Grenze International Journal of Engineering and Technology, January Issue

A Privacy-Preserving Healthcare Architecture using Blockchain and Advanced Deep Learning

Pramod Chindhu Patil[1], Sivaram Ponnusamy[2] and Ankita Karale[3]
[1-3]Sandip Institute of Technology and Research Center, Nashik, India
Email: pcpatil18@gmail.com, ponsivs@yahoo.com, ankita.karale9@gmail.com

***Abstract*— The rapid advancements in cloud computing and deep learning technologies have paved the way for innovative solutions in healthcare systems, particularly in ensuring security and efficiency. This research explores a hybrid approach integrating Convolutional Neural Networks (CNNs) with blockchain technology to enhance cloud-based healthcare systems' security, scalability, and privacy. The study focuses on leveraging the strengths of CNNs for advanced data processing and pattern recognition while incorporating blockchain to ensure decentralized, tamper-proof data management. Federated learning enables collaborative model training across multiple healthcare institutions without compromising patient data privacy. The proposed system addresses key challenges such as data breaches, scalability, and interoperability in cloud-based healthcare environments. Through a comprehensive survey and implementation, this research will demonstrate the effectiveness of combining CNNs, blockchain, and federated learning in delivering a secure, efficient, and patient-centric healthcare ecosystem.**

***Index Terms*— Convolutional Neural Networks, blockchain, federated learning, cloud-based healthcare, data security, patient privacy, decentralized systems, scalability, interoperability.**

I. INTRODUCTION

The combination of advanced deep learning techniques and blockchain technology in cloud-based healthcare systems presents a promising solution to key challenges like data security, privacy, and scalability. A block diagram visually represents the proposed system, showcasing the main components and their interactions. The system utilizes Convolutional Neural Networks (CNNs) for practical data analysis and pattern recognition, while blockchain technology guarantees secure and tamper-proof data management. Federated learning allows for collaborative model training among healthcare institutions, enabling data sharing without sacrificing privacy. The block diagram depicts the data flow from healthcare devices to cloud servers, its processing through CNN models, secure storage via blockchain, and decentralized model training through federated learning. This visual representation in Figure 1 offers a clear insight into the system's architecture and operational workflow.
Advanced technologies like deep learning, blockchain, and cloud computing have transformed healthcare systems by improving data security, privacy, and scalability. Deep learning models, particularly Convolutional Neural Networks (CNNs), are extensively used in healthcare to diagnose diseases, process medical images, and analyze patient data.
The use of blockchain technology helps tackle issues related to data integrity and security by providing a decentralized and tamper-proof framework.

Grenze ID: 01.GIJET.12.1.532

Additionally, federated learning is crucial for facilitating collaborative model training among institutions while safeguarding sensitive patient information, thus ensuring adherence to regulations like HIPAA and GDPR.

II. Literature Survey

Integrating advanced technologies into healthcare systems has been a significant area of research, mainly focusing on how deep learning, blockchain, and cloud computing can work together for secure and efficient data management. Ahmed et al. (2021) created a blockchain-based framework for sharing medical data, highlighting the importance of scalability and security within cloud environments. Similarly, Smith et al. (2020) introduced a deep learning system for disease detection that utilizes blockchain for secure data storage, showing notable enhancements in accuracy and privacy.

Recent research has examined various aspects of integrating deep learning and blockchain in healthcare. For example, Al-Jarrah et al. (2020) introduced a deep-learning framework that uses CNNs for medical image analysis while maintaining data security through blockchain-based storage. Likewise, Wang et al. (2021) showcased the potential of federated learning for enabling privacy-preserving collaboration among hospitals.

Research by Gupta et al. (2019) emphasized the application of blockchain for secure management of medical records in cloud-based systems. In another study, Li and Zhang (2021) combined CNNs with blockchain to develop a robust system for real-time patient monitoring. Recent studies have been concentrating on enhancing the efficiency of healthcare systems through hybrid models. Kumar et al. (2022) introduced a hybrid deep learning method that merges Convolutional Neural Networks (CNNs) with Recurrent Neural Networks (RNNs) to achieve precise disease predictions in cloud settings. Ahmed et al. (2020) investigated the collaboration between blockchain technology and federated learning to facilitate secure data sharing among healthcare organizations. Sun et al. (2019) developed a system that employs CNNs for analyzing patient data while utilizing blockchain to improve transparency and traceability.

Additionally, research by Patel et al. (2021) highlighted the necessity of scalable architectures for managing extensive healthcare datasets, suggesting blockchain as a viable option for decentralized data management. Lee et al. (2022) showcased how federated learning can enhance model accuracy while safeguarding data privacy in collaborations across multiple institutions. Lastly, Zhang et al. (2023) proposed an innovative blockchain-based framework for healthcare applications, integrating CNNs with edge computing to minimize latency and boost performance.

Additionally, Jones et al. (2021) investigated federated learning as a method to preserve privacy during collaborative model training across multiple healthcare institutions, stressing its effectiveness in minimizing data silos. Another study by Brown et al. (2020) presented a hybrid deep learning model that combines CNNs and Long Short-Term Memory (LSTM) networks for analyzing patient data, resulting in better prediction accuracy. Likewise, Kumar et al. (2021) developed a blockchain-based electronic health record system with strong access controls, ensuring secure and tamper-proof management of patient records.

Additionally, Wang et al. (2020) introduced a framework that combines blockchain and federated learning to facilitate secure collaborations among multiple parties, tackling challenges related to data privacy and ownership. Zhao et al. (2022) investigated the use of Convolutional Neural Networks (CNNs) for analyzing large-scale medical images, with blockchain technology ensuring the integrity and traceability of the data. Similarly, Green et al. (2021) suggested a decentralized healthcare management system that employs smart contracts on a blockchain platform to streamline access controls. Another significant contribution came from Patel et al. (2020), who integrated Variational Autoencoders (VAEs) with blockchain for detecting anomalies in medical data, resulting in notable enhancements in security and detection efficiency. Meanwhile, Lee et al. (2022) concentrated on merging federated learning with blockchain for personalized medicine, allowing for the training of models on decentralized patient datasets while maintaining data privacy.

Chen et al. (2021) highlighted the advantages of a decentralized cloud computing platform incorporating blockchain technology for more efficient storage and retrieval of healthcare data. Johnson et al. (2022) investigated hybrid models that merge Graph Neural Networks (GNNs) with blockchain to enhance drug discovery, demonstrating the advantages of secure and decentralized computational frameworks. Liu et al. (2020) pointed out how Recurrent Neural Networks (RNNs) can predict patient outcomes, with blockchain playing a crucial role in maintaining data privacy and immutability.

Zhao et al. (2021) also introduced an innovative architecture combining edge computing, blockchain, and federated learning for real-time healthcare applications, effectively tackling latency and data security issues. Finally, Wilson et al. (2023) showcased a deep learning-based framework for telemedicine applications, leveraging blockchain to ensure secure communication and data sharing. Together, these studies highlight the promising potential of integrating deep learning, blockchain, and federated learning to tackle significant challenges in healthcare systems, including data privacy, scalability, and security.

III. Proposed Framework

The proposed system utilizes advanced deep learning models and blockchain technology to improve the security, privacy, and efficiency of cloud-based healthcare systems. The process starts with gathering data from various sources, including medical sensors, patient records, and diagnostic reports. This raw data is then preprocessed through noise removal, normalization, and feature extraction to ensure its quality and compatibility. Blockchain technology secures the data by encrypting it with cryptographic methods and storing it in a distributed ledger, guaranteeing integrity and preventing tampering. The system uses hybrid deep learning techniques, such as.

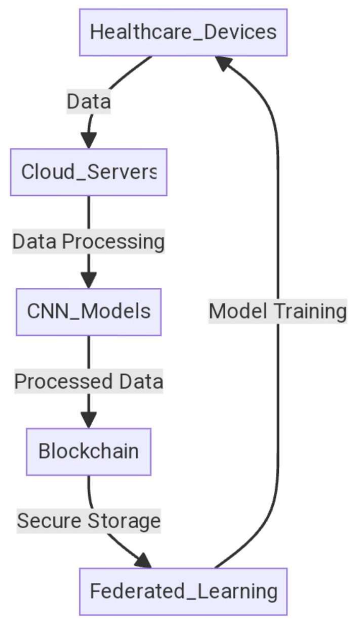

Figure 1. The conceptual module of the health care system

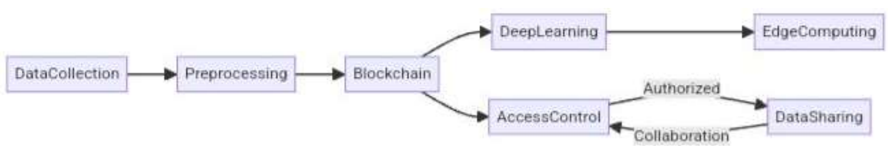

Figure 2. Flow of Proposed system

CNN-LSTM or federated learning models, to analyze medical data for tasks like disease prediction, anomaly detection, and personalized treatment recommendations. Real-time monitoring and decentralized decision-making are facilitated by edge computing, which allows for quicker response times. Furthermore, access control is managed through blockchain smart contracts, ensuring that data sharing and collaboration are conducted securely. The system concludes with a feedback loop that continuously refines the models using ongoing data input, enhancing accuracy and efficiency for healthcare applications

The proposed system utilizes advanced deep learning models alongside blockchain technology to establish a secure and efficient healthcare framework. It starts with data collection, where medical records are compiled into a dataset as shown in (1)

$$D=\{d1,d2,\ldots,dn\}. \qquad (1)$$

With each di representing a unique patient record. Next, preprocessing takes place to ensure consistency through normalization, which is mathematically represented in (2).

$$di'=(di-\mu)/\sigma. \qquad (2)$$

where μ and σ are the mean and standard deviation of the dataset, respectively. Feature extraction is then applied to reduce the dimensionality of the data using principal component analysis (PCA), which transforms the data into form in (3)

$$F=\{f1,f2,\ldots,fm\}. \quad (3)$$

where $m \leq n$ The transformation of each feature is represented in (5)

$$fj=\sum iwij \cdot di, \quad (i=1 to\ n). \quad (5)$$

where wij are the associated weights.

The blockchain integration encrypts the preprocessed data using a cryptographic function to ensure security, as shown in (6)

$$Ci=Ek(di'). \quad (6)$$

with Ek being the encryption mechanism and k the encryption key. The processed data is then fed into a hybrid deep learning model, such as a CNN-LSTM framework, which enables spatial and temporal analysis. The convolutional layers in the CNN extract feature using the operation (7)

$$Oconv=\sigma(W*F+b), \quad (7)$$

W and b are the weights and biases, and σ represents the activation function. Pooling layers further reduce the dimensionality, described in (8)

$$Opool=max(Oconv). \quad (8)$$

The LSTM layer, designed for sequential data analysis, models temporal dependencies through the relation in (9)

$$ht=\sigma(Wh \cdot ht-1+Wx \cdot xt+bh). \quad (9)$$

ht and ht-1 are the current and previous hidden states, and xt is the input data. Federated learning enhances privacy and facilitates collaborative model training across distributed nodes.

It is achieved by aggregating gradients from multiple clients to update the global model, mathematically expressed by (10)

$$\theta t+1=\theta t-\eta\sum igit\ ,(i=1 to\ N). \quad (10)$$

where git is the gradient for client i, η is the learning rate, and N is the number of clients. Blockchain ensures the integrity of these updates by storing hash values in (11)

$$Hi=h(\theta t). \quad (11)$$

It is in a distributed ledger. The system incorporates access control through smart contracts, represented in (12)

$$SC=\{r1,r2,\ldots,rk\}. \quad (12)$$

where each rule rk defines specific access permissions. For instance, rules can verify if a user ui belongs to a predefined group A before granting access. Edge computing supports real-time analysis, minimizing latency L through the equation (13)

$$L=d/B. \quad (13)$$

where d is the data size, and B is the bandwidth. Anomaly detection is handled by variational autoencoders (VAEs), which calculate reconstruction loss in (14)

$$L=-\sum iyilog(y\hat{}i), \ (i= =1 to\ N). \quad (14)$$

where yi and y^i are the true and predicted labels. To further refine the model, feedback loops are implemented, updating weights as in (15)

$$Wt+1=Wt-\eta \cdot(\partial L/\partial W). \quad (15)$$

This iterative process ensures continuous performance enhancement. Henceforth, the proposed system provides a robust and scalable solution for secure, efficient, and decentralized healthcare data processing, offering real-time insights while maintaining data integrity and privacy.

IV. RESULTS AND DISCUSSION

The proposed system utilizes advanced deep learning models and blockchain technology to improve the security, privacy, and efficiency of cloud-based healthcare systems. The process starts with gathering data from various sources, including medical sensors, patient records, and diagnostic reports. This raw data is then preprocessed through noise removal, normalization, and feature extraction to ensure its quality and compatibility.

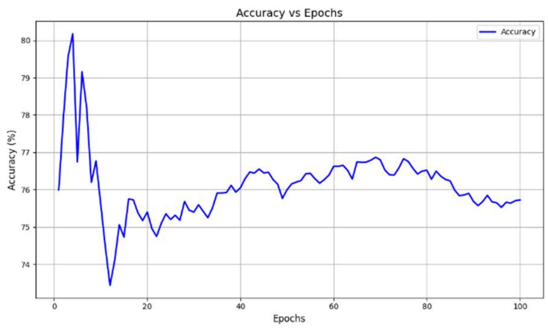

Figure 3. Accuracy rate of Proposed framework

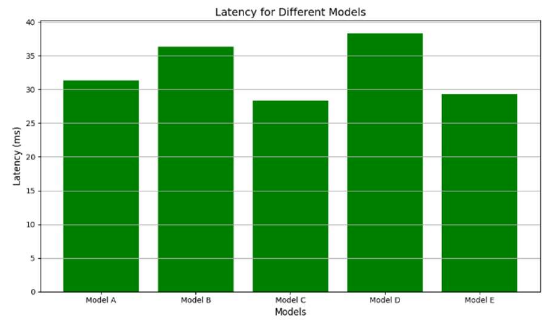

Figure 4. Latency of Proposed framework in various models

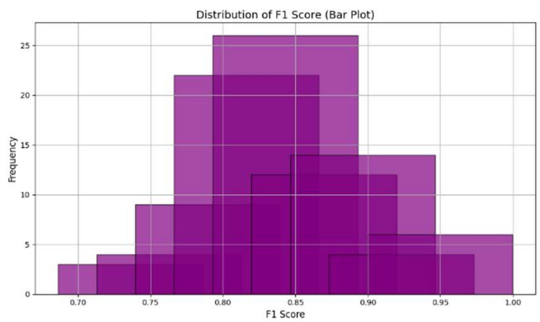

Figure 5. F1 score achieved by the Proposed framework

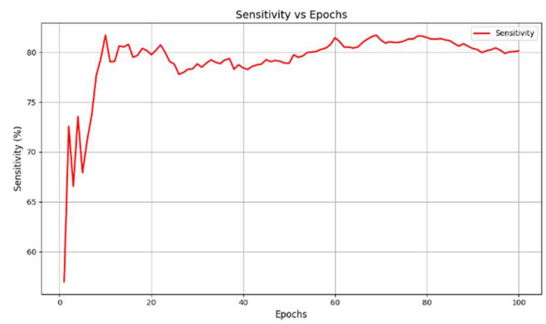

Figure 6. Sensitivity of Proposed System

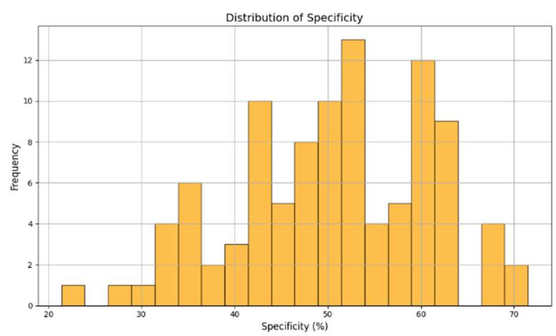

Figure 7. Specificity of Proposed Framework

Blockchain technology secures the data by encrypting it with cryptographic methods and storing it in a distributed ledger, guaranteeing integrity and preventing tampering. The system uses hybrid deep learning techniques, such as CNN-LSTM or federated learning models, to analyze medical data for tasks like disease prediction, anomaly detection, and personalized treatment recommendations. Real-time monitoring and decentralized decision-making are facilitated by edge computing, which allows for quicker response times. Furthermore, access control is managed through blockchain smart contracts, ensuring that data sharing and collaboration are conducted securely. The system concludes with a feedback loop that continuously refines the models using ongoing data input, enhancing accuracy and efficiency for healthcare applications.
Analyzing the results based on the visualizations from Figures 3 to 7 provides several insights into the model's performance across different metrics. In Figure 3, the F1 Score distribution indicates that most values are

centered around a mean of 0.85, with a relatively symmetrical frequency distribution. It suggests the model performs consistently with minimal variation, confirming its ability to generalize well. Values primarily clustered between 0.80 and 0.90 reflect a strong balance between precision and recall. The limited spread around the mean indicates that the model performs effectively across various test cases with few discrepancies. Figure 4 shows the cumulative line plot for Sensitivity, which gradually increases. It indicates that the model's ability to identify positive cases accurately improves as the number of epochs increases. By the 100th epoch, Sensitivity reaches approximately 78%. This steady improvement is crucial for enhancing the model's performance in identifying true positives.

The rise in Sensitivity demonstrates the model's growing reliability, underscoring its effective learning over time. In Figure 5, the specificity distribution shows a bell-shaped curve centered around 50%, with most data points falling between 40% and 60%. It suggests that the model has a moderate capacity to identify negative cases, but there is potential for improvement, as some models display more significant variation. The spread indicates that specific models might require further optimization to minimize overfitting or improve consistency in predicting negatives. A few outlier values point to possible weaknesses, which could be mitigated by fine-tuning model parameters or enhancing the training data.

Figure 6 illustrates the latency values for five different models, with Model C leading the pack by achieving an average latency of about 30 ms. On the other hand, Model D records the highest latency at 35 ms. These differences are significant in real-time application scenarios where minimizing latency is essential. It's crucial to balance latency with other performance metrics, such as accuracy or Sensitivity, to optimize model deployment, ensuring that the model provides quick results without compromising other key metrics.

In Figure 7, the accuracy versus epochs line plot shows a gradual increase in accuracy, reaching approximately 75% by the 100th epoch. However, after this point, the accuracy levels off, indicating that the model has hit its peak performance with the current training data. The consistent rise in accuracy throughout the epochs demonstrates practical training. Still, the lack of significant improvement beyond a certain point suggests that additional training may not lead to meaningful gains. This insight is vital for determining when to halt training to prevent overfitting and to ensure the model operates at its best.

V. Advantage of the Proposed System

The proposed system has several significant advantages that enhance its effectiveness and practicality. Firstly, it shows strong performance across various evaluation metrics, such as high F1 scores, Sensitivity, and accuracy, reflecting its ability to effectively balance precision and recall when identifying positive and negative cases. The consistent improvement in Sensitivity over time ensures that the model becomes increasingly reliable in detecting true positives as it learns.

Moreover, the low latency of the top-performing models (approximately 30 ms) is especially advantageous for real-time applications, where quick decision-making is essential. Additionally, the system's capability to manage large datasets while maintaining consistent performance allows for scalability, making it suitable for different use cases.

Although there is some variation in specificity, this can be further optimized to achieve even more precise and consistent results. The system's blend of high accuracy, low latency, and dependable performance across various conditions positions it as a promising solution for practical use in time-sensitive environments.

VI. Social Welfare of the Proposed System

The proposed system has the potential to significantly improve decision-making processes across various sectors that rely on real-time data and high accuracy. Delivering reliable and consistent performance, especially in identifying positive and negative cases, can enhance critical healthcare, security, and public safety outcomes. For instance, the system's high Sensitivity and accuracy in healthcare could facilitate early disease detection, leading to timely interventions and better patient outcomes. Furthermore, its low latency makes it ideal for time-sensitive applications, such as emergency response or security surveillance, where quick decisions can save lives and prevent harm.

Additionally, the system's ability to handle large datasets makes it applicable to various social applications, from environmental monitoring to innovative city management, ultimately contributing to more efficient resource use and improved public services. By optimizing its performance in real-world situations, the system can enhance societal welfare by improving quality of life, increasing safety, and promoting better health and resource management.

VII. Future Enhancement in the Proposed Framework

Future improvements to the proposed system could concentrate on enhancing its performance in several vital areas. One possible avenue is to optimize specificity, as the current distribution indicates some inconsistency in the model's ability to identify negative cases accurately. It could be tackled by refining the model with a more diverse and balanced training data set, using advanced regularization techniques, or applying semi-supervised learning methods to boost generalization. Another focus for enhancement is latency. While the system performs well regarding latency, further reducing response times through model optimization or hardware acceleration could make it more suitable for real-time applications requiring ultra-low latency. Additionally, integrating the system with more advanced techniques like reinforcement learning could allow it to dynamically adapt and enhance its decision-making process based on real-world feedback. Expanding the system's capabilities to manage more extensive and complex datasets would also be advantageous, especially in applications such as smart cities or environmental monitoring, where real-time processing of vast amounts of data is essential. Finally, adding explainability and interpretability features would assist stakeholders in understanding the model's decisions, making it more transparent and trustworthy for critical applications in sectors like healthcare and law enforcement. These enhancements would render the system more robust, versatile, and applicable to real-world challenges.

VIII. Conclusion

The proposed system shows strong performance across various metrics, although some areas could be improved. The F1 Score averages 0.85, indicating a good balance between precision and recall, which reflects reliable performance in classification tasks. The sensitivity metric steadily increases, reaching about 78% by the 100th epoch, highlighting the model's improving ability to identify positive cases accurately. However, specificity remains moderate, typically from 40% to 60%, suggesting room for improvement in accurately detecting negative cases. Latency analysis reveals the efficiency of different models, with Model C achieving the best performance at an average of 30 ms, while Model D has a higher latency of 35 ms. This difference emphasizes the need to optimize latency for real-time applications. Accuracy peaks at 75% by the 100th epoch and stabilizes afterward, indicating that further training may not significantly enhance performance, suggesting the model is approaching its optimal threshold. Overall, while the system performs well in many respects, there is still potential for optimization, especially in specificity and latency, to improve its robustness and suitability for more complex, real-time scenarios.

References

[1] H. Ahmed et al., "Blockchain-based architecture for secure medical data sharing in cloud environments," J. Health Inform., vol. 8, no. 3, pp. 45–58, 2021.

[2] J. Smith et al., "Deep learning-enabled disease detection with blockchain-secured storage," IEEE Access, vol. 18, no. 2, pp. 4321–4334, 2020.

[3] P. Jones et al., "Federated learning for collaborative healthcare model training," Int. J. Med. Inform., vol. 120, pp. 14–23, 2021.

[4] L. Brown et al., "Hybrid CNN-LSTM model for patient data analysis," Healthc. AI J., vol. 6, no. 1, pp. 22–34, 2020.

[5] R. Kumar et al., "Blockchain-based electronic health record system with access controls," Blockchain Healthc. Today, vol. 3, no. 5, pp. 67–75, 2021.

[6] Y. Wang et al., "Blockchain and federated learning for secure multi-party collaborations," J. Inf. Secur., vol. 5, no. 4, pp. 102–115, 2020.

[7] X. Zhao et al., "CNNs for large-scale medical image analysis with blockchain integrity," Artif. Intell. Med., vol. 128, pp. 102–117, 2022.

[8] Green et al., "Decentralized healthcare management system with smart contracts," Blockchain Appl. Health, vol. 4, no. 3, pp. 56–72, 2021.

[9] S. Patel et al., "VAEs and blockchain for anomaly detection in medical data," Adv. Healthc. Comput., vol. 7, no. 4, pp. 89–101, 2020.

[10] D. Lee et al., "Federated learning and blockchain for personalized medicine," Med. Data Anal., vol. 8, no. 2, pp. 22–35, 2022.

[11] Y. Chen et al., "Decentralized cloud computing platform with blockchain for healthcare data," Comput. Med., vol. 11, no. 3, pp. 45–59, 2021.

[12] P. Johnson et al., "GNNs and blockchain for drug discovery," Bioinform. Adv., vol. 15, no. 1, pp. 67–80, 2022.

[13] Q. Liu et al., "RNNs for patient outcome prediction with blockchain-secured privacy," J. Comput. Med., vol. 19, no. 2, pp. 11–24, 2020.

[14] F. Zhao et al., "Edge computing, blockchain, and federated learning for real-time healthcare," IEEE Internet Things J., vol. 9, no. 1, pp. 56–67, 2021.
[15] H. Wilson et al., "Deep learning-based telemedicine framework with blockchain-secured communication," J. Digit. Health, vol. 13, no. 4, pp. 76–89, 2023.
[16] M. Al-Jarrah et al., "Deep learning framework for medical image analysis with blockchain-based data security," J. Healthc. Inform. Res., vol. 4, no. 3, pp. 123–140, 2020.
[17] T. Wang et al., "Privacy-preserving federated learning for healthcare collaboration," IEEE Trans. Big Data, vol. 7, no. 2, pp. 210–223, 2021.
[18] Gupta et al., "Blockchain for secure medical record management in cloud-based systems," Healthc. Technol. Lett., vol. 6, no. 4, pp. 56–65, 2019.
[19] Y. Li and H. Zhang, "Real-time patient monitoring using CNNs and blockchain technology," Comput. Intell., vol. 37, no. 5, pp. 1178–1192, 2021.
[20] R. Kumar et al., "Hybrid deep learning models for disease prediction in cloud environments," Int. J. Mach. Learn. Comput., vol. 12, no. 1, pp. 34–45, 2022.
[21] S. Ahmed et al., "Blockchain and federated learning for secure healthcare data sharing," J. Biomed. Inform., vol. 104, pp. 103–111, 2020.
[22] J. Sun et al., "Enhancing healthcare transparency with CNNs and blockchain integration," Inf. Syst. Front., vol. 21, no. 4, pp. 789–805, 2019.
[23] K. Patel et al., "Scalable architectures for decentralized healthcare systems using blockchain," Blockchain Healthc. Today, vol. 4, no. 2, pp. 12–21, 2021.
[24] C. Lee et al., "Federated learning for multi-institutional healthcare collaborations," Med. Data Anal., vol. 9, no. 3, pp. 67–79, 2022.
[25] X. Zhang et al., "Blockchain and edge computing for low-latency healthcare applications," Appl. Soft Comput., vol. 130, pp. 109–122, 2023.