Research Paper

Real-Time Medical Resource Allocation in 6G-Enabled IoT Environment using Multi-Head Attention-Based Graph Convolutional Network

Arun Chandran[1*], Santhoshkumar Sundar[2], Jegathesh Amalraj Joseph[3], Suresh Thiyagarajan[4]

1. *Department of Computer Science, Arumugam Pillai Seethai Ammal College, Tiruppattur - 630211, Tamil Nadu, India.*
2. *Department of Computer Science, Alagappa University, Karaikudi – 630003, Tamil Nadu, India.*
3. *Department of Computer Science, Government Arts and Science College, Tittagudi, Cuddalore-606106, Tamil Nadu, India.*
4. *Department of AIML, K. Ramakrishnan College of Engineering (8115), Samayapuram, Trichy-621112, Tamil Nadu, India.*

Citation Chandran A, Sundar S, Joseph JA, Thiyagarajan S. Real-Time Medical Resource Allocation in 6G-Enabled IoT Environment using Multi-Head Attention-Based Graph Convolutional Network. International Journal of Medical Toxicology and Forensic Medicine. 2026; 16:E52673.

https://doi.org/10.22037/ijmtfm.v16.52673

Article info:

Received: 01 July, 2026

Accepted: 16 July, 2026

Published: 31 August, 2026

Keywords:

Resource utilization efficiency, Medical resource allocation, 6G-Network, Smart healthcare, Internet of things

ABSTRACT

Background: The incorporation of sixth-generation (6G) networks with the Internet of Things (IoT) is transforming smart healthcare. The wide range of IoT devices and their offloading demands increase latency and bandwidth requirements, impacting performance.

Methods: This study proposes an Attention-based Graph Convolutional Approach for Efficient Medical Resource Allocation (AGCA-EMRA) methodology. The proposed model uses a conditional mutual information maximization-based feature selection approach to identify the most informative features while removing the redundant ones. For classification, a multi-head spatiotemporal attention graph convolutional network captures both spatial and temporal correlations in the data. The model is further optimized using the Ranger optimizer.

Results: An extensive experimental analysis is conducted on the IoT-Driven MR Allocation for 6G Network Dataset. The proposed AGCA-EMRA technique exhibits superior performance with an accuracy of 95.97%.

Conclusion: The comparative result analysis demonstrates the improvement of the AGDRL-ECFA method and confirms its effectiveness for intelligent cybercrime forensic investigation and decision-support applications.

*** Corresponding Author:**

Arun Chandran, PhD

Department of Computer Science, Arumugam Pillai Seethai Ammal College, Tiruppattur - 630211, Tamil Nadu, India.

E-mail: *drcarunofficial@gmail.com*

Introduction

The convergence of sixth-generation (6G) networks with the Internet of Things (IoT) is reshaping smart healthcare by providing intelligent connectivity, high bandwidth, and ultra-low latency across medical systems. Early detection of infections affecting critical parts of the human body can help control disease progression and preserve the integrity of those parts [1]. Unlike conventional diagnostic approaches for early cancer detection, 6G-powered IoT can process massive volumes of medical data. The current healthcare IoT approach is limited by challenges in dynamic resource allocation, data protection, and infrastructure scalability. This method supports dependable and protected data transfer, a key requirement for any resource allocation (RA) structure. Also, 6G will enable AI-powered systems to support data-driven decision-making in resource management using insights collected from IoT devices [2]. However, most existing approaches using the IoT-Driven MR Allocation for 6G Network Dataset primarily focus on basic system parameters and conventional optimization techniques, often overlooking the complex relationships between resource usage patterns and network dynamics.

A brief review on previous works on healthcare resource allocation using IoT. Almadhor et al. [3] introduced a combined AI-driven 6G-IoT healthcare system comprising 3 elements: scalable network optimization, adaptive healthcare resource allocation, and privacy-preserving anomaly detection. Anand and Karthikeyan [4] proposed an innovative method for resource allocation and task scheduling in an edge computing approach employed for health-related monitoring. Singh and Hada [5] developed an innovative Fuzzy Logic–Driven Decision Support Model (FL-DSM) for adaptive resource allocation in an IoT-assisted Smart medical system. Rafiqu et al. [6] developed the ML-RASPF model, an ML-based architecture for effective service distribution in smart healthcare models. Gowda et al. [7] examined the integration of IoT devices and ML-based predictive analytics in medical services to enhance resource allocation. In [8], a DDQN-based resource allocation technique for AI-native medical systems in 6G dense networks has been developed. Muazu et al. [9] implemented resource management in the IoMT by utilizing the edge-empowered BC and FL model. The gradient parameters must be encoded by employing Paillier encryption. The authors [10] introduced an efficient healthcare data management method (EE-HDMM) that can improve medical monitoring by employing an IoT-based wearable sensor environment.

To address these challenges, this work presents an Attention-based Graph Convolutional Approach for Efficient Medical Resource Allocation (AGCA-EMRA) model. The core contribution of this research includes the following:

- Present a conditional mutual information maximization (CMIM)-based feature selection approach, allowing the model to select the informative features.

- Introduction of a multi-head spatiotemporal attention graph network (MHSAGN), which incorporates the graph convolutional network (GCN), a gated recurrent unit (GRU), and an attention mechanism (AM) for capturing spatiotemporal dynamics and prioritizing relevant features.

- Use the ranger optimizer helps to attain model convergence stability.

Materials and Methods

The proposed system introduces an efficient approach for classifying resource utilization efficiency in a 6G-enabled IoT environment. The AGCA-EMRA method follows a systematic approach that integrates comprehensive preprocessing, feature selection, classification, and optimization. Figure 1 demonstrates the complete process of the AGCA-EMRA framework.

Graph Construction Strategy

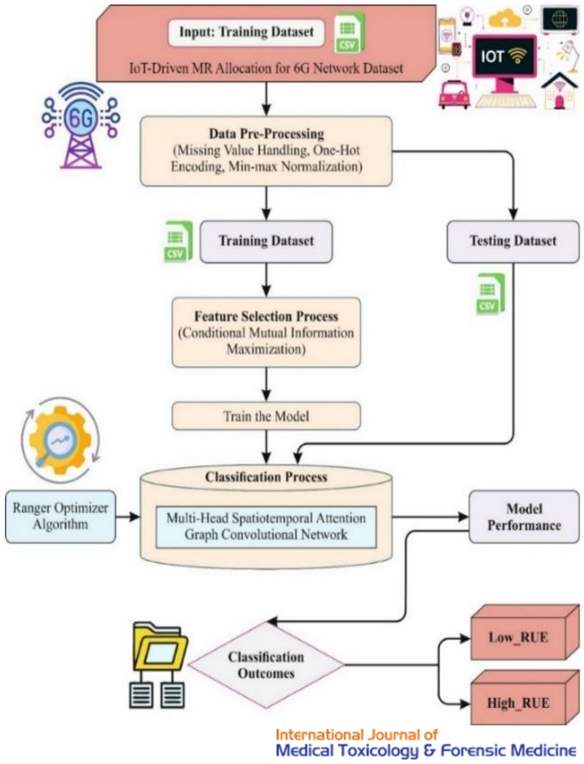


Figure 1. Complete process of the AGCA-EMRA framework.

The proposed architecture represents the healthcare IoT environment as a dynamic graph (G= (V, E, A)), where (V) represents the set of nodes that represent the IoT devices, medical resources, and patient entities, while (E) denotes the communication and interaction relationships over these nodes. The adjacency matrix (A) is constructed by analyzing connectivity, proximity, service dependencies, and resource interaction patterns among the nodes. Initially, a binary adjacency matrix is generated, where $(A_{\{ij\}} = 1)$ shows the existence of a relationship between node (i) and node (j), whereas $(A_{\{ij\}} = 0)$ shows no direct interaction.

The edge weight $(W_{\{ij\}})$ between two connected nodes depicts the intensity of interactions and influence among them. Stronger dependencies are indicated by higher weights.

Network conditions, device status, resource utilization levels, and patient conditions are continuously monitored, allowing the adjacency matrix and edge weights to adapt to real-time changes.

CMIM-based Feature Selection

CMIM is applied to enhance data representation by selecting relevant features that contribute effectively to classification tasks.

CMI

It is a general idea in information theory that calculates dependence amongst 2 variables when the state of a 3rd variable is known [11]. This norm is determined with respect to predictable Kullback-Leibler (KL) divergence as given by,

$$\mathrm{I}(\mathrm{X};\mathrm{Y}|\mathrm{Z}) = \mathrm{H}(\mathrm{X}|\mathrm{Z}) - \mathrm{H}(\mathrm{X}|\mathrm{YZ}) = \sum_{\mathrm{z}\in\mathrm{Z}} p(z) \sum_{\mathrm{x}\in\mathrm{X}} \sum_{\mathrm{y}\in\mathrm{Y}} p\,(\mathrm{xy}|\mathrm{z}) \ln \frac{\mathrm{p}(\mathrm{xy}|\mathrm{z})}{\mathrm{p}(\mathrm{x}|\mathrm{z})\mathrm{p}(\mathrm{y}|\mathrm{z})} \quad (1)$$

It could be interpreted as the shared information between X and Y given a third variable Z, where higher conditional mutual information corresponds to reduced conditional entropies $H(X \mid Z)$ and $H(Y \mid Z)$, and probabilities $p(x)$ are estimated from event frequencies.

CMIM

The CMI is more beneficial in the FS framework as it scores features that meet this standard. This model gains complementarity, redundancy, and relevance, and also offers a better trade-off between stability and accuracy. CMIM is a multivariate feature selection method that uses conditional mutual information to select features by maximizing relevance to the target Y while minimizing redundancy with previously selected features S.

$$\mathrm{CMIM}\,(\mathrm{X_k}) = \min_{\mathrm{j}\in\mathrm{S}} \mathrm{I}\big(\mathrm{X_k};\mathrm{Y}\big|\mathrm{X_j}\big) \quad (2)$$

A greater CMIM value means a feature $\mathrm{X_k}$ is relevant to the target Y and extremely matches the chosen alternative feature $\mathrm{X_j}$ with parameter j ∈ S.

Attention-based Spatio-Temporal Classification

MHSAGN integrates a GCN, a GRU, and an AM to capture spatio-temporal dependencies in the data.

GCN

A graph representation is constructed for medical resources and IoT devices. The GCN models the spatial interactions, and its mathematical form is represented by $G = (V, E)$, where every individual node V with respect to an IoT device or healthcare service, and the interlink among them are denoted as edges E [12]. Patient load with regard to time, denoted as $X \in \mathbb{R}^{N\times P}$. Figure 2 shows the architecture of GCN module.

The GCN computes the matrix multiplication and the 1st-order neighborhood-weighted average of the feature vectors, as given below.

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}X\Theta^{(l)}\right) \quad (3)$$

Here, σ is a non-linear activation function, $\tilde{A} = A + I_N$ is the adjacency matrix with self-loops, $\tilde{D}$ is the degree matrix, $\Theta^{(l)}$ implies learnable parameters at layer l, and $\tilde{D}^{-1/2}$ normalizes node contributions by degree.

$$f(X.A) = \sigma\big(\hat{A}\,\mathrm{ReLU}\big(\hat{A}XW_0\big)W_1\big) \quad (4)$$

Here, the ReLU is the activation function, $\hat{A} = \tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}$, and W_0 and W_1 indicate weight matrices.

GRU

To obtain the model, the temporal dependencies of IoT data sequences, with the spatial features removed using the GCN, have been fed into the GRU. Figure 3 portrays the structure of the GRU model. To gain the input data corresponding to time t, the parameter r_t refers to the reset gate and the term u_t specifies the update gates are given below,

$$u_t = \sigma(W_u[X_t, h_{t-1}] + b_u) \quad (5)$$

$$r_t = \sigma(W_r[X_t, h_{t-1}] + b_r) \quad (6)$$

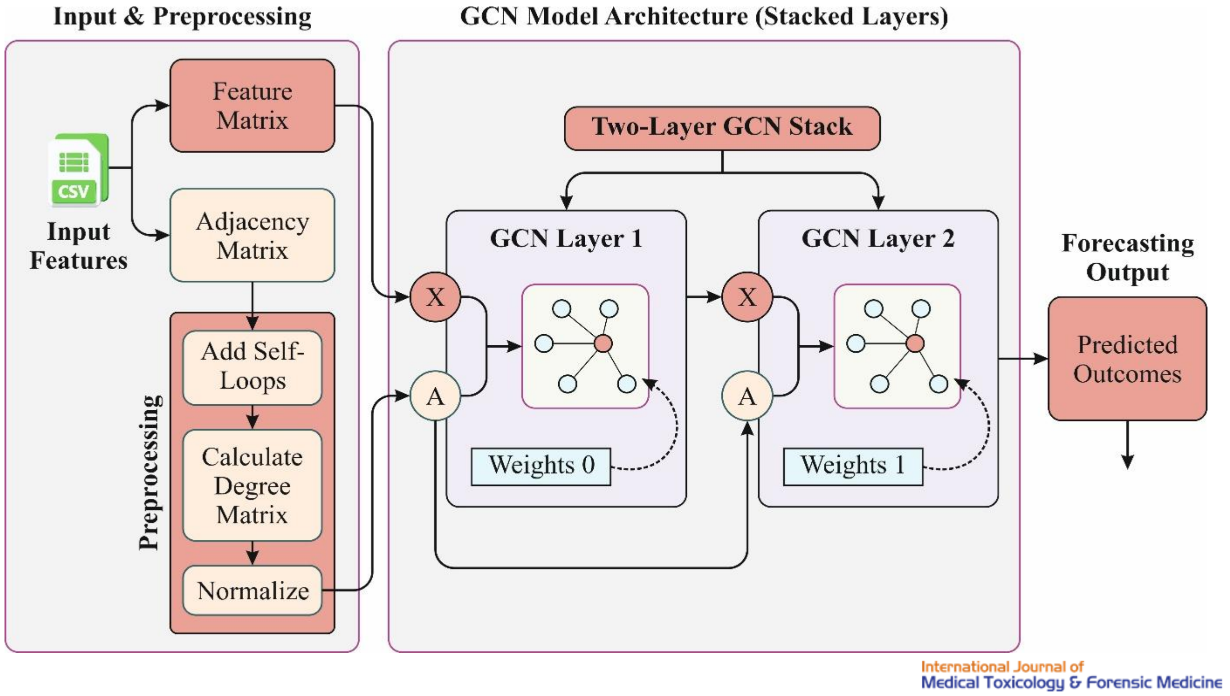


Figure 2. Architecture of the GCN module.

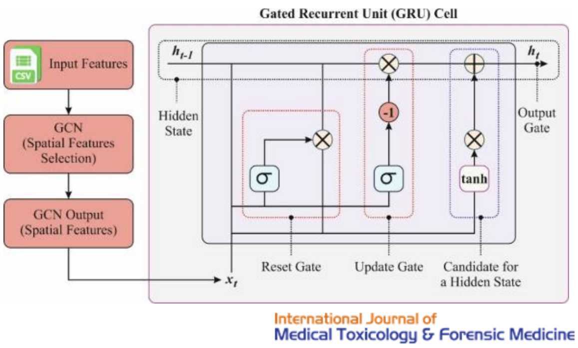


Figure 3. Structure of GRU model.

Let X_t is the feature vector at time t, h_{t-1} is the prior hidden layer (HL), $b$and W describes trainable biases and weight factors, and σ represents the sigmoid function.

The HL (h_t) and candidate memory content (c_t) must be computed by,

$$c_t = \tanh(W_c[X_t, r_t * h_{t-1}] + b_c) \quad (7)$$

$$h_t = u_t * h_{t-1} + (1 - u_t) * c_t \quad (8)$$

AM

AM is used to capture spatiotemporal features of IoT data in the medical domain; multi-head attention (MHA) is combined with the GCN-GRU architecture to incorporate temporal and spatial representations of the input data.

Primarily, the input data X is entered into the GCN to represent the spatial dependency and later the GRU to the features of the temporal model, thus generating HL, $h_i(i = 1,2, \dots, n)$. The MHA comprises value, key, and query heads; all the individuals have a dual multilayer perceptron (MLP) for obtaining a scoring function that can be used to calculate the outcomes HL $H = [h_1, h_2, \dots, h_n]$ by using the softmax function α_{ij}.

$$e_{ij} = W_2(W_1 H + b_1) + b_2 \quad (9)$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{n} \exp(e_{ik})} \quad (10)$$

Now the parameter H represents the final HL output, and $b$and W indicates learning biases and weights for the 1st and 2nd FFN layers. The scores should be multiplied by HL vectors, h, and lastly, it can be denoted by C_t

$$C_t = \sum_{j=1}^{n} \alpha_{ij} h_j \quad (11)$$

Parameter Optimization

The ranger optimizer is an advanced optimization technique for tuning network parameters that combines dual strategies to improve learning efficiency. [13].

$$\theta_{t,i+1} = \theta_{t,i} + A(L, \theta_{t,i-1}, d) \quad (12)$$

Here, the parameter L is the main IoT-MR objective, $A$is the optimizer, $d$is mini-batch data, $i$indexes batches, $t$is iterations, and $\zeta$is the slowly updated weight over batches.

$$\zeta_{t+1} = \zeta_t + \alpha(\theta_{t,k} - \zeta_t)$$

$$= \alpha[\theta_{t,k} + (1-\alpha)\theta_{t-1,k} + \cdots + (1-\alpha)^{t-1}\theta_{0,k}] + (1-\alpha)^t \zeta_0 \qquad (13)$$

Now the mathematical term $\alpha \in (0,1)$ remains the interpolation factor that can be used to collect the input of the faster weight to the slower weight.

Results

This article studies the performance of the AGCA-EMRA technique on the IoT-based MR Allocation for 6G Network Dataset [14]. The dataset contains 500 samples divided into two classes, as given in Table 1.

Table 1. Details on dataset.

Classes	Labels	No. of Samples
Low Resource Utilization Efficiency	Low_RUE	222
High Resource Utilization Efficiency	High_RUE	278
Total Samples		**500**

In Table 2, the confusion matrices and RUE results of the AGCA-EMRA technique are presented for 80%:20% TRPE/TEPE splits. The proposed method achieved an average value of $accur_y$of 95.97% and 95.63% for TRPE and TEPE, respectively.

In Table 3, the confusion matrices and RUE results of the AGCA-EMRA technique are presented for 70%:30% TRPE/TEPE splits. The proposed method obtained an average value of $accur_y$of 94.92%, and 92.90%, for the TRPE and TEPE.

Discussion

To validate the AGCA-EMRA method, a detailed comparison study is presented in Figure 4 [15-17]. The outcome implied that the KNN, SVM, ATI-CNN, CNN, LSTM, QLN, Deep CNN, and DQN methods have resulted in lower performance, whereas the AGCA-EMRA system has achieved a higher result with $accur_y$ of 95.97%, $preci_n$ of 96.51%, $recal_l$ of 95.97%, and $F1_{score}$ of 96.19%, correspondingly.

Table 4 shows the computational efficiency of the AGCA-EMRA technique with other models. Based on inference latency, the AGCA-EMRA method confers lesser value of 4.26ms, Flops of 1.37G and Memory footprint of 1.02 while the KNN, SVM, ATI-CNN, CNN, LSTM, QLN, Deep CNN, and DQN techniques attains greater time of 14.82ms, 12.47ms, 8.94ms, 7.86ms, 10.31ms, 6.94ms, 9.58 and 6.17ms, respectively.

Table 5 indicates the ablation evaluation of diverse components in the AGCA-EMRA model. The AGCA-EMRA (CMIM + GCN + GRU + MHA + Ranger) model achieved an accuracy of 94.92%, precision of 94.07%, recall of 94.92%, and F1-score of 94.24%.

Table 2. Confusion matrices and RUE analysis of AGCA-EMRA approach with 80:20.

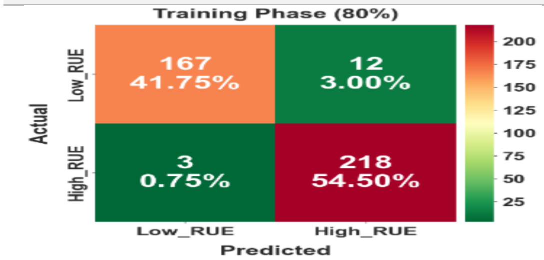
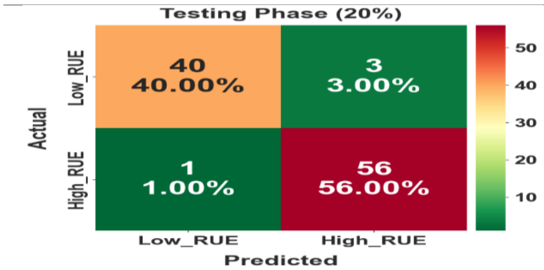
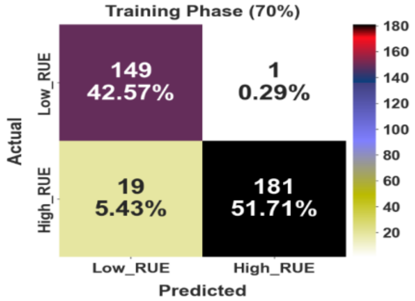
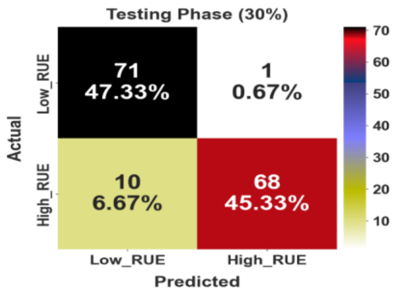
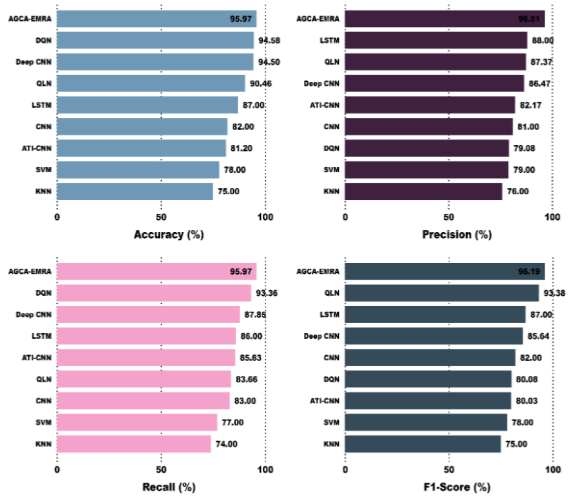
Confusion Matrix	Class Labels	$Accur_y$	$Preci_n$	$Recal_l$	$F1_{Score}$	MCC
	Low_RUE	93.30	98.24	93.30	95.70	92.48
	High_RUE	98.64	94.78	98.64	96.67	92.48
	Average	**95.97**	**96.51**	**95.97**	**96.19**	**92.48**
	Low_RUE	93.02	97.56	93.02	95.24	91.87
	High_RUE	98.25	94.92	98.25	96.55	91.87
	Average	**95.63**	**96.24**	**95.63**	**95.89**	**91.87**

Table 3. Confusion matrices and RUE analysis of AGCA-EMRA approach with 70:30.

Confusion Matrix	Class Labels	$Accur_y$	$Preci_n$	$Recal_l$	$F1_{Score}$	MCC
	Low_RUE	99.33	88.69	99.33	93.71	88.98
	High_RUE	90.50	99.45	90.50	94.76	88.98
	Average	**94.92**	**94.07**	**94.92**	**94.24**	**88.98**
	Low_RUE	98.61	87.65	98.61	92.81	86.00
	High_RUE	87.18	98.55	87.18	92.52	86.00
	Average	**92.90**	**93.10**	**92.90**	**92.66**	**86.00**

International Journal of Medical Toxicology & Forensic Medicine



International Journal of Medical Toxicology & Forensic Medicine

Figure 4. Comparative analysis of AGCA-EMRA approach with other systems.

Conclusion

This manuscript introduces an AGCA-EMRA model for a 6G-enabled IoT system. The goal of the proposed model is to precisely categorize resource utilization efficiency into low and high classes, enabling reliable monitoring and improved resource management. For classification, an MHSAGN captures correlations between spatial and temporal data. The technique is further optimized using the ranger optimizer to stabilize training and accelerate convergence. An extensive experimental analysis is conducted on the IoT-based MR

Table 4. Computational complexity of the proposed AGCA-EMRA model.

Model	Inference Latency (ms)	FLOPs (G)	Memory Footprint (GB)
KNN	14.82	0.95	1.18
SVM	12.47	0.82	1.06
ATI-CNN	8.94	3.87	2.43
CNN	7.86	3.24	2.11
LSTM	10.31	4.68	2.79
QLN	6.94	2.41	1.74
Deep CNN	9.58	5.36	3.08
DQN	6.17	2.08	1.53
AGCA-EMRA	4.26	1.37	1.02

International Journal of Medical Toxicology & Forensic Medicine

Table 5. Ablation Study of the Proposed AGCA-EMRA model.

Configuration	Accuracy	Precision	Recall	F1-Score
GCN + GRU	90.81	89.92	90.74	90.33
GCN + GRU + MHA	92.38	91.56	92.27	91.91
CMIM + GCN + GRU	93.24	92.46	93.13	92.79
CMIM + GCN + GRU + MHA	94.11	93.34	94.06	93.7
AGCA-EMRA (CMIM + GCN + GRU + MHA + Ranger)	94.92	94.07	94.92	94.24

International Journal of Medical Toxicology & Forensic Medicine

Allocation for the 6G Network Dataset. The proposed AGCA-EMRA model achieves 95.97% accuracy, outperforming existing methodologies. Although the proposed AGCA-EMRA demonstrates strong performance, it is currently evaluated on a single 6G-enabled IoT healthcare dataset, which limits its generalizability.

In the future, the research will focus on extending the proposed architecture with federated graph learning to enable privacy-preserving, collaborative resource allocation across distributed healthcare institutions. In such a decentralized learning model, healthcare organizations can jointly optimize and train the resource allocation model while keeping sensitive operational and patient data locally stored. The research will also emphasize validating the proposed architecture in heterogeneous, multi-center healthcare IoT environments characterized by diverse network frameworks, dynamic resource availability, and varied device capabilities.

Ethical Approval

None.

Acknowledgment

None.

Funding

No external funding was received for this research.

Conflicts of Interest

The authors report that there are no competing interests to declare.

References

[1] Xu Z, Ban F, Fotia P. Efficient QoS processing for internet of medical things using non-cooperative game theory: Resource allocation in cloud framework. Ann Oper Res. 2026;356(1):427-42. [DOI: 10.1007/s10479-024-05857-1]

[2] Anandaram H, Shreenidhi KS, Joshi K, Sidhu KS, Kiruthigha R, Singh V, et al. A case study on resource management industrial IoT application. In: Building Data-Driven Edge Systems for Business Success. IGI Global Scientific Publishing; 2026. p. 1-26. [DOI: 10.4018/979-8-3373-1147-0.ch001]

[3] Almadhor A, Ayari M, Alqahtani A, Al Hejaili A, Bouallegue B, Juanatas RA, et al. An AI and 6G-IoT enabled computational framework for intelligent medical resource allocation and adaptive personalized healthcare. Computing. 2025;107(12):234. [DOI: 10.1007/s00607-025-01571-3]

[4] Anand J, Karthikeyan B. Dynamic priority-based task scheduling and adaptive resource allocation algorithms for efficient edge computing in healthcare systems. Results Eng. 2025;25:104342. [DOI: 10.1016/j.rineng.2025.104342]

[5] Yaraziz MS, Hill R. A review of resource allocation for maximizing performance of IoT systems. IEEE Access. 2025;13:98426-51. [Link]

[6] Rafique W. ML-RASPF: A machine learning-based rate-adaptive framework for dynamic resource allocation in smart healthcare IoT. Algorithms. 2025;18(6):325. [DOI: 10.3390/a18060325]

[7] Gowda VD, Sharma A, Mukherjee S, Rahman M, Barshe SB. Optimizing resource allocation in healthcare facilities through IoT and machine learning predictive analysis. Lect Notes Netw Syst. 2024;1007:177-87. [DOI: 10.1007/978-981-97-5146-4_16]

[8] Lv J, Chen CM, Kumari S, Li K. Resource allocation for AI-native healthcare systems in 6G dense networks using deep reinforcement learning. Digit Commun Netw. 2025;11(6):2016-29. [DOI: 10.1016/j.dcan.2025.06.011]

[9] Muazu T, Yingchi M, Muhammad AU, Ibrahim M, Samuel O, Tiwari P. IoMT: A medical resource management system using edge empowered blockchain federated learning. IEEE Trans Netw Serv Manag. 2023;21(1):517-34. [DOI: 10.1109/TNSM.2023.3308331]

[10] Ahamed B, Sellamuthu S, Karri PN, Srinivas IV, Zabeeulla AM, Kumar MA. Design of an energy-efficient IoT device-assisted wearable sensor platform for healthcare data management. Meas Sens. 2023;30:100928. [DOI: 10.1016/j.measen.2023.100928]

[11] Morán-Fernández L, Blanco-Mallo E, Sechidis K, Bolón-Canedo V. Breaking boundaries: Low-precision conditional mutual information for efficient feature selection. Pattern Recognit. 2025;162:111375. [DOI: 10.1016/j.patcog.2025.111375]

[12] Oluwasanmi A, Aftab MU, Qin Z, Sarfraz MS, Yu Y, Rauf HT. Multi-head spatiotemporal attention graph convolutional network for traffic prediction. Sensors. 2023;23(8):3836. [DOI: 10.3390/s23083836]

[13] Xie F, Wang L, Zhu H, Xie S. Research on a rolling bearing fault diagnosis method based on multi-source deep sub-domain adaptation. Appl Sci. 2023;13(11):6800. [DOI: 10.3390/app13116800]

[14] Ziya07. IoT-driven medical resource allocation dataset [Internet]. Kaggle; [cited 2026 Aug 27]. [Link]

[15] Alharbe N, Almalki M. IoT-enabled healthcare transformation leveraging deep learning for advanced patient monitoring and diagnosis. Multimed Tools Appl. 2025;84(19):21331-44. [DOI: 10.1007/s11042-024-19919-w]

[16] Islam MR, Kabir MM, Mridha MF, Alfarhood S, Safran M, Che D. Deep learning-based IoT system for remote monitoring and early detection of health issues in real-time. Sensors. 2023;23(11):5204. [DOI: 10.3390/s23115204]

[17] Roy PP, Teju V, Kandula SR, Sowmya KV, Stan AI, Stan OP. Secure healthcare model using multi-step deep Q learning network in internet of things. Electronics. 2024;13(3):669. [DOI: 10.3390/electronics13030669]