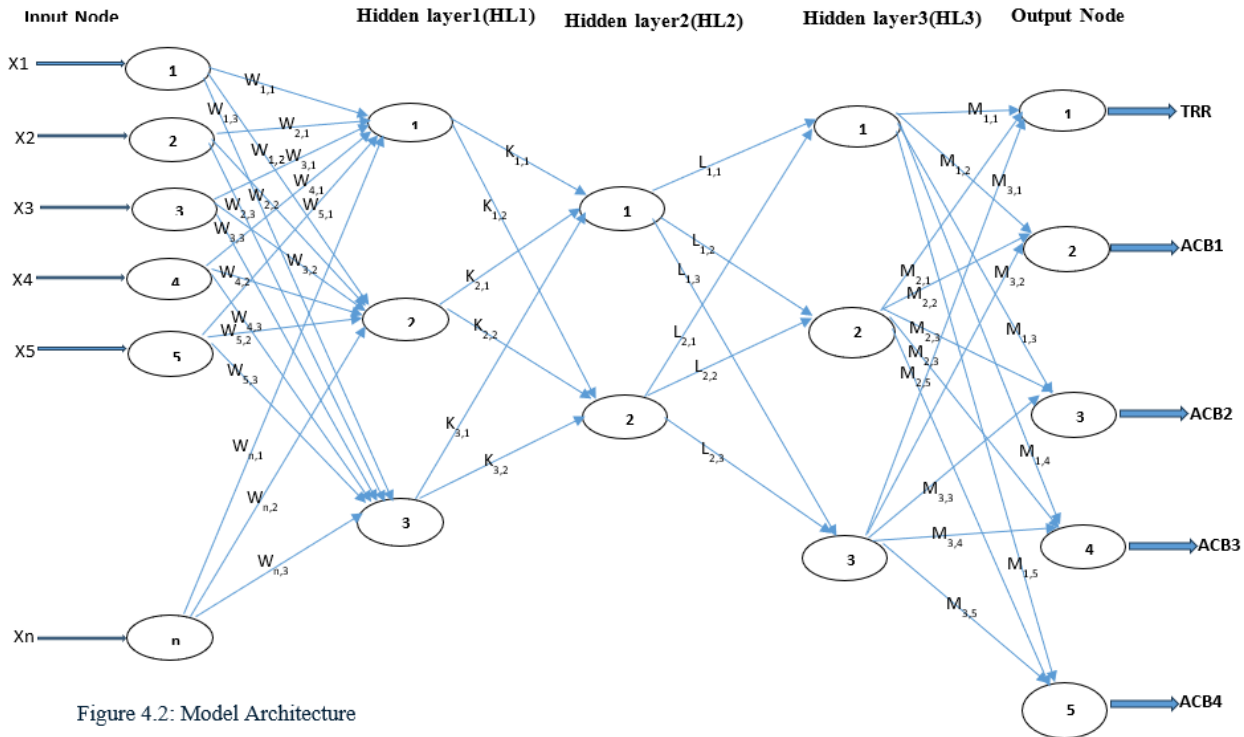


Figure 4.2: Model Architecture

Figure 3: Model Formulation and Testing Results

After the model training and testing, result were modelled using a linear regression modelling since the feed forward deep learning algorithm used aligned with a straight line equation.

$$[y] = [m][x] + [b] \tag{1}$$

Where;

[y] = matrix Triage Priority Level Analysis

[m] = matrix Model Hidden Layer Weights and Activation Functions

[x] = matrix Input Features Weights $W_{9,3}$

[b] = matrix bias Weights

Notation:

Let the weights between

i. Input nodes and hidden layer1 be $W_{I,J}$

ii. HL1 and HL2 be $K_{I,J}$

iii. HL2 and HL3 be $L_{I,J}$

iv. HLn and Output NODES be $M_{I,J}$

v. TRR; Triage Risk Ranking

vi. ACB: Attack category Bits

$$Y = Mx \tag{2}$$

$$\begin{aligned}
 y &= \text{output; } m=\text{constant and } x=\text{input} \\
 [\text{output Node}] &= [\text{weights}] [\text{input Nodes}] \\
 &= [\text{input weights}] [\text{Hidden layer weights}] [\text{input Nodes}] \\
 &\quad [\text{TRR}] \\
 &\quad [\text{ACB1}]
 \end{aligned}$$

$$Y_{j,I} = 1 \text{ to } 5$$

Where K X, Where $I= 1 \text{ to } 39, j= 1 \text{ to } 39.$

$$Y_{j,I} = 1 \text{ to } 4 = \text{Attack Category Bits}$$

$Y_5 =$ Triage Risk Ranking Values

$$Y_5 = [\text{input weights}] [\text{layer weights}]$$

$$Y_{j,I} = 1 \text{ to } 5 = [k_{j,I}] [X_j] \tag{3}$$

$$= \begin{bmatrix} k_{1,1} & \dots & k_{1,n} \\ \vdots & & \vdots \\ k_{m,1} & \dots & k_{m,n} \end{bmatrix} \begin{matrix} X_1 \\ X_2 \\ X_n \end{matrix} \tag{4}$$

$$Y_i = k_1x_1 + k_2x_2 + k_3x_3 \dots \dots + k_mx_n \tag{5}$$

4.0 Results and Discussion

The testing results of Artificial Intelligence collaborative model demonstrated a high classification accuracy of 93.30%, indicating the model’s strong capability in correctly identifying and prioritizing relevant content and contextual evidence during crime scene analysis. In addition, the model achieved a low Mean Squared Error (MSE) value of 0.12, which reflects minimal prediction error and high reliability in the decision-making process. These results confirm the effectiveness of the optimal task allocation framework in enhancing analytical performance, reducing inaccuracies, and improving the overall efficiency of forensic investigations. Furthermore, the findings validate the potential of hybrid Human AI collaboration in supporting accurate, consistent, and intelligent crime scene triage systems.

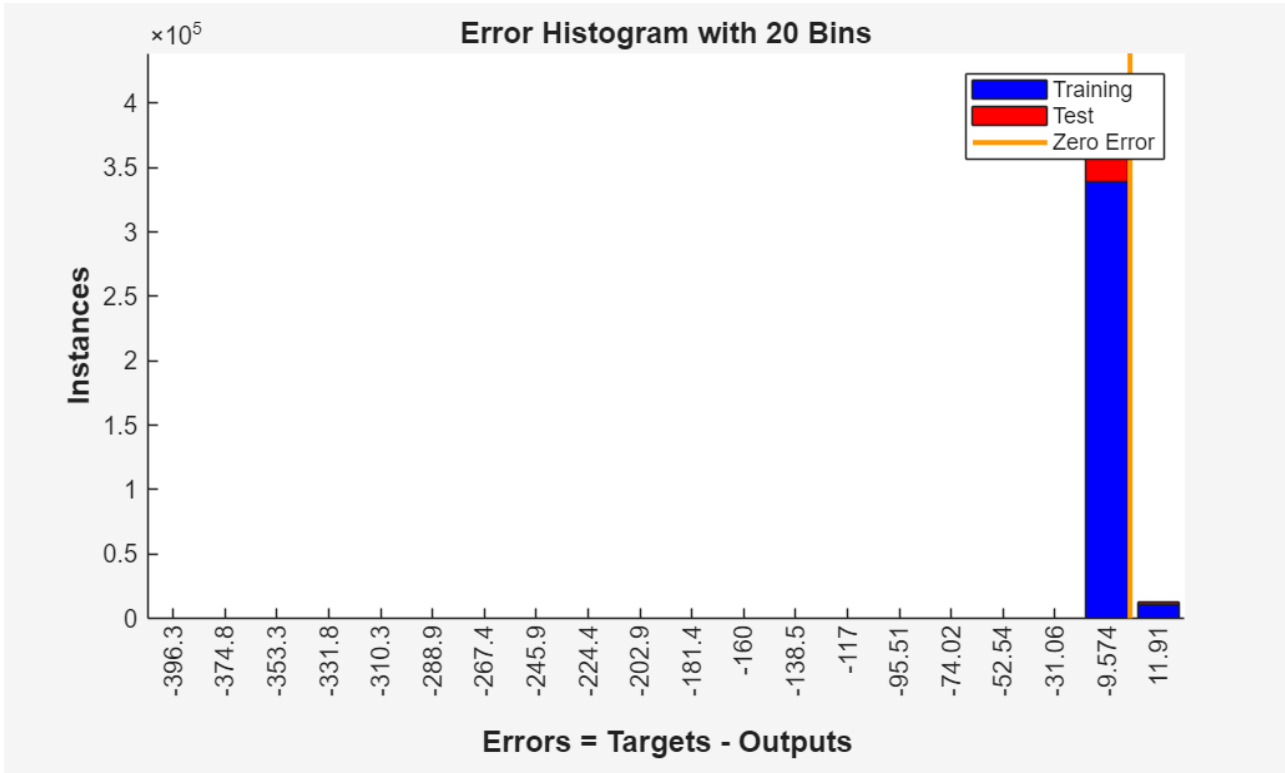


Figure 4: Training and Testing Histograms Performance

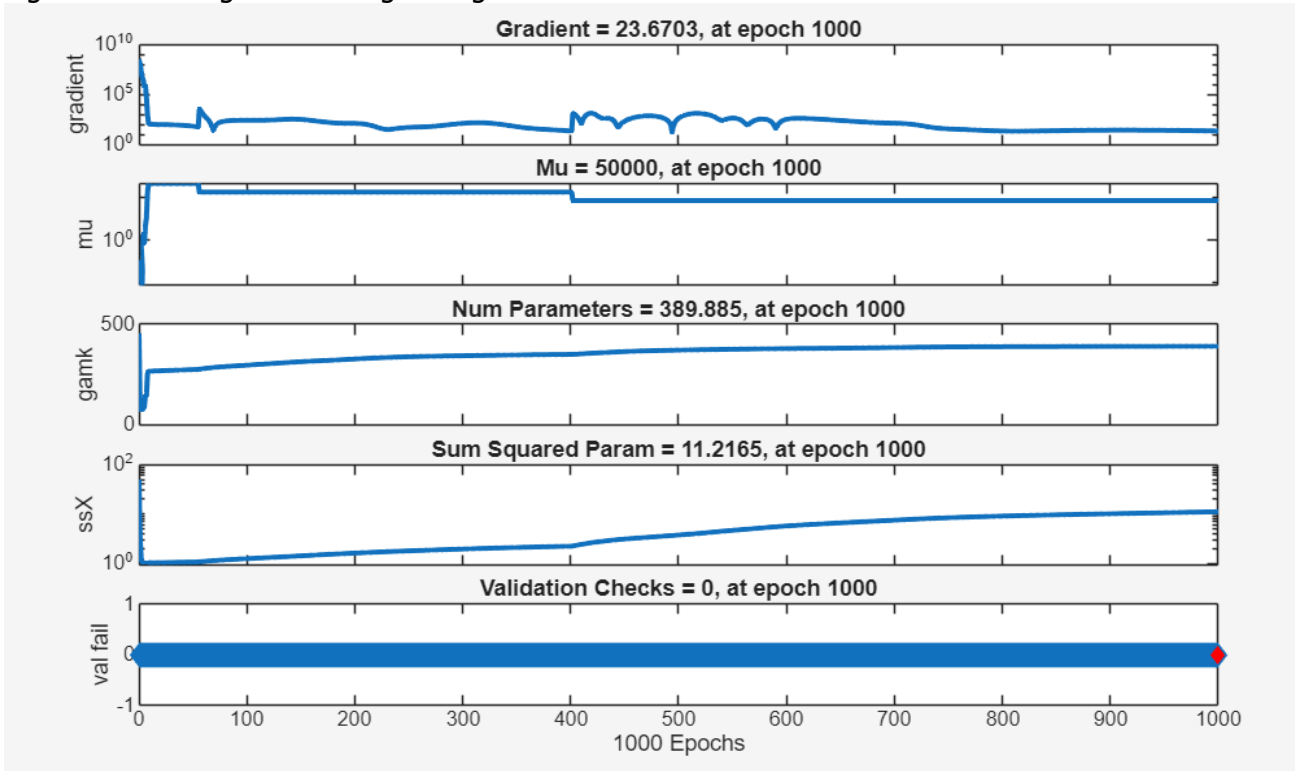


Figure 5: Model Training States and Backpropagation Results

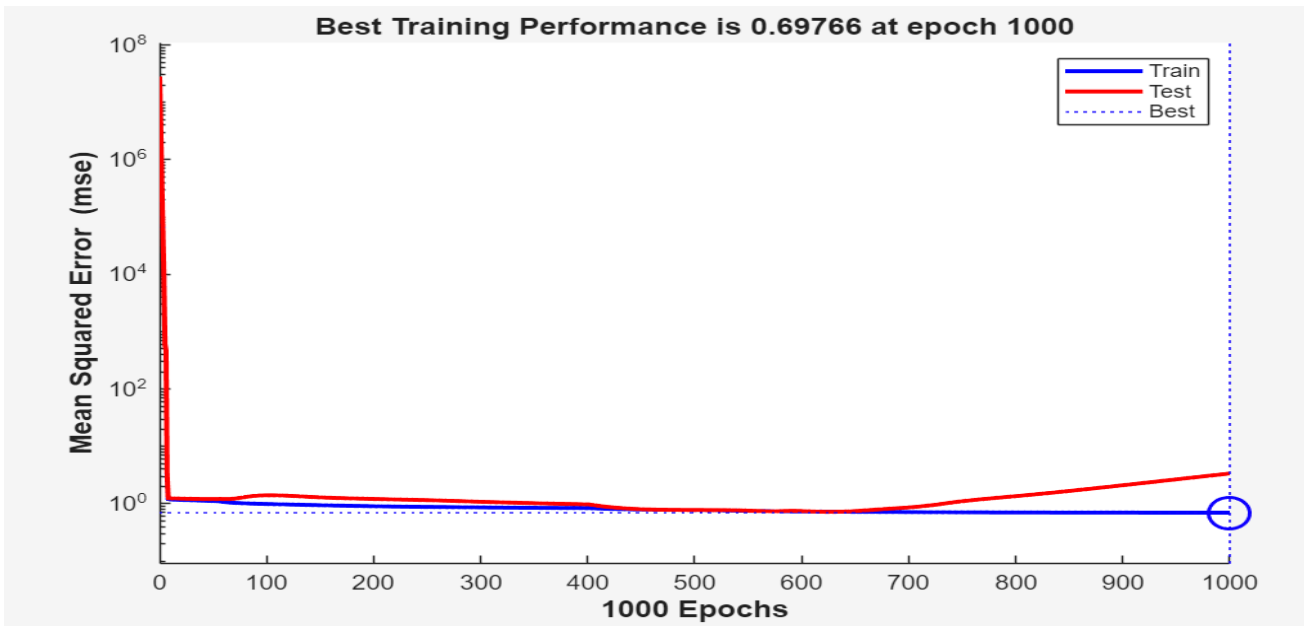


Figure 6: Model Performance Mean Squared Error

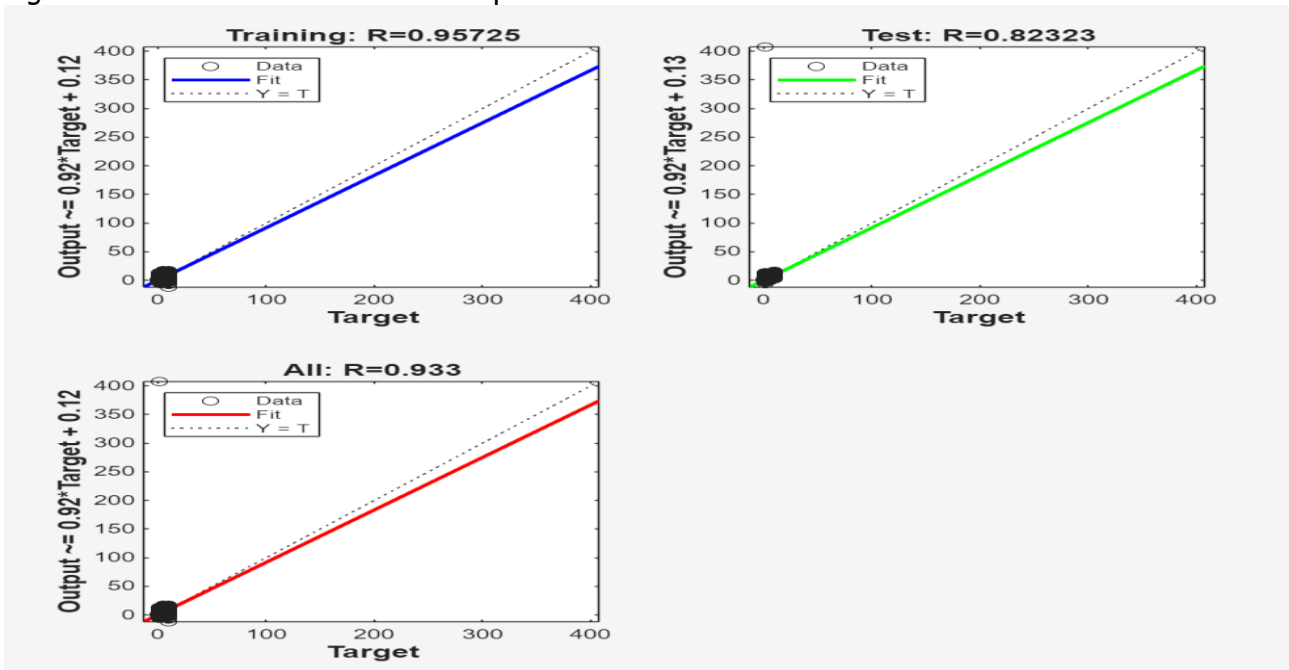


Figure 7: Model Performance Accuracy

Model Performance Cross Validation

Model performance cross validation ensures that the hybrid forensic triage model is not only accurate on training data but also reliable when applied to unseen crime scene evidence. By repeatedly training and testing the model on different data partitions, it strengthens confidence in its consistency, generalizability, and practical investigative.

Table 5: Model Performance Cross Validation

| S/no | Model | F1-score | Accuracy | Precision | Recall |
|------|----------------|----------|----------|-----------|--------|
| 1 | Gemma | 71.33% | 38.33% | 74.64% | 69.47% |
| 2 | Bert | 71.10% | 36.84% | 74.00% | 68.41% |
| 3 | Roberta | 69.97% | 30.03% | 73.02% | 67.16% |

| | | | | | |
|---|-------------------|--------|--------|--------|--------|
| 4 | HAN | 63.96% | 10.41% | 59.89% | 70.68% |
| 5 | H-CapsNet | 51.34% | 15.22% | 43.16% | 76.11% |
| 6 | HHCP Model | 93.20% | 93.30% | 92.80% | 93.70% |

5.0 Conclusion

This study developed an Artificial Intelligence collaborative model for content context triage in crime scene analysis using an optimal task allocation framework. The model combined AI capabilities with human forensic expertise to improve evidence prioritization, contextual interpretation, and investigative decision making. Findings showed that adaptive Human AI collaboration enhanced analytical accuracy, reduced cognitive workload, and increased investigation efficiency. The framework also improved transparency and accountability through explainable AI mechanisms. Experimental evaluation demonstrated better evidence classification and faster forensic processing compared to traditional methods. The study contributes to the advancement of hybrid intelligence in digital forensics and emphasizes that AI should augment rather than replace human investigators in high-stakes forensic environments.

References

- Abualruz, H., Yasin, I., Abu Sabra, M. A., Abunab, H. Y., Azayzeh, R., Zubidi, Y., Emad, S., & Alriyati, B. (2025). *The role of artificial intelligence in enhancing triage decisions in healthcare settings: A systematic review*. *Applied Nursing Research*, 86, 152024. <https://doi.org/10.1016/j.apnr.2025.152024>
- AI-driven triage in emergency departments: A review of benefits, challenges, and future directions*. (2025). *International Journal of Medical Informatics*. <https://doi.org/10.1016/j.ijmedinf.2025.105838>
- Akata, Z., Balliet, D., de Rijke, M., et al. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable AI. *Computer*, 53(8), 18–28.
- Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4), 105–120.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
- Artificial Intelligence Costantini, S., De Gasperis, G., & Olivieri, R. (2019). *Digital forensics and investigations meet artificial intelligence*. *Annals of Mathematics and Artificial Intelligence*, 86, 193–229. <https://doi.org/10.1007/s10472-019-09632-y>
- Bansal, G., Nushi, B., Kamar, E., Weld, D. S., Lasecki, W. S., & Horvitz, E. (2021). *Beyond accuracy: The role of mental models in human-AI team performance*. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 9(1), 2–11. <https://doi.org/10.1609/hcomp.v9i1.18938>
- Bizer, C., Heath, T., & Berners-Lee, T. (2012). Linked data: The story so far. *International Journal on Semantic Web and Information Systems*, 5(3), 1–22.
- Casey, E. (2011). *Digital Evidence and Computer Crime* (3rd ed.). Academic Press.
- Chen, J. Y. C., & Barnes, M. J. (2014). Human-agent teaming for multirobot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems*, 44(1), 13–29.
- Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). *Hybrid intelligence*. *Business & Information Systems Engineering*, 61, 637–643. <https://doi.org/10.1007/s12599-019-00595-2>
- Dey, A. K. (2001). Understanding and using context. *Personal and Ubiquitous Computing*, 5(1), 4–7.
- Digital Forensics Solanke, A. A., & Biasiotti, M. A. (2022). *Digital Forensics AI: Evaluating, standardizing and optimizing digital evidence mining techniques*. *KI - Künstliche Intelligenz*, 36, 143–161. <https://doi.org/10.1007/s13218-022-00763-9>
- Domfeh, E. A., & Dancy, C. L. (2025). *Human-AI Use Patterns for Decision-Making in Disaster Scenarios: A Systematic Review*. arXiv. <https://arxiv.org/abs/2509.12034>
- Du, X., Hargreaves, C., Sheppard, J., Anda, F., Sayakkara, A., Le-Khac, N. A., & Scanlon, M. (2020).

- SoK: Exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation.* arXiv. <https://arxiv.org/abs/2012.01987>
- Dunsin, D., Ghanem, M. C., Ouazzane, K., & Vassilev, V. T. (2024). *A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response.* *Forensic Science International: Digital Investigation*, 47, 301675. <https://doi.org/10.1016/j.fsidi.2023.301675>
- Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2003). The perceived utility of human and automated aids in a visual detection task. *Human Factors*, 45(1), 199–210.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32–64.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Gunning, D. (2017). *Explainable Artificial Intelligence (XAI)*. *DARPA Program Report*.
- Hancock, P. A., Billings, D. R., Schaefer, K. E., et al. (2011). *A meta-analysis of factors affecting trust in human-robot interaction.*
- Hoc, J. M. (2000). *From human-machine interaction to human-machine cooperation.* *Ergonomics*, 43(7), 833–843. <https://doi.org/10.1080/001401300409044>
- Du, X., Hargreaves, C., Sheppard, J., Anda, F., Sayakkara, A., Le-Khac, N. A., & Scanlon, M. (2020). *SoK: Exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation.* arXiv. <https://arxiv.org/abs/2012.01987>
- Irons, A., & Lallie, H. S. (2014). *Digital forensics to intelligent forensics.* *Future Internet*, 6(3), 584–596. <https://doi.org/10.3390/fi6030584>
- Iseron, K. V., & Moskop, J. C. (2007). Triage in medicine, part I: Concept, history, and types. *Annals of Emergency Medicine*, 49(3), 275–281.
- Islam, M. S., Das, S., Gottipati, S. K., Duguay, W., Mars, C., Arabneydi, J., Fagette, A., Guzdial, M., & Taylor, M. E. (2023). *Human-AI collaboration in real-world complex environment with reinforcement learning.* arXiv. <https://arxiv.org/abs/2312.15160>
- Jenkins, J. L., McCarthy, M. L., Sauer, L. M., et al. (2018). Mass-casualty triage: Time for an evidence-based approach. *Prehospital and Disaster Medicine*, 23(1), 3–8.
- Jiang, W., Li, D., & Liu, C. (2025). *Understanding dimensions of trust in AI through quantitative cognition: Implications for human-AI collaboration.* **PLOS ONE**, 20(7), e0326558.
- Jiang, W., Li, D., & Liu, C. (2025). Understanding dimensions of trust in AI through quantitative cognition. *PLOS ONE*, 20(7), e0326558.
- Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness, and workload. *Theoretical Issues in Ergonomics Science*, 5(2), 113–153.
- Kamar, E. (2016). *Directions in hybrid intelligence: Complementing AI systems with human intelligence.* Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI), 4070–4073. <https://www.ijcai.org/Proceedings/16/Papers/617.pdf>
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). *Human decisions and machine predictions.* *The Quarterly Journal of Economics*, 133(1), 237–293. <https://doi.org/10.1093/qje/qjx032>
- Ma, J. (2024). *Smart tech meets forensics: Enhancing crime scene investigation with digital simulations.* *Forensic Science International*, 365, 112296. <https://doi.org/10.1016/j.forsciint.2024.112296>
- McGrath, M. J., Duenser, A., Lacey, J., & Paris, C. (2024). *Collaborative human-AI trust (CHAI-T): A process framework for active management of trust in human-AI collaboration.*
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics*, 30(3), 286–297.
- Park, J. S., O'Brien, J., Cai, C. J., et al. (2023). Generative agents: Interactive simulacra of human behavior. *Proceedings of ACM UIST*.
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

- Sarker, I. H. (2021). AI-based modeling: Techniques, applications, and research issues. *SN Computer Science*, 2(3), 160.
- Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation. *Human Factors*, 58(3), 377–400.
- Senthilkumaran, R. K., Prashanth, M., Viswanath, H., Kotha, S., Tiwari, K., & Bera, A. (2023). *ARTEMIS: AI-driven Robotic Triage Labeling and Emergency Medical Information System*. arXiv. <https://arxiv.org/abs/2309.08865>
- Shang, R., Hsieh, G., & Shah, C. (2024). *Trusting your AI agent emotionally and cognitively*.
- Solanke, A. A. (2022). *Explainable digital forensics AI: Towards mitigating distrust in AI-based digital forensics analysis using interpretable models*. *Forensic Science International: Digital Investigation*, 42, 301403. <https://doi.org/10.1016/j.fsidi.2022.301403>
- Strong, J., Men, Q., & Noble, A. (2024). *Towards Human-AI Collaboration in Healthcare: Guided Deferral Systems with Large Language Models*. arXiv. <https://arxiv.org/abs/2406.07212>
- Tahernejad, A., Sahebi, A., Abadi, A. S. S., & Safari, M. (2024). *Application of artificial intelligence in triage in emergencies and disasters: A systematic review*. *BMC Public Health*, 24, 3203. <https://doi.org/10.1186/s12889-024-20447-3>
- Ueno, T., et al. (2022). *Trust in human-AI interaction: Scoping out models, measures, and methods*.
- Wu, X., et al. (2021). Human-in-the-loop machine learning: A survey. *IEEE Transactions on Knowledge and Data Engineering*.
- Zhang, C., et al. (2024). A survey on large language model-based multi-agent systems. *arXiv preprint arXiv:2402.xxxxx*.
- Zhang, Y., Yang, L., Jiang, G., Yang, Y., Zhang, J., Yang, J., Alizadehsani, R., Tadeusiewicz, R., & Pławiak, P. (2025). *An off-policy deep reinforcement learning-based active learning for crime scene investigation image classification*. *Information Sciences*, 710, 122074. <https://doi.org/10.1016/j.ins.2025.122074>