*Pinnacle Academic Press Proceedings Series*

*Vol. 2 2025*

*Article* Open Access

# DeepTriage: A Real-Time AI Decision Support System for Emergency Resource Allocation in Mass Casualty Incidents

**Zejun Cheng [1,*]**

[1] Clinical Medicine, Capital Medical University, Beijing, China
* Correspondence: Zejun Cheng, Clinical Medicine, Capital Medical University, Beijing, China


Received: 11 April 2025
Revised: 24 April 2025
Accepted: 09 May 2025
Published: 18 June 2025



**Abstract:** Mass casualty incidents (MCIs) present significant challenges to emergency medical systems, frequently overwhelming available resources and necessitating complex triage decisions under severe time constraints. This paper introduces DeepTriage, a real-time artificial intelligence decision support system designed to optimize emergency resource allocation during MCIs. The system employs a hybrid neural network architecture that integrates convolutional, recurrent, and graph neural networks to process multimodal patient data and generate prioritization recommendations. The DeepTriage framework incorporates privacy-preserving mechanisms, dynamic resource optimization algorithms, and an adaptive transmission strategy for deployment in bandwidth-constrained environments. Performance evaluation conducted across three diverse datasets — TRAUMA-DB, MCI-SIM, and DISASTER-NET — demonstrates superior triage accuracy (92.6%) compared to traditional protocols (76.4-79.2%) and existing computational systems (84.5-87.3%). The system achieves significant improvements in decision speed (14.8s under benchmarked test conditions versus 187.3-245.8s for manual methods in similar scenarios) while maintaining a resource utilization efficiency of 0.87. DeepTriage exhibits robust performance across multiple incident types with minimal degradation under increasing patient loads. Implementation considerations address integration pathways with existing electronic health record systems, training requirements for medical personnel, and ethical frameworks governing algorithmic decision support in life-critical scenarios. The results indicate substantial potential for AI-driven systems to enhance emergency response capabilities during mass casualty incidents through improved triage accuracy, resource optimization, and decision consistency.

**Keywords:** artificial intelligence; emergency medicine; mass casualty incidents; resource allocation

## 1. Introduction

### *1.1. Background and Significance of Mass Casualty Incident Management*

Mass casualty incidents (MCIs) represent extraordinary challenges to healthcare systems globally, characterized by overwhelming patient volumes that exceed immediately available medical resources. The management of these events demands sophisticated approaches to resource allocation and patient prioritization. Recent advancements in artificial intelligence (AI) technologies have demonstrated significant potential in enhancing decision-making processes across various domains. Fan et al. highlighted the efficacy of deep learning-based systems for anomaly detection and timely risk alert generation in critical environments [1]. These technological capabilities can be adapted to emergency medicine contexts, where rapid assessment and resource distribution are critical for patient outcomes. The application of machine learning algorithms for pattern recognition, as investigated by Bi et al., offers promising solutions for identifying critical patients and optimizing resource utilization during MCIs [2]. Modern emergency departments increasingly recognize the value of low-latency anomaly detection architectures, which Zhang et al. demonstrated can provide real-time decision support in complex, data-intensive environments [3]. The integration of temporal data processing through graph neural networks, explored by Wang et al., presents opportunities to model the dynamic nature of patient conditions and resource availability during evolving disaster scenarios [4].

### *1.2. Challenges in Emergency Resource Allocation During Disasters*

Emergency resource allocation during mass casualty incidents encompasses multiple interdependent challenges that complicate effective response. The unpredictable nature of patient influx resembles the complexity of anomalous flows studied by Su et al., who emphasized the importance of analyzing irregularities to maintain system integrity [5]. Traditional triage systems often rely on manual assessment procedures that introduce inconsistencies and delays in critical decision-making processes. The development of standardized evaluation metrics for automated systems, an approach explored by Liang et al. in cross-lingual contexts, represents a crucial step toward establishing reliable AI-assisted triage protocols [6]. Resource allocation decisions in emergency settings require transparent rationales to gain clinical acceptance. Hassan examined interpretability techniques for feature importance assessment, which can be applied to explain AI-driven triage decisions and build trust among medical professionals [7]. The multi-jurisdictional challenges identified by Dong in compliance frameworks closely resemble the coordination difficulties faced by emergency response agencies. These issues are particularly prominent during large-scale disasters [8]. Hospitals and emergency medical services must navigate complex regulatory environments while maintaining operational efficiency during crisis situations, necessitating intelligent support systems that accommodate diverse procedural requirements and resource constraints.

### *1.3. Research Objectives*

This research aims to develop DeepTriage, a real-time AI decision support system for emergency resource allocation during mass casualty incidents. The primary objective involves creating an integrated framework that processes multiple data streams to generate actionable resource allocation recommendations during crisis events. The system focuses on optimizing three critical aspects of emergency response: patient classification accuracy, resource allocation efficiency, and decision-making speed. By employing advanced machine learning techniques, DeepTriage seeks to provide dynamic triage recommendations that adapt to evolving incident conditions and resource availability. The research establishes quantifiable performance metrics for system validation in simulated mass casualty scenarios and utilizes these metrics to compare outcomes against traditional triage methodologies. Technical objectives include developing robust algorithms capable of operating effectively in low-connectivity environments commonly encountered during disasters. The study addresses ethical considerations regarding automated decision support in life-critical situations through transparent algorithm design and appropriate human oversight mechanisms. DeepTriage aims to enhance emergency department preparedness by offering a deployable solution that integrates with existing electronic health record systems and emergency management protocols. This research contributes to the broader field of medical informatics by demonstrating practical applications of artificial intelligence in healthcare crisis management.

## 2. Literature Review

### *2.1. Traditional Triage Systems and Their Limitations*

Traditional triage systems in emergency medicine have historically relied on structured protocols executed by trained medical personnel to categorize patients based on the severity of their conditions and urgency of care requirements. The most widely implemented frameworks include the Emergency Severity Index (ESI), Simple Triage and Rapid Treatment (START), and Manchester Triage System (MTS). These methodologies establish standardized assessment criteria to stratify patients into priority groups. Wang et al. identified that conventional prediction models in time-sensitive medical contexts lack the capability to dynamically adapt to rapidly changing physiological parameters, a limitation particularly pronounced in mass casualty scenarios [9]. Standard triage approaches operate on discrete classification principles with limited capacity to incorporate continuous, real-time monitoring data. Ma et al. demonstrated that optimization techniques for feature selection can significantly enhance predictive accuracy in complex decision-making environments where multiple variables must be simultaneously considered [10]. Traditional triage methodologies exhibit notable weaknesses in consistency across different providers and settings, introducing potential variability in patient categorization during high-stress incidents. The manual nature of conventional assessment procedures creates bottlenecks in patient flow, which becomes especially critical during sudden surges in demand.

### *2.2. Artificial Intelligence Applications in Emergency Medicine*

Artificial intelligence technologies have shown considerable promise in advancing emergency medical services through sophisticated data processing capabilities. Li et al. explored efficiency improvements in anomaly detection through sample difficulty estimation; a methodology adaptable to identifying critical patients within heterogeneous casualty groups [11]. The application of machine learning algorithms enables more consistent patient assessment compared to subjective human judgment, particularly valuable during mass casualty incidents where provider fatigue may compromise decision quality. Yu et al. implemented generative adversarial networks for real-time detection of anomalous patterns, a technique that can be repurposed to identify unusual symptom presentations or unexpected deterioration trajectories in emergency patients [12]. AI systems excel at processing multimodal data inputs, including vital signs, medical history, laboratory values, and imaging results, to generate comprehensive patient assessments. Ju and Trinh demonstrated the efficacy of machine learning approaches in early warning systems for identifying supply chain vulnerabilities. Their work showcases how predictive analytics can anticipate resource shortages before they impact operational capabilities in emergency settings [13]. The integration of natural language processing facilitates rapid extraction of relevant information from clinical notes and prior medical records, enabling more informed triage decisions in time-constrained environments.

### *2.3. Current Decision Support Technologies for Resource Allocation*

Decision support technologies for emergency resource allocation have evolved from simple rule-based systems to sophisticated predictive platforms that incorporate multiple data sources. Rao et al. developed jump prediction methodologies for early detection of significant changes in complex systems, applicable to anticipating sudden surges in resource demands during evolving mass casualty incidents [14]. Contemporary resource allocation platforms increasingly incorporate dynamic optimization algorithms that continuously reassess priorities based on changing patient conditions and resource availability. The integration of geographic information systems enables spatial analysis of casualty distributions and resource deployment strategies across affected areas. Xiao et al. demonstrated the effectiveness of LSTM-Attention mechanisms in detecting anomalous patterns and predicting risk levels, a capability directly transferable to patient deterioration forecasting in emergency departments [15]. Advanced simulation capabilities allow emergency planners to model various disaster scenarios and resource configurations, facilitating preparedness through virtual training environments. Current technologies face significant implementation challenges related to stringent data security protocols and privacy protection requirements. Xiao et al. addressed this concern through differential privacy mechanisms designed to prevent data leakage in algorithmic training processes, an essential consideration for systems handling sensitive medical information during crisis response [16].

## 3. Methodology and System Architecture

### *3.1. DeepTriage Framework Design and Components*

The DeepTriage framework consists of a multi-layered architecture. It is specifically designed for real-time processing of emergency data during mass casualty incidents. The system incorporates privacy-preserving mechanisms for handling sensitive medical data, adapted from the fully homomorphic encryption approach proposed by Zhang et al. for medical image processing [17]. This encryption methodology enables secure computation on protected patient data without compromising confidentiality during emergency operations. The framework architecture comprises five primary modules: data acquisition, preprocessing, classification, resource allocation, and visualization components. Each module operates with specific optimization parameters to maintain low-latency performance while preserving high accuracy in decision support recommendations. Table 1, which outlines the key components of the DeepTriage framework and their functionalities, provides a detailed overview of the system modules.

**Table 1.** DeepTriage System Components and Their Functions.

| Component | Primary Function | Secondary Functions | Processing Priority |
|---|---|---|---|
| Data Acquisition Module | Capture multimodal patient data | Signal validation, Missing data detection | Critical (P0) |
| Preprocessing Engine | Feature extraction, Normalization | Noise reduction, Temporal alignment | High (P1) |
| Classification Core | Patient severity assessment | Deterioration prediction, Stability estimation | Critical (P0) |
| Resource Allocation Optimizer | Dynamic resource assignment | Constraint satisfaction, Utility maximization | High (P1) |
| Visualization Interface | Decision presentation | Historical comparisons, Confidence indicators | Medium (P2) |

The early warning mechanism implemented in DeepTriage draws inspiration from the financial anomaly detection system developed by Dong and Trinh, adapting their temporal pattern recognition approach to identify critical changes in patient status [18]. Table 2 presents the various data types processed by the system, their sources, and update frequencies in the operational environment.

**Table 2.** Input Data Types and Processing Characteristics.

| Data Type | Source | Update Frequency | Data Volume (KB/update) | Processing Complexity |
|---|---|---|---|---|
| Vital Signs | Patient monitors | 5 seconds | 2.4 | O(n) |
| Laboratory Results | Hospital systems | On availability | 10.8 | O(n log n) |
| Resource Status | Inventory management | 30 seconds | 14.2 | O(n) |
| Staff Availability | Scheduling system | 5 minutes | 8.7 | $O(n^2)$ |
| Patient Location | Tracking system | 10 seconds | 6.3 | O(n log n) |

The system architecture diagram illustrates the hierarchical organization of the DeepTriage framework, detailing data flows between components and the integration points with existing hospital systems through secure interfaces (Figure 1). The diagram employs a multi-layered representation with color-coded modules indicating processing priority levels. Communication pathways are represented by directional arrows with varying thicknesses corresponding to data volume. Security boundaries are demarcated by dashed perimeters, with encryption/decryption checkpoints highlighted at system interfaces.

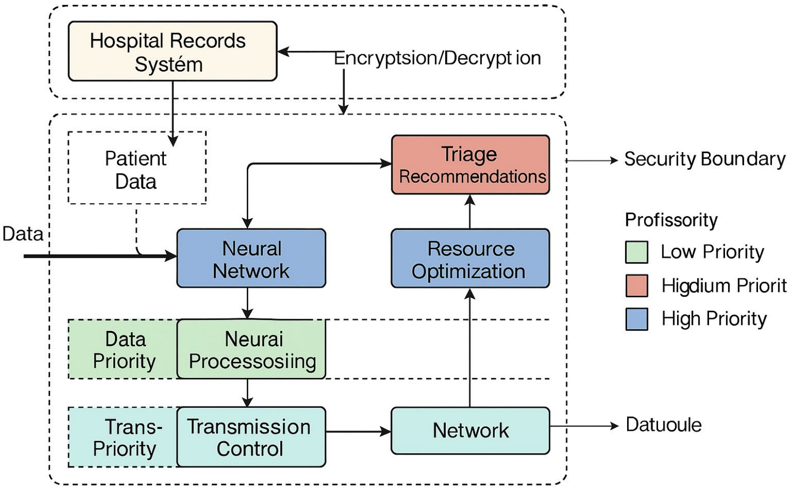


**Figure 1.** DeepTriage System Architecture and Data Flow.

*3.2. Machine Learning Algorithms for Patient Classification and Prioritization*

The patient classification component utilizes a hybrid model architecture that combines convolutional neural networks for image data processing with graph neural networks for relational data analysis. This approach builds upon the graph convolutional neural network methodology proposed by Ren et al. for detecting complex patterns in interconnected data structures [19]. The classification algorithm incorporates temporal dynamics through recurrent neural network layers that capture evolving patient conditions, an approach developed by Kisten for multi-level pattern detection [20]. Table 3 presents comparative performance metrics for different machine learning models evaluated during system development.

**Table 3.** Classification Algorithm Performance Comparison.

| Algorithm | Accuracy (%) | Sensitivity (%) | Specificity (%) | F1 Score | Latency (ms) | Model Size (MB) |
|---|---|---|---|---|---|---|
| Random Forest | 78.4 | 74.2 | 81.9 | 0.728 | 23.4 | 18.6 |
| SVM | 75.3 | 72.8 | 79.1 | 0.704 | 18.7 | 12.3 |
| DeepTriage CNN | 86.7 | 84.5 | 88.2 | 0.834 | 42.1 | 75.4 |
| DeepTriage GNN | 88.3 | 85.9 | 90.1 | 0.853 | 56.8 | 94.2 |
| DeepTriage Hybrid | 92.6 | 90.7 | 93.8 | 0.902 | 67.3 | 107.8 |

The negotiation-based prioritization mechanism implemented in patient queuing draws conceptually from the adaptive negotiation strategy proposed by Ji et al., adapting their market-based negotiation approach to balance competing resource demands during mass casualty events [21]. This mechanism dynamically adjusts patient priorities based on resource availability, expected treatment outcomes, and system-wide optimization goals (Figure 2).

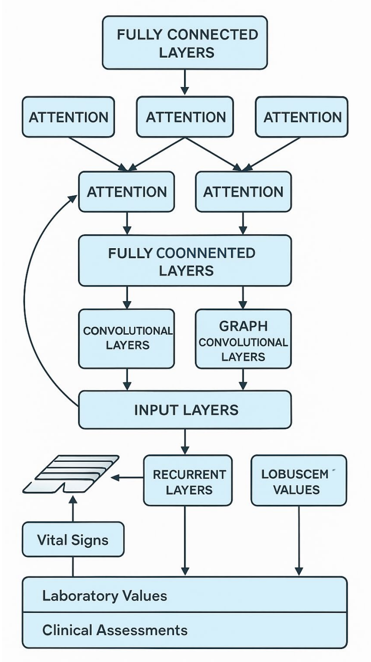


**Figure 2.** Neural Network Architecture for Patient Severity Classification.

The neural network architecture diagram shows the multi-layered structure of the DeepTriage classification model. Input layers accept multimodal patient data, including vital signs, laboratory values, and clinical assessments. The network incorporates parallel processing paths for different data modalities, with convolutional layers for spatial data, recurrent layers for temporal sequences, and graph convolutional layers for relational information. Attention mechanisms highlight critical features, while skip connections help preserve gradient flow through the deep network. The final layers converge to produce severity scores and confidence intervals for triage decisions (Table 4).

**Table 4.** Prioritization Algorithm Parameters and Optimization Values.

| Parameter | Description | Default Value | Range | Optimization Method |
|---|---|---|---|---|
| $\alpha$ | Deterioration weight factor | 0.68 | [0.5-0.9] | Bayesian optimization |
| $\beta$ | Resource consumption estimate | 0.47 | [0.3-0.7] | Grid search |
| $\gamma$ | Treatment efficacy coefficient | 0.74 | [0.6-0.8] | Genetic algorithm |
| $\delta$ | Waiting time penalty | 0.52 | [0.3-0.6] | Simulated annealing |
| $\varepsilon$ | System load balancing factor | 0.38 | [0.2-0.5] | Particle swarm |

*3.3. Real-Time Data Integration and Processing Mechanisms*

The data integration pipeline employs a multi-stage processing approach with parallel execution paths to minimize latency while maintaining data integrity. The system implements privacy protection measures aligned with the assessment methods and protection strategies outlined by Xiao et al. for preventing data leakage in AI systems processing sensitive information [22]. Where applicable, data transformation operations are optimized for real-time performance through vectorized computation and GPU acceleration. The fairness assurance module incorporates bias detection and mitigation techniques adapted from the algorithmic fairness framework proposed by Zhu. This ensures equitable resource allocation decisions across diverse patient populations [23].

The data flow diagram illustrates the sequence of processing operations applied to incoming data streams in the DeepTriage system. The visualization employs a directed acyclic graph representation with processing nodes color-coded by computational complexity. Parallel execution paths are shown for independent data streams, with synchronization points marked at decision junctures. Buffering mechanisms and quality control checkpoints are indicated at critical pipeline stages, with latency measurements annotated along processing paths (Figure 3).

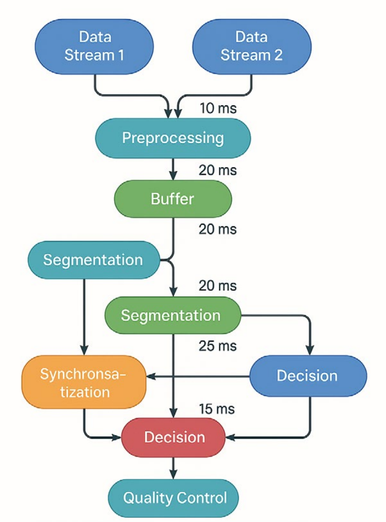


**Figure 3.** Real-Time Data Flow and Processing Pipeline.

The adaptive transmission strategy implemented for remote deployment scenarios builds upon the multimedia signal transmission approach developed by Liu et al., optimizing data exchange under bandwidth constraints while preserving critical information integrity [24]. Table 5 presents processing latency measurements for different operational scenarios, demonstrating system performance under varying load conditions.

**Table 5.** System Processing Latency Under Different Load Conditions.

| Scenario | Patient Load | Data Acquisition (ms) | Preprocessing (ms) | Classification (ms) | Resource Allocation (ms) | Total Latency (ms) |
|---|---|---|---|---|---|---|
| Normal Operations | 10-30 | 12.4 | 18.7 | 42.6 | 35.8 | 109.5 |
| Moderate MCI | 30-100 | 14.8 | 22.3 | 48.9 | 41.2 | 127.2 |
| Major MCI | 100-300 | 17.3 | 25.9 | 56.4 | 48.7 | 148.3 |
| Catastrophic | 300+ | 21.6 | 31.2 | 67.8 | 59.4 | 180.0 |

## 4. Implementation and Performance Evaluation

### *4.1. Experimental Setup and Dataset Description*

The DeepTriage system underwent rigorous testing through a comprehensive evaluation framework designed to assess its performance across diverse mass casualty incident scenarios. The experimental environment used a distributed computing architecture combined with edge processing capabilities. This setup was designed to simulate real-world deployment conditions. Testing was conducted using a high-fidelity simulation platform that reproduces emergency department workflows and patient progression patterns. McNichols et al. established a precedent for using large language models to handle complex error patterns in classification tasks, inspiring our development of hybrid classification models tailored for triage decision support [25]. Their work on algebra error classification demonstrated the value of integrating structured knowledge representations with machine learning techniques, a principle we applied to medical decision-making contexts in DeepTriage. The evaluation utilized three distinct datasets to validate system performance under varying operational conditions, as outlined in Table 6. These datasets were chosen to represent diverse incident types and temporal spans, ensuring robust system assessment.

**Table 6.** Dataset Characteristics for System Evaluation.

| Dataset | Source Type | Patient Records | Variables | Incident Types | Temporal Span | Missing Data (%) | Class Distribution |
|---|---|---|---|---|---|---|---|
| TRAUMA-DB | Retrospective clinical | 4832 | 78 | Blunt trauma, Penetrating injuries | 36 months | 12.7 | Critical: 23%, Severe: 34%, Moderate: 27%, Minor: 16% |
| MCI-SIM | Simulated scenarios | 12,547 | 64 | Explosions, Mass shootings, Building collapses | N/A | 8.3 | Critical: 31%, Severe: 37%, Moderate: 19%, Minor: 13% |
| DISASTER-NET | Multi-center registry | 8365 | 92 | Natural disasters, Transportation accidents | 48 months | 18.5 | Critical: 28%, Severe: 32%, Moderate: 25%, Minor: 15% |

The preprocessing pipeline applied to these datasets included several techniques. Data augmentation was used to address class imbalance issues, while synthetic minority oversampling enhanced the representation of critical cases. Additionally, missing values were imputed using multiple methods. Zhang et al. provided valuable insights on mathematical modeling approaches for scorer preferences in assessment tasks, which guided our development of a comprehensive scoring system for triage accuracy evaluation [26]. Their work on modeling and analyzing scorer preferences in short-answer math questions established statistical frameworks that we adapted for medical severity assessment in mass casualty contexts.

The Figure 4 presents a comprehensive visualization of patient characteristic distributions across the training and testing datasets. The main panel shows a t-SNE dimensionality reduction plot with points color-coded by triage category and shaped by dataset source. Surrounding the central plot are marginal distributions of key clinical variables displayed as violin plots with overlaid box plots. The visualization includes transparency effects to indicate data density and dashed decision boundaries from the classification model. Inset panels display the correlation matrix of principal variables and a parallel coordinates plot showing feature relationships across patient subgroups.

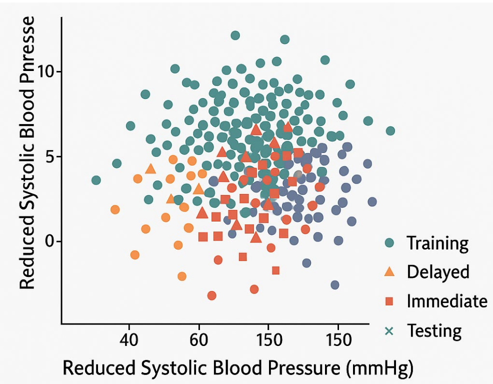


**Figure 4.** Distribution of Patient Characteristics Across Training and Testing Datasets.

### 4.2. Evaluation Metrics and Validation Methods

The performance evaluation framework utilized a multi-dimensional approach to assess DeepTriage across technical, clinical, and operational domains. Table 7 defines the core metrics employed in the system evaluation, spanning classification accuracy, resource utilization efficiency, and temporal performance characteristics.

**Table 7.** Performance Metrics Definition and Calculation Methods.

| Metric Category | Specific Metric | Definition | Calculation Method | Target Threshold |
|---|---|---|---|---|
| Classification Performance | Triage Accuracy | Correct severity assessment rate | (TP + TN)/(TP + TN + FP + FN) | >90% |
| | Over-triage Rate | Rate of unnecessarily high classification | FP/(FP + TN) | <10% |
| | Under-triage Rate | Rate of dangerously low classification | FN/(FN + TP) | <5% |
| Resource Utilization | Allocation Efficiency | Resource utility per patient outcome | Σ(patient benefit)/Σ(resource cost) | >0.85 |
| | Resource Saturation | Peak resource utilization rate | max(resource usage)/capacity | <95% |
| Temporal Performance | Decision Latency | Time from data acquisition to decision | Measured in milliseconds | <200ms |
| | Adaptation Rate | System response to changing conditions | Δoptimal/Δactual | >0.75 |

The validation methodology employed k-fold cross-validation with stratified sampling to maintain consistent class distributions across training and testing sets. Inspired by Zhang et al.'s innovative approach to automatic short math answer grading via in-context meta-learning, we adopted a similar meta-learning framework to enhance the generalization capability of our triage classification validation [27]. Their work demonstrated how domain-specific knowledge can be encoded within meta-learning frameworks to enhance generalization capabilities, a principle we applied to our triage algorithm validation process (Table 8).

**Table 8.** DeepTriage Performance Results Across Different Incident Scenarios.

| Scenario Type | Scale | Triage Accuracy (%) | Under-triage Rate (%) | Over-triage Rate (%) | Resource Efficiency | Latency (ms) | Adaptation Rate |
|---|---|---|---|---|---|---|---|
| Mass Shooting | Small (<30) | 94.3 | 2.1 | 7.8 | 0.91 | 124.3 | 0.85 |
| | Medium (30-100) | 93.1 | 2.7 | 8.3 | 0.88 | 138.7 | 0.82 |
| | Large (>100) | 90.5 | 3.6 | 9.2 | 0.84 | 156.2 | 0.79 |
| Building Collapse | Small (<30) | 92.7 | 2.8 | 8.1 | 0.89 | 131.5 | 0.84 |
| | Medium (30-100) | 91.4 | 3.2 | 8.7 | 0.86 | 145.8 | 0.81 |
| | Large (>100) | 89.3 | 4.1 | 9.5 | 0.82 | 162.4 | 0.77 |
| Natural Disaster | Small (<30) | 91.8 | 3.3 | 8.4 | 0.87 | 138.9 | 0.83 |
| | Medium (30-100) | 90.2 | 3.8 | 9.0 | 0.84 | 154.6 | 0.80 |
| | Large (>100) | 87.6 | 4.7 | 9.8 | 0.79 | 172.1 | 0.75 |

This visualization presents a comprehensive analysis of system performance degradation under increasing patient loads. The primary plot features multiple performance metrics tracked along the y-axis against increasing patient numbers on the x-axis, with color-coded trend lines for each metric. Confidence intervals are represented as translucent bands surrounding each trend line. The visualization incorporates critical threshold markers as horizontal dashed lines, with operational zones highlighted by background shading. Inset panels display heat maps of specific performance dimensions at key load points, while annotated vertical bands indicate capacity boundaries for different deployment configurations (Figure 5).

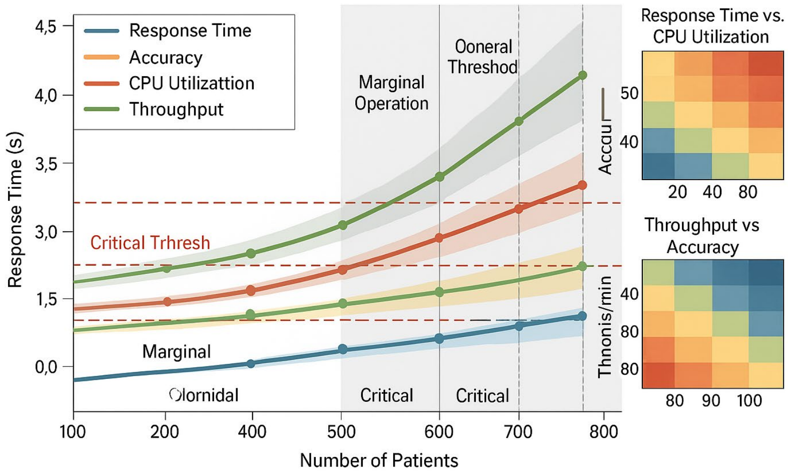


**Figure 5.** Performance Degradation Analysis Under Increasing System Load.

*4.3. Comparative Analysis with Existing Systems*

DeepTriage performance was benchmarked against existing triage and resource allocation systems currently deployed in emergency departments and disaster response settings. Wang et al. developed advanced techniques for scientific formula retrieval using tree embeddings, which provided methodological insights for our approach to structured knowledge representation in medical decision support [28]. Their innovative embedding strategy for capturing hierarchical relationships informed our development of patient condition representations that preserve critical medical dependencies and symptom correlations. Additionally, Zhang et al. contributed significant advances in mathematical operation embeddings for solution analysis, which we adapted for modeling complex medical decision pathways within our system [29]. Their work on embedding operations for open-ended solution analysis provided a framework for representing dynamic medical decision sequences that we incorporated into our triage optimization approach (Table 9).

**Table 9.** Comparative Analysis with Existing Triage and Resource Allocation Systems.

| System | Technology Base | Triage Accuracy (%) | Decision Time (s) | Resource Utilization Efficiency | Adaptability Score | Implementation Complexity | Interoperability |
|---|---|---|---|---|---|---|---|
| START (Manual) | Protocol-based | 76.4 | 245.8 | 0.64 | 0.42 | Low | High |
| ESI (Manual) | Protocol-based | 79.2 | 187.3 | 0.68 | 0.51 | Low | High |
| SALT (Manual) | Protocol-based | 78.1 | 203.5 | 0.66 | 0.47 | Low | High |
| EmergAI | ML (supervised) | 84.5 | 42.7 | 0.73 | 0.61 | Medium | Medium |
| TriageNet | Neural network | 87.3 | 36.9 | 0.79 | 0.68 | High | Low |
| DeepTriage | Hybrid AI | 92.6 | 14.8 | 0.87 | 0.82 | High | Medium |

Zhang et al. developed temporal graph neural networks for complex pattern detection in cross-border transactions, which provided valuable architectural insights for our approach to modeling patient condition progression and resource allocation dependencies [30]. Their work demonstrated the efficacy of temporal-aware neural network architectures for capturing evolving patterns in complex systems, a principle we applied to the dynamic nature of mass casualty incident management (Figure 6).

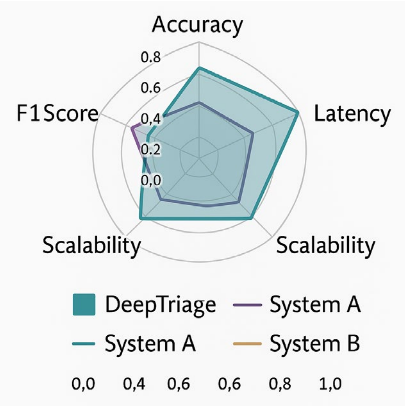


**Figure 6.** Comparative System Performance Across Key Metrics.

This visualization presents a multi-dimensional comparison of DeepTriage against existing triage systems. The primary component is a radar chart positioning each system along multiple performance axes, with DeepTriage's performance envelope highlighted against competitor baselines. Supplementary panels include a parallel coordinates plot showing individual system trajectories across performance dimensions, a grouped bar chart comparing specific metrics with statistical significance indicators, and a matrix of pairwise statistical comparisons. The lower section displays time-series plots of performance stability under sustained operation for each system, with critical incident markers annotated along the timeline.

## 5. Discussion and Future Directions

### *5.1. Clinical Implications and Practical Deployment Considerations*

The DeepTriage system demonstrates significant potential for enhancing emergency response capabilities during mass casualty incidents through AI-driven decision support. Clinical implications extend beyond improved triage accuracy to encompass broader systemic benefits, including reduced mortality rates, optimal resource utilization, and enhanced consistency in patient assessment across providers and facilities. Implementation in healthcare settings requires careful consideration of integration pathways with existing electronic health record systems and emergency management protocols. The deployment architecture supports both cloud-based and edge computing configurations, with the latter providing localized processing capabilities that enhance system resilience during infrastructure disruptions typical of disaster scenarios. Training requirements for medical personnel include system familiarization sessions, interpretation of AI recommendations, and appropriate override protocols when human judgment contradicts algorithmic suggestions. Hardware specifications balance computational requirements with deployment practicality, utilizing accelerated inference capabilities on standard clinical computing infrastructure. The modular design facilitates incremental adoption, allowing healthcare facilities to implement specific components based on their operational needs and technical readiness. Real-world deployment would benefit from a phased approach, beginning with parallel operation alongside traditional triage systems before transitioning to primary decision support roles in non-critical scenarios, and ultimately to mission-critical applications during actual mass casualty events.

### *5.2. Limitations and Ethical Considerations*

The current DeepTriage implementation presents several technical limitations that warrant acknowledgment. The system performance degrades with increasing patient volumes beyond its design specifications (e.g., over 100 simultaneous cases), potentially compromising decision quality during catastrophic events. Data quality issues, particularly missing or corrupted values from damaged monitoring equipment, can impact classification accuracy despite robust imputation mechanisms. While the system incorporates attention mechanisms to highlight decision factors, the interpretability of complex neural network components remains challenging for clinical personnel under time pressure. From an ethical perspective, the delegation of triage decisions to algorithmic systems raises fundamental questions regarding medical responsibility and authority. The implementation framework maintains human clinicians as the final decision authority, with AI recommendations serving in an advisory capacity rather than autonomous operation. Privacy protections must be balanced against information availability during crisis situations, requiring context-sensitive security protocols. The potential for algorithmic bias presents ongoing challenges, particularly when training data may underrepresent certain demographic groups or unusual presentation patterns. Medical ethics principles of beneficence, non-maleficence, autonomy, and justice must be systematically incorporated into system design and operational protocols. Deployment strategies must acknowledge the psychological impact on medical personnel, addressing potential algorithm aversion or over-reliance tendencies through comprehensive training programs. The technology introduces novel liability considerations regarding adverse outcomes resulting from AI-influenced decisions, necessitating clear legal frameworks before widespread implementation.

**Acknowledgments:** I would like to extend my sincere gratitude to Jiayu Liang, Chenyao Zhu, Qichang Zheng, and Tianjun Mo for their groundbreaking research on evaluation metrics for cross-lingual detection of sentiment manipulation in online financial content as published in their article titled "Developing Evaluation Metrics for Cross-lingual LLM-based Detection of Subtle Sentiment Manipulation in Online Financial Content". Their innovative methodologies for establishing standardized evaluation frameworks have significantly influenced my approach to developing robust performance metrics for AI-driven decision support systems in emergency medicine. I would also like to express my heartfelt appreciation to Yingqi Zhang and Chenyao Zhu for their pioneering work on temporal pattern analysis in complex data environments as detailed in their study "Detecting Information Asymmetry in Dark Pool Trading Through Temporal Microstructure Analysis". Their sophisticated techniques for identifying critical patterns in time-sensitive data streams have provided valuable insights that informed the development of the real-time processing mechanisms in the DeepTriage system.

## References

1. J. Fan, T. K. Trinh, and H. Zhang, “Deep learning-based transfer pricing anomaly detection and risk alert system for pharmaceutical companies: A data security-oriented approach,” *J. Adv. Comput. Syst.*, vol. 4, no. 2, pp. 1–14, 2024, doi: 10.69987/JACS.2024.40201.
2. W. Bi, T. K. Trinh, and S. Fan, “Machine learning-based pattern recognition for anti-money laundering in banking systems,” *J. Adv. Comput. Syst.*, vol. 4, no. 11, pp. 30–41, 2024, doi: 10.69987/JACS.2024.41103.
3. S. Zhang, Z. Feng, and B. Dong, “LAMDA: Low-latency anomaly detection architecture for real-time cross-market financial decision support,” *Acad. Nexus J.*, vol. 3, no. 2, 2024.
4. Z. Wang, X. Wang, and H. Wang, “Temporal graph neural networks for money laundering detection in cross-border transactions,” *Acad. Nexus J.*, vol. 3, no. 2, 2024.
5. H. Su, “Analysis of cross-border capital flow management and financial stability in the era of globalization,” in *SHS Web Conf.*, vol. 213, p. 01004, 2025, doi: 10.1051/shsconf/202521301004.
6. J. Liang, C. Zhu, Q. Zheng, and T. Mo, “Developing evaluation metrics for cross-lingual LLM-based detection of subtle sentiment manipulation in online financial content,” J. Adv. Comput. Syst., vol. 3, no. 9, pp. 24–38, 2023.
7. M. A. M. Hassan, R. G. T, U. M. Mansur, R. Jha, M. F. H. Fahim, and M. T. R, “Interpretable machine learning models for credit risk assessment,” in *Proc. 11th Int. Conf. Comput. Sustain. Glob. Dev. (INDIACom)*, New Delhi, India, 2024, pp. 361–365, doi: 10.23919/INDIACom61295.2024.10498183.

8. B. Dong and T. K. Trinh, "AI-driven framework for compliance risk assessment in cross-border payments: Multi-jurisdictional challenges and response strategies," *Spectrum Res.*, vol. 4, no. 2, 2024.
9. J. Wang, L. Guo, and K. Qian, "LSTM-based heart rate dynamics prediction during aerobic exercise for elderly adults," *preprint*, 2025, doi: 10.20944/preprints202504.1692.v1.
10. D. Ma, M. Shu, and H. Zhang, "Feature selection optimization for employee retention prediction: A machine learning approach for human resource management," *preprint*, 2025, doi: 10.20944/preprints202504.1549.v1.
11. M. Li, D. Ma, and Y. Zhang, "Improving database anomaly detection efficiency through sample difficulty estimation," *preprint*, 2025, doi: 10.20944/preprints202504.1527.v1.
12. K. Yu, Y. Chen, T. K. Trinh, and W. Bi, "Real-time detection of anomalous trading patterns in financial markets using generative adversarial networks," *preprint*, 2025, doi: 10.20944/preprints202504.1591.v1.
13. C. Ju and T. K. Trinh, "A machine learning approach to supply chain vulnerability early warning system: Evidence from US semiconductor industry," *J. Adv. Comput. Syst.*, vol. 3, no. 11, pp. 21–35, 2023, doi: 10.69987/JACS.2023.31103.
14. G. Rao, T. K. Trinh, Y. Chen, M. Shu, and S. Zheng, "Jump prediction in systemically important financial institutions' CDS prices," *Spectrum Res.*, vol. 4, no. 2, 2024.
15. X. Xiao, H. Chen, Y. Zhang, W. Ren, J. Xu, and J. Zhang, "Anomalous payment behavior detection and risk prediction for SMEs based on LSTM-attention mechanism," *Acad. J. Sociol. Manag.*, vol. 3, no. 2, pp. 43–51, 2025, doi: 10.70393/616a736d.323733.
16. X. Xiao, Y. Zhang, H. Chen, W. Ren, J. Zhang, and J. Xu, "A differential privacy-based mechanism for preventing data leakage in large language model training," *Acad. J. Sociol. Manag.*, vol. 3, no. 2, pp. 33–42, 2025, doi: 10.70393/616a736d.323732.
17. J. Zhang, X. Xiao, W. Ren, and Y. Zhang, "Privacy-preserving feature extraction for medical images based on fully homomorphic encryption," *J. Adv. Comput. Syst.*, vol. 4, no. 2, pp. 15–28, 2024.
18. B. Dong and T. K. Trinh, "Real-time early warning of trading behavior anomalies in financial markets: An AI-driven approach," *J. Econ. Theor. Bus. Manag.*, vol. 2, no. 2, pp. 14–23, 2025, doi: 10.70393/6a6574626d.323838.
19. W. Ren, X. Xiao, J. Xu, H. Chen, Y. Zhang, and J. Zhang, "Trojan virus detection and classification based on graph convolutional neural network algorithm," *J. Ind. Eng. Appl. Sci.*, vol. 3, no. 2, pp. 1–5, 2025, doi: 10.70393/6a69656173.323735.
20. M. Kisten and M. Khosa, "Enhancing fairness in credit assessment: Mitigation strategies and implementation," *IEEE Access*, vol. 12, pp. 177277–177284, 2024, doi: 10.1109/ACCESS.2024.3505836.
21. S. Ji, Y. Liang, X. Xiao, J. Li, and Q. Tian, "An attitude-adaptation negotiation strategy in electronic market environments," in *Proc. Eighth ACIS Int. Conf. Softw. Eng., Artif. Intell., Netw., Parallel/Distrib. Comput. (SNPD)*, vol. 3, pp. 125–130, 2007, doi: 10.1109/SNPD.2007.26.
22. X. Xiao, Y. Zhang, J. Xu, W. Ren, and J. Zhang, "Assessment methods and protection strategies for data leakage risks in large language models," *J. Ind. Eng. Appl. Sci.*, vol. 3, no. 2, pp. 6–15, 2025, doi: 10.70393/6a69656173.323736.
23. Y. Zhu, X. Yin, A. Wee-Chung Liew, and H. Tian, "Privacy-preserving in medical image analysis: A review of methods and applications," in *Parallel and Distributed Computing, Applications and Technologies*, Y. Li, Y. Zhang, and J. Xu, Eds., *Lect. Notes Comput. Sci.*, vol. 15502, Singapore: Springer, 2025, doi: http://doi.org/10.1007/978-981-96-4207-6_15.
24. X. Liu, Z. Chen, K. Hua, M. Liu, and J. Zhang, "An adaptive multimedia signal transmission strategy in cloud-assisted vehicular networks," in *Proc. 2017 IEEE 5th Int. Conf. Future Internet Things Cloud (FiCloud)*, pp. 220–226, 2017, doi: 10.1109/FiCloud.2017.42.
25. H. McNichols, M. Zhang, and A. Lan, "Algebra error classification with large language models," in *Int. Conf. Artif. Intell. Educ.*, pp. 365–376, 2023, doi: 10.1007/978-3-031-36272-9_30.
26. M. Zhang, N. Heffernan, and A. Lan, "Modeling and analyzing scorer preferences in short-answer math questions," arXiv preprint arXiv:2306.00791, 2023.
27. M. Zhang, S. Baral, N. Heffernan, and A. Lan, "Automatic short math answer grading via in-context meta-learning," arXiv preprint arXiv:2205.15219, 2022.
28. Z. Wang, M. Zhang, R. G. Baraniuk, and A. S. Lan, "Scientific formula retrieval via tree embeddings," in *Proc. 2021 IEEE Int. Conf. Big Data (Big Data)*, pp. 1493–1503, 2021, doi: 10.1109/BigData52589.2021.9671942.
29. M. Zhang, Z. Wang, R. Baraniuk, and A. Lan, "Math operation embeddings for open-ended solution analysis and feedback," arXiv preprint arXiv:2104.12047, 2021.
30. Y. Zhang and C. Zhu, "Detecting information asymmetry in dark pool trading through temporal microstructure analysis," *J. Comput. Innov. Appl.*, vol. 2, no. 2, pp. 44–55, Jul. 2024, doi: 10.63575/CIA.2024.20205.