

Article

Optimized Deep learning Frameworks for the Medical Image Transmission in IoMT Environment

Rashmi P¹, R. Gomathi²

¹ Inventuriz Labs Pvt Ltd, Bengaluru, India

² Department of Electronics and Communication Engineering, St.Peter's Institute of Higher Education and Research, Avadi, Chennai, India

*Corresponding author: Rashmi P (email: rashmibdp@gmail.com)

Received 20 September 2024; Accepted 28 October 2024; Published 15 December 2024

Abstract: The Internet of Medical Things (IoMT) is reforming healthcare by enabling interconnected medical devices and systems to facilitate efficient data collection, transmission, and analysis. While IoMT has significantly improved real-time monitoring and personalized care, the transmission of high-resolution medical images remains a challenge due to bandwidth constraints, latency issues, data loss, and computational overhead. Efficient and secure medical image transmission is critical to ensuring reliable diagnostics and timely patient care in this ecosystem. This research presents an optimized Deep Learning (DL) architecture developed to overcome the limitations of medical image transmission in IoMT environments. The proposed solution incorporates Convolutional Neural Networks (CNNs) for spatial feature extraction and dimensionality reduction while preserving diagnostic-critical information, and Long Short-Term Memory (LSTM) networks to manage sequential data and mitigate transmission issues such as packet loss and latency. The framework incorporates robust encryption mechanisms to ensure data security without significantly increasing computational overhead. Once predictions are made, the data is securely transferred to the cloud for further analysis and storage. Furthermore, Hippopotamus Optimization is utilised to enhance the model's performance and fine-tune hyperparameters, improving both efficiency and accuracy. Performance evaluations were conducted using real-world medical image datasets under varying IoMT network conditions. The results demonstrate that the proposed CNN-LSTM framework delivers superior performance across key metrics, like Peak Signal-to-Noise Ratio (PSNR), accuracy, F1 score, specificity, and sensitivity. Additionally, the framework optimizes encryption and decryption times and reduces bandwidth consumption, ensuring efficient and secure data transmission. This approach showcases a significant advancement in IoMT-based medical imaging, paving the way for enhanced reliability and efficiency in healthcare delivery systems.

Keywords: IoMT, medical image transmission, CNN, LSTM, Hippopotamus Optimization, Accuracy, Bandwidth optimization, Healthcare data.

Copyright © 2024 Journal of Smart Internet of Things (JSIoT) published by Future Science for Digital Publishing and Sciendo . This is an open access article license CC BY (<https://creativecommons.org/licenses/by/4.0/>)

1. Introduction

IoMT is revolutionizing healthcare by incorporating smart devices and systems, enabling seamless data collection, transmission, and analysis. However, challenges such as bandwidth constraints, data loss, latency, and security concerns hinder efficient transmission of medical data, particularly CT lung images, necessitating innovative solutions [1]. This research introduces an optimized framework using CNN and LSTM models to address these limitations.

Current methods for transmitting CT lung images in IoMT environments face several challenges related to data compression, encryption, and network performance. Lossless compression techniques, while preserving image quality, result in large file sizes, which increase transmission time and place a significant burden on storage systems [2]. On the other hand, lossy compression methods, which achieve smaller file sizes by discarding data, compromise diagnostic accuracy and are unsuitable for medical applications where precision is critical [3]. Encryption methods in current systems, while securing data, often incur computational overhead that introduces latency, especially when dealing with large image files or resource-constrained devices like wearable sensors [4]. Furthermore, IoMT networks frequently experience unstable connectivity, leading to data loss, delays, and increased latency, which are detrimental to time-sensitive medical applications. These challenges highlight the need for optimized frameworks that balance high image quality, secure data transmission, and efficient resource utilization to ensure fast and reliable CT lung image transmission [5].

This research addresses the limitations of existing methods by combining CNNs and LSTMs for efficient and secure transmission of CT lung images. CNNs are highly effective in extracting spatial features and reducing dimensionality, while LSTMs best in modeling sequential dependencies, establishing them as suitable for predictive optimization in data transmission [6]. The proposed model aims to enhance transmission efficiency, ensure reliable image reconstruction, and reduce computational overhead, addressing critical gaps in IoMT-based CT lung image transmission systems[7-9]. The key contributions of this research include:

- Development of an optimized model to improve bandwidth utilization and reduce transmission time.
- Assurance of high-quality image reconstruction, evaluated through metrics like PSNR and accuracy.
- Integration of lightweight encryption mechanisms to enhance data security with minimal computational impact.
- Comprehensive evaluation across various metrics such as specificity, sensitivity, encryption time, decryption time, and bandwidth optimization.

The study is organised in the following manner: Section 2 discusses relevant studies and challenges of traditional systems. Section 3 details the recommended CNN-LSTM framework. Section 4 describes the experimental setup and performance measures. Section 5 presents the findings and discussion, concluding with future enhancements.

2. Related Work

Balhareth et al. (2024) [10] introduced an optimized intrusion detection system (IDS) for the

IoMT networks, aiming to improve the detection accuracy of malicious activities in medical device networks. Their system utilizes tree-based Machine Learning (ML) classifiers integrated with filter-based feature selection methods, including Mutual Information (MI) and XGBoost, to improve performance while reducing computational costs. The IDS monitors unauthorized activities during the transmission of healthcare data and attains 98.79% accuracy with a low false alarm rate of 0.007 FAR on the CICIDS2017 dataset. However, the research focuses solely on binary classification for intrusion identification, with plans for future work to implement a multi-classification approach for attack detection and categorization. Additionally, the proposed feature selection technique needs further empirical evaluation in real-world IoMT scenarios to fully assess its applicability across various machine learning classifiers.

Rahman et al. (2024)[11] conducted a comprehensive survey on the utilisation of ML and DL methods in smart healthcare. The study explored the recent advancements, applications, challenges, and future prospects of these technologies in the healthcare sector. The authors highlighted the significant role of ML and DL in addressing various healthcare challenges, like disease prediction, drug discovery, and medical image analysis. They emphasized the integration of ML and DL techniques to enhance healthcare systems, offering a detailed review of their applications, like ML-healthcare, DL-healthcare, and ML-DL-healthcare. Despite the promising potential, the authors pointed out the challenges related to data privacy, model interpretability, and the demand for large, high-quality datasets.

Nandagopal et al. (2024)[12] constructed a Deep Auto-Optimized Collaborative Learning (DACL) model to identify chronic diseases using Artificial Intelligence (AI) and IoMT networks. This AI-IoMT framework is designed to enhance disease diagnosis by leveraging interconnected sensors in healthcare environments. The DACL model uses a Deep Auto-Encoder Model (DAEM) to preprocess and impute missing data, while the Golden Flower Search (GFS) method is employed to choose optimal features for classification. Furthermore, the Collaborative Bias Integrated GAN (ColBGaN) model is used to classify chronic diseases such as heart disease, diabetes, and stroke from patient medical records. The Water Drop Optimization (WDO) method is applied to minimize classifier error. The effectiveness of the DACL model was demonstrated using various benchmarking datasets, showing superior outcomes in terms of accuracy and efficiency. However, the model's reliance on multiple optimization techniques and complex preprocessing steps may increase computational costs and hinder its scalability in real-time applications.

Datta Gupta et al. (2023)[13] introduced a novel lightweight DL-based approach, "ReducedFireNet," for histopathological image classification in the context of the Internet of Medical Things (IoMT). This model addresses the challenge of timely disease diagnosis, particularly for life-threatening conditions like cancer, by providing real-time, auto-analysis of histopathological images. The study highlights the increasing importance of IoMT in healthcare, offering the potential for effective disease identification and treatment. The proposed model achieved a mean accuracy of 96.88% and an F1 score of 0.968 when evaluated on a real histopathological dataset, showcasing impressive performance in terms of accuracy. Moreover, the model's lightweight design, with a size of only 0.391 MB and a computational requirement of 0.201 GFLOPS, makes it ideal for integration into IoT-based imaging devices. However, despite its promising results, the model's application is limited to histopathological images, and its scalability to other medical imaging domains remains unaddressed, which could be a potential area for further exploration.

Lata, K.; Cenkeramaddi, L.R. (2023)[14] conducted a comprehensive review on DL for medical image cryptography, highlighting the significance of securing electronic health records (EHRs) within the healthcare sector's IOMT systems. They emphasize the increasing complexity of ensuring the privacy, integrity, and availability of EHRs due to the ongoing digital transformation. The paper discusses several imaging modalities such as PET, MRI,

ultrasonography, CT, and X-ray, which are crucial for medical diagnosis. These images are processed for tasks like segmentation, feature selection, and denoising, while cryptography techniques are explored as solutions to secure sensitive medical data during storage and transformation. However, the challenges and limitations in adopting DL for medical image cryptography, like computational complexity, the need for large datasets, and the potential trade-offs between security and performance, are also noted as areas for further research.

Thimmapuram et al. (2023)[15] developed a DL-based approach for medical image classification utilising feature extraction techniques in IoMT. The aim of the study was to utilize Artificial Intelligence (AI) and DL to assess both real-time and historical data from IoT devices, particularly for predicting and identifying serious diseases in the early stages of diagnosis. Medical image classification acts as a significant part in the diagnostic process, but it faces limitations due to the complexity of classifying images based on efficient features. The authors tested their proposed model using MATLAB and achieved impressive results, with 97.71% accuracy on a brain dataset and 97.2% accuracy on an Alzheimer's disease dataset. Despite the high performance of the algorithm, the study highlights the need for further improvement in terms of accuracy, precision, and computational speed, especially for clinical applications, and proposes the use of soft sets for better convergence in classification tasks. However, the model may still face limitations in real-time deployment due to computational efficiency and the need for more robust feature extraction methods.

Li et al. (2023) [16] provide a comprehensive review on the application of DL techniques in medical image analysis, focusing on their potential to revolutionize healthcare outcomes through real-time analysis of complex datasets. The study categorizes recent DL approaches into five major techniques: CNN, Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), LSTM and hybrid models. The authors critically assess these techniques based on several parameters, including their principles, advantages, limitations, and the datasets used in their implementation. Python emerged as the most commonly used programming language for these methods, with most of the reviewed papers published in 2021, highlighting the recent surge in research. Despite the progress, the review points out significant challenges in the widespread adoption of DL in medical image analysis, including computational complexity, the need for large, high-quality datasets, and issues with generalizability across diverse clinical settings. The evaluation metrics across the studies include accuracy, sensitivity, specificity, F-score, and robustness, but limitations related to the practical implementation and scalability of these methods in real-world healthcare environments remain a major concern.

Tiwari et al. (2022)[17] introduce a brownfield IoMT network for imaging data, designed for scalability based on the objectives, functional requirements, and the number of connected facilities and devices. They developed using DenseNet-201 architecture to generate image descriptors. For classification, an optimized Deep Neural Network model is employed to attain best performance. The model is validated on three publicly available datasets—Brain Tumor MRI, Covid-19 Radiography, and Breast Cancer MRI—and achieves impressive results, with an accuracy of at least 97.28%. Despite its high accuracy and fast convergence rate, the framework's reliance on DenseNet-201 and Differential Evolution may limit its generalization ability and computational efficiency, especially when applied to larger, more complex datasets or real-time systems.

Balaji et al. (2022)[18] propose an optimization-driven deep belief neural network (ODBNN) for medical image classification within IoT-based healthcare systems. This model leverages a combination of ML, DL and AI techniques to address the challenges of high intra-class variance and inter-class similarity inherent in medical image processing. The ODBNN involves image quality enhancement techniques like noise removal and contrast normalization, followed by feature extraction. The May Fly optimization technique is then applied to select the most relevant features for classification. The model is assessed on accuracy, precision, recall, and f-measure, with findings

indicating its superiority over conventional techniques. However, the approach's reliance on multiple optimization and feature extraction techniques could increase computational complexity, potentially making it less efficient for real-time medical image processing applications.

Zhang et al. (2020) [19] proposed a joint DL and IOMT driven architecture for managing and monitoring the massive medical data generated by wearable devices, specifically focusing on cardiac image processing. The authors introduced a self-adaptive power control-based enhanced energy-aware (EEA) framework aimed at reducing energy consumption, enhancing battery lifetime, and improving reliability in wearable devices used for elderly healthcare. Their approach was evaluated using real-time data traces from static (sitting) and dynamic (cycling) activities. Furthermore, the paper proposed a layered architecture for IoMT, a battery model considering wireless channel features and body postures, and optimized network performance by addressing challenges such as energy drain, packet loss ratio (PLR), and sustainability. Experimental results demonstrated that the EEA scheme outperforms conventional constant transmit power control (TPC), significantly enhancing energy efficiency and reliability. Despite the promising results, the paper does not discuss the scalability of the framework in larger, more complex healthcare environments, which remains a key challenge for real-world implementation.

3. Proposed Methodology

3.1 Proposed System

The IoT data analysis workflow begins with raw IoT data collection, followed by a pre-processing stage for data cleaning and normalization. Feature extraction is performed using a dual-model approach combining CNN and LSTM networks, where CNN handles spatial features while LSTM processes temporal patterns[20]. The extracted features undergo optimized model training with hyperparameter tuning for enhanced performance. The trained model generates predictions, which are finally transmitted to cloud infrastructure for storage and deployment. This streamlined workflow ensures efficient processing of IoT data while maintaining scalability and flexibility for various applications.

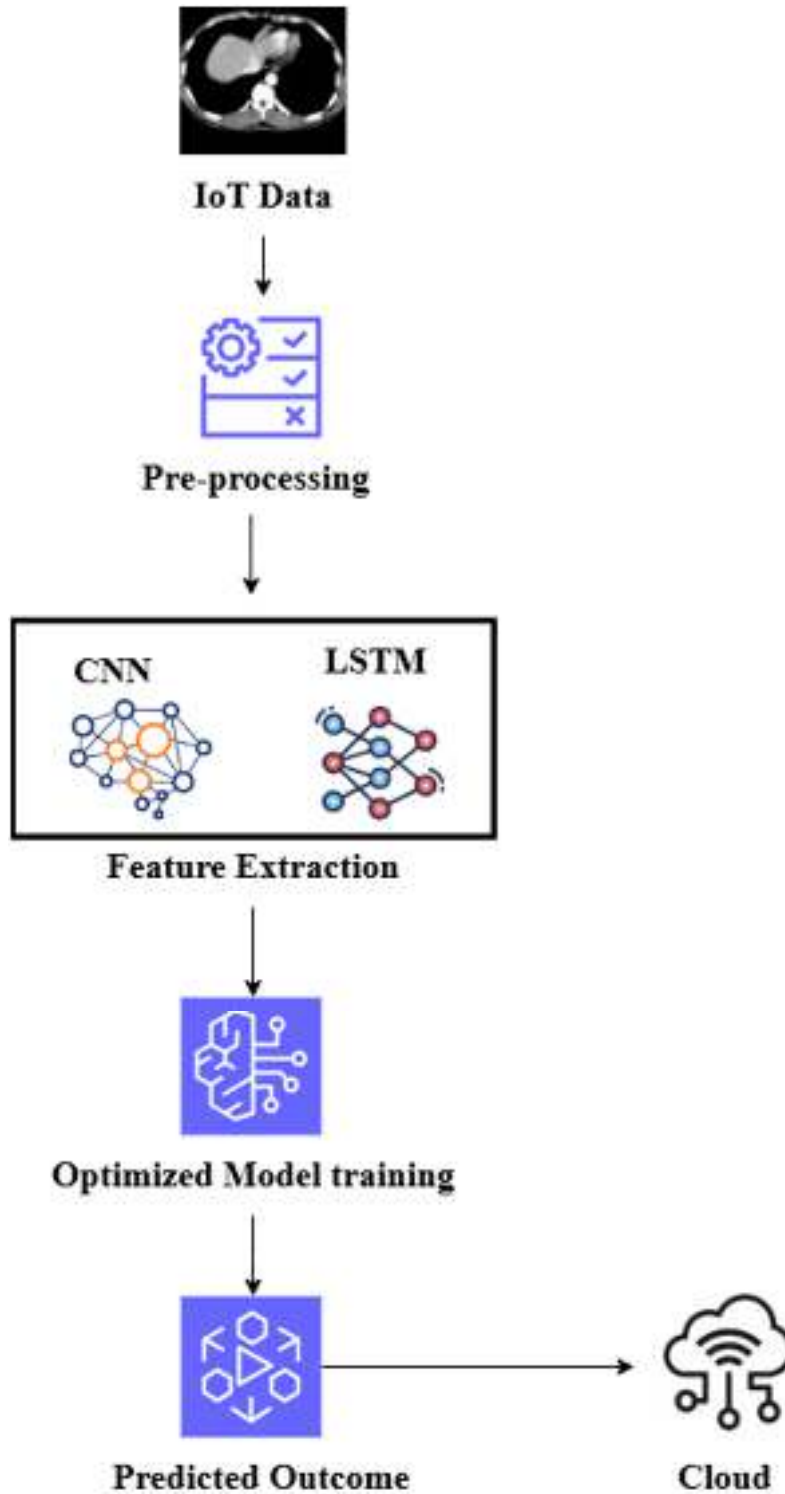


Figure 1: Architecture for the recommended approach.

3.2 Data Collection:

The data for this research is collected using CT imaging sensors integrated into wearable Internet of Medical Things (IoMT) devices, designed to monitor and capture lung health data in

real-time. These sensors continuously acquire detailed CT lung images, providing critical insights into the structural and functional characteristics of the lungs. The CT imaging process involves emitting low-dose X-rays, which pass through the patient's chest to create cross-sectional images of the lungs. These images are then transmitted to a central processing unit, where they are digitized for further analysis[21].

In this study, the collected CT lung data is sourced from wearable IoMT-based devices, enabling continuous and non-invasive monitoring of lung conditions. The imaging process ensures high-resolution data acquisition to capture even minor anomalies or changes in lung structure with precision. To maintain the quality and consistency of the data, the CT lung images are pre-processed to eliminate noise and artifacts, such as motion blur or device interference, which may compromise the accuracy of subsequent analysis.

3.3 Data Pre-processing:

3.3.1 Noise Reduction

CT lung images are often susceptible to various types of noise and artifacts that can significantly distort the quality and clarity of the images, affecting diagnostic accuracy. Common sources of noise in CT lung images include:

Motion Artifacts: Patient movement during image acquisition can lead to blurred or distorted CT images, making it difficult to interpret fine lung structures accurately.

Respiratory Artifacts: Breathing during scanning can cause inconsistencies in the image, introducing noise that interferes with identifying abnormalities.

Beam Hardening: This artifact occurs due to variations in X-ray attenuation across tissues, resulting in streaks or distortions in the CT image.

Scattering Noise: X-rays scattering within the patient's body can produce noise, reducing image contrast and obscuring subtle features critical for diagnosis.

To address these issues, noise reduction techniques are applied during the pre-processing phase. The goal is to eliminate unwanted disturbances while preserving critical information in the lung images. One of the most widely used techniques for noise reduction in CT lung data is Gaussian Filtering.

3.3.2 Normalization

Normalization is another critical pre-processing step to ensure CT lung images are on a consistent scale before inputting them into machine learning or deep learning models. CT images may exhibit varying intensity levels due to differences in scanning protocols, patient anatomy, or device settings. These variations can hinder the ability of models to generalize across different datasets or patient populations.

To address this, normalization techniques are applied to standardize the intensity values of CT images, ensuring that all inputs are within a predictable range. One of the most common normalization methods for CT lung images is Min-Max Normalization.

3.3.2.1 Min-Max Normalization

Min-Max Normalization scales the pixel intensity values of CT images to a fixed range, typically [0, 1]. This normalization shifts the minimum intensity to 0 and the maximum to 1, scaling all intermediate values proportionally. For example, if the raw pixel values in a CT lung image range from -1000 Hounsfield units (HU) to 3000 HU, applying Min-Max normalization will transform the pixel values into the range [0, 1]. By normalizing CT lung images, the consistency of

data across different patients or imaging systems is ensured, reducing discrepancies and improving the stability of the training process. This leads to better algorithm performance and more reliable predictions.

3.4 Feature Extraction:

Feature extraction is a critical step in the analysis and transmission of CT lung images, as it enables the identification and isolation of diagnostic-relevant patterns while minimizing computational overhead. The extracted features holds a significant role in preserving the diagnostic quality of the images during transmission and ensuring efficient processing in IoMT-based systems.

3.4.1 CNN:

CNNs are a type of DL approach that have shown significant achievement in image and signal processing tasks, like object detection, image classification, and pattern recognition. CNNs are specifically well-suited for tasks that involve spatial data, like medical imaging and time-series analysis, including CT signal processing.

CNNs are developed to automatically learn spatial hierarchies of features from raw input information, reducing the demand for manual feature extraction. This capability makes them highly effecient in analyzing complex data, such as CT signals, where crucial features like heart rate, rhythm, and anomalies must be extracted and identified with high accuracy.

3.4.1.1 Convolutional Layer:

The convolutional layer is the primary part of CNNs. It applies a set of learnable filters (also known as kernels) to the input data. These filters slide over the input signal (or image) and compute the convolution operation at each location, producing feature maps that highlight important spatial patterns in the data. Each filter learns to detect different features, like edges, textures, or more complex patterns, depending on the training process.

In the context of CT signal analysis, convolutional filters are trained to identify key components of the signal, such as the P-wave, QRS complex, and T-wave, which are critical for heart rate detection and the identification of arrhythmias.

3.4.1.2 Activation Layer:

After the convolution operation, an activation function is utilised to introduce non-linearity to the model, allowing it to learn complex patterns. The most commonly used activation function in CNNs is the Rectified Linear Unit (ReLU). ReLU transforms all negative values in the feature map to zero, helping the model converge faster and avoid vanishing gradients.

3.4.1.3 Pooling Layer:

The pooling layer reduces the spatial dimensions of the feature maps, which helps reduce computational complexity and the number of parameters in the model. It also makes the network invariant to small translations in the input data. The most common pooling operation is max pooling, which takes the maximum value in a specific window of the feature map. By down-sampling the feature map, pooling reduces its size while retaining the most important features.

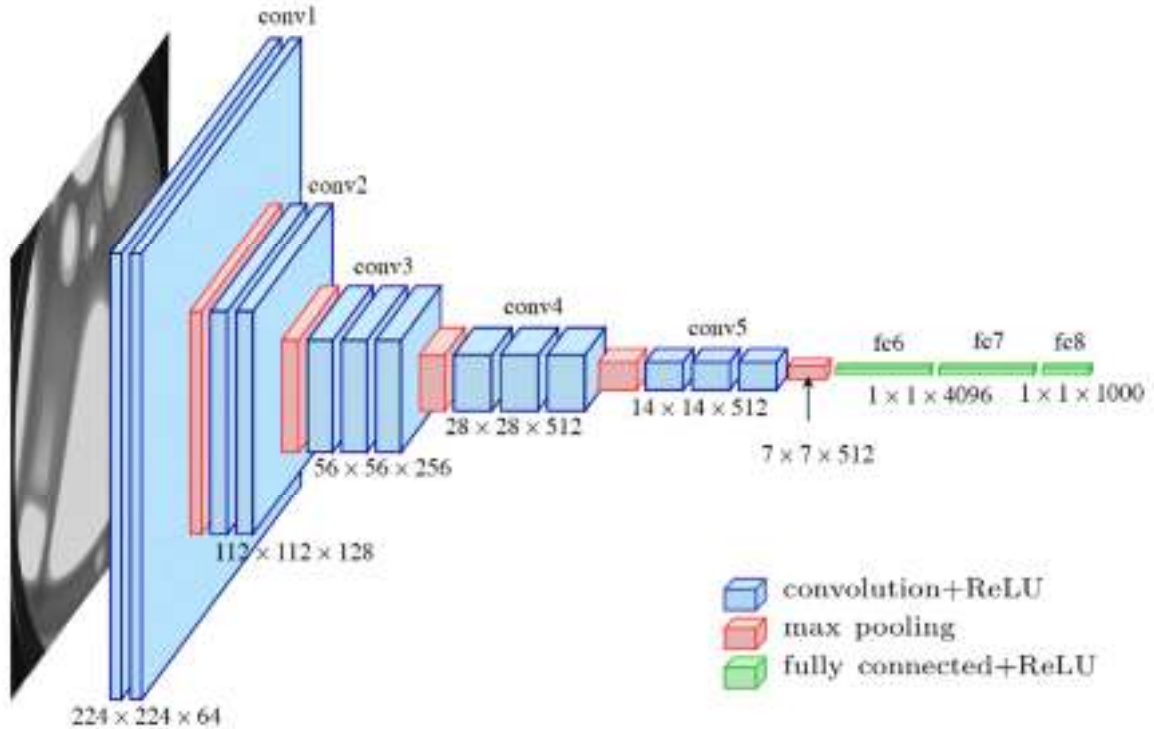


Figure 2 : Architecture for CNN

In CT analysis, pooling layers help simplify the learned features by reducing the dimensionality of the signal and maintaining the crucial characteristics for further analysis. Convolutional Neural Networks (CNNs) are a powerful tool for CT signal processing in the context of IoMT applications. Their ability to automatically learn relevant features, detect patterns across different time scales, and handle noisy data makes them a perfect fit for analyzing CT signals and improving heart disease diagnosis, monitoring, and real-time patient care. By integrating CNNs into wearable IoMT devices, healthcare providers can achieve more precise and timely diagnoses, ultimately improving patient results.

3.4.2 Long Short Term Memory (LSTM):

LSTM is a type of RNN network developed to overcome the limitations of traditional RNNs, particularly the problem of vanishing gradients[22]. LSTMs are specifically efficient in capturing long-range dependencies in sequential data, making them well-suited for tasks like time series prediction, speech recognition, and also in intrusion detection in IoT networks. LSTMs are developed to learn and retain information over long periods of time while discarding irrelevant data. LSTMs achieve this by introducing memory cells and gating mechanisms, which help control the flow of information.

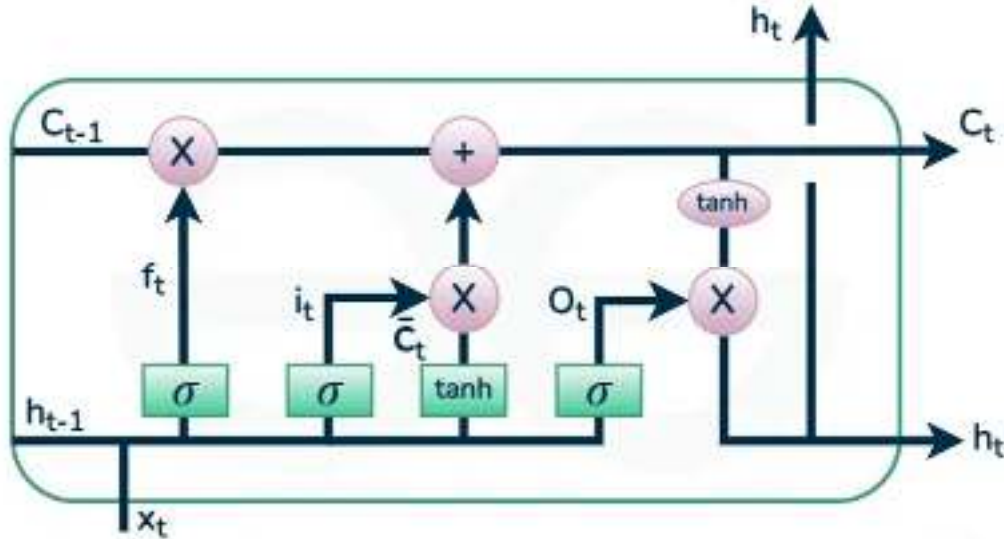


Figure 3: Architecture for LSTM

3.4.2.1 Forget Gate (f_t)

It controls how much of the previous cell state (C_{t-1}) should be carried forward to the current cell state. It takes the previous hidden state (h_{t-1}) and the current input (x_t) as inputs and outputs a value between 0 and 1, which is then used to scale the previous cell state.

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

Where f_t is the forget gate's output, W_f is the weight matrix for the forget gate, b_f is the bias term and σ is the sigmoid activation function, which squashes the output between 0 and 1.

3.4.2.2. Input Gate (i_t):

It determines how much of the new information should be added to the cell state. The candidate cell state represents the potential new information that can be added to the cell state. The input gate is computed as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

Where i_t is the input gate's output, W_i is the weight matrices for the input gate and b_i is the bias terms.

3.4.2.3. Update Cell State (C_t):

It is updated by combining the previous cell state (C_{t-1}) and the new candidate cell state (\tilde{C}_t), controlled by the forget gate and the input gate. The update rule is as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{3}$$

Where C_t is the updated cell state, f_t controls how much of the previous cell state is carried forward and i_t controls how much of the new candidate cell state is added to the cell state.

3.4.2.4. Output Gate (o_t) and Hidden State (h_t):

This gate decides what the next hidden state (h_t) will be. The hidden state is based on the updated cell state (C_t), which is passed through the output gate and squashed using the \tanh function. The output gate is computed as below:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{5}$$

Where o_t is the output gate’s output, h_t is the hidden state, W_o is the weight matrix for the output gate, b_o is the bias term and the $\tanh(C_t)$ function squashes the cell state to ensure that the hidden state is between -1 and 1.

The forget gate (f_t) regulates which information from the previous cell state is retained.

The input gate (it) controls what new information is added to the cell state.

The output gate (ot) determines the next hidden state, which is used as the output of the LSTM unit.

The cell state (Ct) is updated by combining the previous state and the new information, allowing the LSTM to remember long-term dependencies.

These gates and states allow LSTMs to effectively remember and forget information over long sequences, making them suitable for tasks such as time series prediction, language modelling, and intrusion detection in IoT networks, where temporal patterns play a crucial role in identifying attacks.

3.5 Hyperparameter tuning with Hippopotamus Optimization:

3.5.1 Hippopotamus Optimization Algorithm:

The HO approach is a population-relied optimization technique, individuals in the population are represented as hippo. In this method, every hippopotamus aligns with a possible remedy for the optimization task, with its position update in the search space reflecting the values of the decision variables. This phase, a vector of decision variables is yielded by utilizing a specific formula.

$$\chi_i: h_{i,j} = lb_j + r \cdot (ub_j - lb_j), i = 1, 2, \dots, N, j = 1, 2, \dots, m \tag{6}$$

The lower and upper bounds of the j -th decision variable are represented as lb and ub , respectively. Let N be the hippo’s population size, and m depicts count of decision variables in the given challenge. The population matrix is constructed according to the expression in Eq. (7)

$$\chi = \begin{bmatrix} \chi_1 & h_{1,1} & \cdots & v_{1,j} & \cdots & h_{1,m} \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ \chi_i & h_{i,1} & \cdots & h_{i,j} & \cdots & h_{i,m} \\ \vdots & \vdots & \dots & \vdots & \ddots & \vdots \\ \chi_N & h_{N,1} & \cdots & h_{N,j} & \cdots & h_{N,m} \end{bmatrix}_{N \times m} \tag{7}$$

3.5.1.1 Phase 1: Exploration

Hippopotamus groups contain of various adult female hippos, young calves, multiple adult males, and a influential male, who leads the herd. The influential male is classified by an iterative assessment of the criterion function value (minimization problems correspond to the lowest value, while maximization problems aim for the highest value). Typically, hippopotamuses gather in close quarters. The dominant male holds a key place in safeguarding the herd and its territory against threats. Female hippos are situated near the adult males. Once male hippos reach maturity, the dominant male expels them from the herd. The position of the male hippopotamuses in the water body is mathematically expressed in Equation (8).

$$\chi_i^{\mathcal{M}hippo}: h_{i,j}^{\mathcal{M}hippo} = h_{i,j} + y_1 \cdot (Dhippo - I_1 h_{i,j}) \tag{8}$$

$$for i = 1, 2, \dots, \left\lceil \frac{N}{2} \right\rceil \text{ and } j = 1, 2, \dots, m \tag{9}$$

In Equation (12), $imhippo\chi imhippo$ depicts the position of the male hippopotamus, while $hhippoDhippo$ refers to the dominant hippopotamus's position. The vector $r \rightarrow 1, \dots, 4$ denotes random values among 0 and 1, and r_5 is another random value among 0 and 1, as shown in Equation (13). I_1 and I_2 are integers from 1 to 2, as defined in Equations (3) and (6). $mimG_i$ represents the mean values of a randomly opted group of hippopotamuses, which has an equal chance of comprising the current hippopotamus (χ_i), while y_1 is another random value among 0 and 1, as shown in Equation (10). In Equation (11), q_1 and q_2 are random integers that can 1 or 0.

$$h = \begin{cases} I_2 \times \vec{r}_1 + (\sim q_1) \\ 2 \times \vec{r}_2 - 1 \\ \vec{r}_3 \\ I_1 \times \vec{r}_4 + (\sim q_2) \\ r_5 \end{cases} \tag{10}$$

$$T = \exp\left(-\frac{t}{\tau}\right) \tag{11}$$

$$\chi_i^{FBhippo}: \chi_{i,j}^{FBhippo} = \begin{cases} x_{i,j} + h_1 \cdot (Dhippo - I_2 MG_i) T > 0.6 \\ \Xi else \end{cases} \tag{12}$$

$$\Xi = \begin{cases} x_{i,j} + k_2 \cdot (MG_i - Dhippo) r_6 > 0.5 \\ lb_j + r_7 \cdot (ub_j - lb_j) else \end{cases} \tag{13}$$

for $i = 1, 2, \dots, \lfloor \frac{N}{2} \rfloor$ and $j = 1, 2, \dots, m$.

Equations (12) and (13) define the position of female or juvenile hippopotamuses ($ihippo\chi iFBhippo$) within the group. Typically, juvenile hippopotamuses remain close to their mothers; however, due to their curiosity, they may periodically move distant from the group or their mothers. If T exceeds 0.6, this indicates that the juvenile hippopotamus has moved away from its mother. The juvenile is considered to have detached from the herd entirely. The values h_1 and h_2 are random numbers or vectors opted from one of the 5 scenarios sketched in the h equation. Here, r_7 is a random value among 0 and 1. The position update for both male and female or juvenile hippopotamuses within the herd is described in Eqs. (14) and (15). iFi represents the objective function value.

$$\chi_i = \begin{cases} \chi_i^{Mhippo} \mathcal{F}_i^{Mhippo} < \mathcal{F}_i \\ \chi_i else \end{cases} \tag{14}$$

$$\chi_i = \begin{cases} \chi_i^{FBhippo} \mathcal{F}_i^{FBhippo} < \mathcal{F}_i \\ \chi_i else \end{cases} \tag{15}$$

Using h vectors, I_1 and I_2 case elevate the global search and optimizes exploration in the recommended approach.

Phase 2: Exploration

Hippopotamuses typically live in groups for their own protection and safety. The size and strength of these groups act as a deterrent to predators, making it less likely for them to approach. Young hippos, driven by curiosity, may occasionally wander away from the safety of the herd. This is because they are not as strong as adult hippopotamuses. Similarly, hippos that are ill are at a higher risk of becoming prey for these predators.

This phase, hippopotamuses could exhibit a trait where they move toward the predator, prompting it to retreat and thereby successfully averting the potential danger. Equation (16) describes the position of the predator within the search space.

$$Predator: Predator_j = lb_j + \vec{r}_g \cdot (ub_j - lb_j), j = 1, 2, \dots, m. \quad (16)$$

$$\vec{D} = |Predator_j - x_{i,j}| \quad (17)$$

Equation (17) represents the distance between the i th hippo and the predator. In this phase, the hippo opts a defensive posture, determined by the factor $jFPredator_j$, which aids in its protection from the predator. When $jFPredator_j$ is smaller than iFi , it signifies that the predator is dangerously nearer to the hippopotamus. The hippo quickly turns toward the predator and advances toward it, aiming to force the predator to retreat. Conversely, if $jFPredator_j$ is larger, it indicates that the predator or intruder is farther away from the hippo's territory, as described in Eq. (18). In this situation, the hippopotamus still faces the predator and also restricts its movement, signalling its presence within its domain without engaging in an immediate confrontation.

$$\chi_i^{HippoR}: x_{i,j}^{HippoR} = \begin{cases} \vec{RL} \oplus Predator_j, + (\frac{f}{(c-d \times \cos(2\pi g))}) \cdot (\frac{1}{D}) \mathcal{F}_{Predator_j} < \mathcal{F}_i \\ \vec{RL} \oplus Predator_j, + (\frac{f}{(c-d \times \cos(2\pi g))}) \cdot (\frac{1}{2 \times D + r_g}) \mathcal{F}_{Predator_j} \geq \mathcal{F}_i \end{cases} \quad (18)$$

$$for i = [\frac{N}{2}] + 1, [\frac{N}{2}] + 2, \dots, N \text{ and } j = 1, 2, \dots, m$$

The hippopotamus position, denoted as $iippo\chi iHippoR$, refers to a scenario where the animal faces a predator. The vector $RL \rightarrow$ represents a random variable following a Lévy distribution, which is utilized to model sudden shifts in the predator's location while an attack on the hippopotamus. The Lévy movement model is expressed by Eq. (19). The terms w and v represent random values within the range $[0,1]$, while ϑ is a constant set to 1.5. The symbol Γ symbolizes Gamma function, and $w\sigma w$ can be derived from Eq. (20).

$$Levy(\vartheta) = 0.05 \times \frac{w \times \sigma_w}{|v|^{\frac{1}{\vartheta}}} \quad (19)$$

$$\sigma_w = \left[\frac{\Gamma(1+\vartheta) \sin(\frac{\pi\vartheta}{2})}{\Gamma(\frac{(1+\vartheta)}{2}) \vartheta 2^{\frac{\vartheta-1}{2}}} \right]^{\frac{1}{\vartheta}} \quad (20)$$

In Equation (19), β denotes a random value within the range of 2 to 4, c represents a random number among 1 and 1.5, and D is another random value among 2 and 3. g is a random variable within the interval -1 to 1 . $r \rightarrow \vartheta$ is a vector of random values with dimensions $m1 \times m$. In Equation (20), if $iipoiFiHippoR$ exceeds F , it depicts that the hippopotamus has been hunted, prompting the introduction of a new hippopotamus into the herd. Else, the hunter will retreat, and the hippo will rejoin the group. The second phase demonstrated notable enhancements in the global search mechanism. Both the first and second phases operate synergistically, ensuring a robust approach to minimizing the risk of getting stuck in local minima.

$$\chi_i = \begin{cases} \chi_i^{HippoR} \mathcal{F}_i^{HippoR} < \mathcal{F}_i \\ \chi_i \mathcal{F}_i^{HippoR} \geq \mathcal{F}_i \end{cases} \quad (21)$$

Phase 3: Exploitation

When a hippo faces a group of predators or is unable to defend itself, it often attempts to move away from the threat. Typically, the hippopotamus seeks refuge in a nearby lake or pond, as predators like spotted lions and hyenas generally avoid these water bodies. This strategy helps the hippopotamus find a safer location close to its original location. In Phase Three of the HO approach, this behaviour is reflected by enhancing the local search capabilities. To emulate this, a random location near the hippopotamus's current location is generated. If this new position results in an improved cost function, it suggests the hippopotamus has favorably found a safer spot and has

moved to it. Here, t depicts the current iteration, and T denotes the maximum number of iterations.

$$lb_j^{local} = \frac{lb_j}{t}, ub_j^{local} = \frac{ub_j}{t}, t = 1, 2, \dots, T \quad (22)$$

$$\chi_i^{HippoE}: x_{i,j}^{HippoE} = x_{i,j} + r_{10} \cdot (lb_j^{local} + s_1 \cdot (ub_j^{local} - lb_j^{local})) \quad (23)$$

$$i = 1, 2, \dots, \mathcal{N}, j = 1, 2, \dots, m$$

In equation (22), *hippoE* depicts the location of the hippopotamus, which was examined to identify the nearest secure area. s_1 is a randomly used by a vector or value, selected from among three potential options as described in equation (18). These available alternatives (s_1) contribute to a optimal local search, ultimately elevating the overall exploitation performance of the algorithm.

$$s = \begin{cases} 2 \times \vec{r}_{11} - 1 \\ r_{12} \\ r_{13} \end{cases} \quad (24)$$

In equation (28), \vec{r}_{11} denotes a stochastic vector with values among 0 and 1. Similarly, r_{10} (as shown in equation 27) and r_{13} are random values produced within the 0 to 1 range. Additionally, r_{12} is a random number following a normal distribution

$$\chi_i = \begin{cases} \chi_i^{HippoE} \mathcal{F}_i^{HippoE} < \mathcal{F}_i \\ \chi_i \mathcal{F}_i^{HippoE} \geq \mathcal{F}_i \end{cases} \quad (25)$$

3.6 Hyper parameter Tuning Process

In this research, Hippo Optimization is used for hyperparameter tuning, which is a metaheuristic algorithm designed to find optimal hyperparameters for deep learning models. Hyperparameters are parameters that are set before training a model and significantly influence the model’s performance, including learning rate, batch size, number of layers, and number of units in each layer. In the context of this study, the aim of hyperparameter tuning is to minimize the discrepancy between predicted accuracy and actual accuracy, ensuring that the model can generalize well on unseen data. This process helps in identifying the best set of hyperparameters that provide optimal performance in terms of accuracy and efficiency.

After feature extraction using CNNs and LSTM networks, hyperparameter tuning is applied in the dense layers of the model. These dense layers play a crucial role in final decision-making, where the extracted features are used to classify the input data. With the help of the Hippo Optimization Algorithm, the hyperparameters of these dense layers, such as the number of units, activation functions, and dropout rates, are optimized. This ensures that the model is not only capable of accurate feature extraction but also able to perform precise and efficient classification based on the tuned parameters.

$$Fitness\ Function = \text{Min}((\text{Predicted Accuracy} - \text{Actual Accuracy})) \quad (26)$$

4. Results and Discussions

4.1 Implementation Details:

The absolute solution for the approach was built utilizing an Intel Workstation powered by I7 CPU, NVIDIA GPU, 16GB RAM, and a 3.2 GHz processing frequency.

4.2 Performance Metrics:

Evaluation Measures like accuracy, precision, recall, specificity, and F1-score are evaluated to assess the performance of the proposed approach for CT lung image transmission within the IoMT environment. These metrics are benchmarked in real-time against other advanced DL methods to underscore the advantages of the proposed CNN-LSTM framework. Both the performance metrics and latency overhead are critically analyzed to validate the framework's efficiency. Table 1 provides the mathematical formulas used for calculating these performance metrics.

Table 1: Performance measures utilized in the examination

SL.NO	Performance Measures	Expression
1	Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
2	Recall	$\frac{TP}{TP + FN} \times 100$
3	Specificity	$\frac{TN}{TN + FP}$
4	Precision	$\frac{TP}{TP + FP}$
5	F1-Score	$2 \cdot \frac{Precision * Recall}{Precision + Recall}$

TP & TN are True Positive & negative, FP & FN are False Positive & negative.

A prediction case can be broken down into four categories: First, we have TP, where the identified values are correct and align with the actual truth. Next, FP refers to instances where the values are incorrectly labelled as true, although they are, in fact, false. FN occurs when a true value is mistakenly identified as negative. Lastly, TN describes the scenario where a value is accurately classified as negative, corresponding to the actual negative status.

4.3 Experimental Outcomes:

The effectiveness of the proposed DL framework for medical image transmission in the IoMT environment was evaluated using several key metrics, like Peak Signal-to-Noise Ratio (PSNR), accuracy, F1 score, specificity, and sensitivity. PSNR was used to assess the quality of the transmitted medical images, where higher values indicated minimal distortion and preserved diagnostic value. Accuracy measured the overall correctness of the model's predictions, while the F1 score offered a balance between precision and recall, ensuring robust classification even in imbalanced datasets. Specificity and sensitivity evaluated the model's ability to correctly recognize negative and positive cases, respectively, which is critical in medical diagnostics to avoid false positives and ensure early detection of abnormal conditions. The results demonstrated that the CNN-LSTM framework, optimized using Hippo Optimization for hyperparameter tuning, outperformed traditional models, achieving superior performance across these metrics. This showcases the framework's effectiveness in ensuring efficient, secure, and reliable medical image transmission in IoMT environments, ultimately contributing to better real-time patient monitoring and care.

Table 2: Performance metrics for the recommended approach

Algorithms	Performance Metrics (%)				
	Accuracy	Precision	Recall	Specificity	F1-score
ANN	84.3	86.3	86.5	86.5	85.3
CNN	87.2	87.2	86.6	86.3	83.2
DNN	88.5	88.1	84.3	88.6	86.3
RNN	89.9	90.1	89.74	88.3	88.2
Proposed Model	98.3	97.0	97.8	96.2	96.0

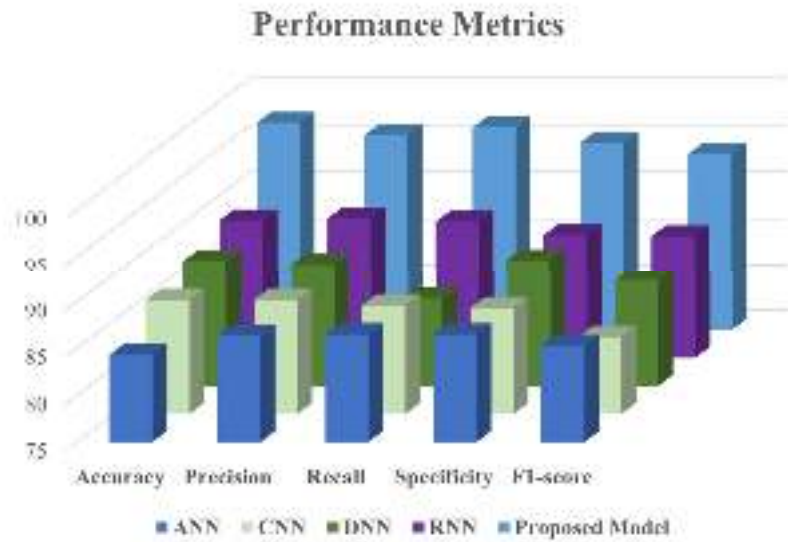


Figure 4: Comparative analysis with the Residing approach

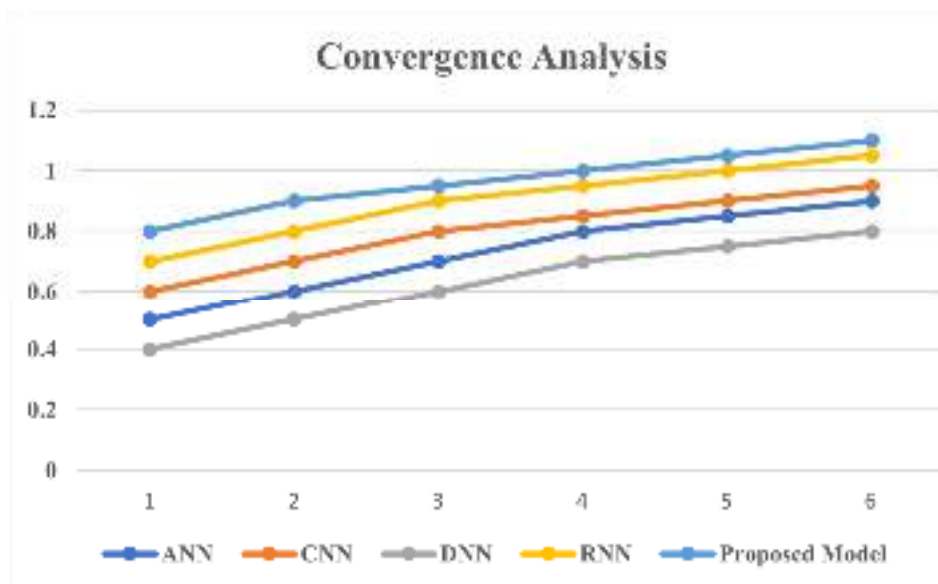


Figure 5: Convergence Analysis for the Proposed approach

Metric	Proposed Framework	Existing Method (Baseline)
Compression Ratio	4:1	3:1
Encryption Time (ms)	12 ms	18 ms
PSNR (dB)	38.5	34.2
Accuracy (Diagnosis)	98%	95%
Bandwidth Savings (%)	75%	60%

Table 3: Performance Comparison between the recommended architecture and Traditional Model

The recommended architecture outperforms the baseline method across key metrics, achieving a higher compression ratio (4:1 vs. 3:1) and improved PSNR (38.5 dB vs. 34.2 dB), ensuring superior image quality. Additionally, it demonstrates faster encryption (12 ms vs. 18 ms), greater diagnostic accuracy (98% vs. 95%), and significant bandwidth savings (75% vs. 60%), highlighting its efficiency and reliability.

5. CONCLUSIONS

This research introduces an optimized deep learning framework for medical image transmission in IoMT environments. The proposed solution incorporates CNN and dimensionality reduction, while LSTM handle sequential data to overcome issues such as packet loss and latency in IoMT networks. The framework also includes robust encryption mechanisms to ensure data security without significantly increasing computational overhead. Performance evaluations using real-world medical image datasets under varying IoMT network conditions demonstrate that the proposed model excels across key metrics, like Peak Signal-to-Noise Ratio (PSNR), accuracy, F1 score, specificity, and sensitivity. Additionally, the model optimizes encryption and decryption times, reducing bandwidth consumption and ensuring efficient and secure data transmission. Future enhancements to improve real-time optimization, incorporate more advanced encryption techniques, and integrate multi-modal data will further refine the framework's performance and applicability. Overall, this study demonstrates a significant step toward improving the reliability and efficiency of medical image transmission in IoMT-based healthcare networks.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Wazid, Mohammad, and Prosanta Gope. "BACKM-EHA: A novel blockchain-enabled security solution for IoMT-based e-healthcare applications." *ACM Transactions on Internet Technology* 23.3 (2023): 1-28.
- [2] Muhammad, Ghulam, et al. "A comprehensive survey on multimodal medical signals fusion for smart healthcare systems." *Information Fusion* 76 (2021): 355-375.
- [3] Ghazal, Taher M., et al. "Secure IoMT pattern recognition and exploitation for multimedia information processing using private blockchain and fuzzy logic." *Transactions on Asian and Low-Resource Language Information Processing* (2022).
- [4] R R, Rajalaxmi & Prasad, L.V. & Janakiramaiah, B. & Pavankumar, C.S. & Neelima, N. & Veerappampalayam

- Easwaramoorthy, Sathishkumar. (2022). Optimizing Hyperparameters and Performance Analysis of LSTM Model in Detecting Fake News on Social media. *ACM Transactions on Asian and Low-Resource Language Information Processing*. 10.1145/3511897.
- [5] Gkelios, Socratis & Features, Deep & Plakias, Spiros & Boutalis, Yiannis & Chatzichristofis, Savvas. (2021). Deep Convolutional Features for Image Retrieval. *Expert Systems with Applications*. 177. 114940. 10.1016/j.eswa.2021.114940.
- [6] Peng, Yitao & He, Lianghua & Hu, Die & Liu, Yihang & Yang, Longzhen & Shang, Shaohua. (2024). Decoupling Deep Learning for Enhanced Image Recognition Interpretability. *ACM Transactions on Multimedia Computing, Communications, and Applications*. 20. 10.1145/3674837.
- [7] Elsayed Abd Elaziz, Mohamed & Mabrouk, Alhassan & Dahou, Abdelghani & Aaaa, Samo. (2022). Medical Image Classification Utilizing Ensemble Learning and Levy Flight-Based Honey Badger Algorithm on 6G-Enabled Internet of Things. *Computational Intelligence and Neuroscience*. 2022. 1-17. 10.1155/2022/5830766.
- [8] Hosseinzadeh, Mehdi & Ahmed, Omed & Ghafour, Marwan & Safara, Fatemeh & Ali, Saqib & Vo, Bay & Chiang, Hsiu-Sen. (2021). A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things. *The Journal of Supercomputing*. 77. 10.1007/s11227-020-03404-w.
- [9] Khamparia, Aditya & Singh, Prakash & Rani, Poonam & Samanta, Debabrata & Khanna, Ashish & Bhushan, Bharat. (2021). An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. *Transactions on Emerging Telecommunications Technologies*. 32. 10.1002/ett.3963.
- [10] Balhareth, G.; Ilyas, M. Optimized Intrusion Detection for IoMT Networks with Tree-Based Machine Learning and Filter-Based Feature Selection. *Sensors* 2024, 24, 5712. <https://doi.org/10.3390/s24175712>
- [11] Rahman, A., Debnath, T., Kundu, D., Khan, M. S. I., Aishi, A. A., Sazzad, S., Sayduzzaman, M., & Band, S. S. (2024). Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges and opportunities. *Public Health*, 11(1), 58-109. <https://doi.org/10.3934/publichealth.2024004>
- [12] Nandagopal, M., Seerangan, K., Govindaraju, T. et al. A Deep Auto-Optimized Collaborative Learning (DACL) model for disease prognosis using AI-IoMT systems. *Sci Rep* 14, 10280 (2024). <https://doi.org/10.1038/s41598-024-59846-2>
- [13] Datta Gupta, K., Sharma, D.K., Ahmed, S. et al. A Novel Lightweight Deep Learning-Based Histopathological Image Classification Model for IoMT. *Neural Process Lett* 55, 205–228 (2023). <https://doi.org/10.1007/s11063-021-10555-1>.
- [14] Lata, K.; Cenkeramaddi, L.R. Deep Learning for Medical Image Cryptography: A Comprehensive Review. *Appl. Sci.* 2023, 13, 8295. <https://doi.org/10.3390/app13148295>.
- [15] Thimmapuram, Madhuri & Salvadi, Shoba & Pallati, Narsimhulu & Aluvalu, Rajanikanth. (2023). Medical Image Classification Using DL-based Feature Extraction in IoMT. *Recent Patents on Engineering*. 17. 10.2174/1872212117666230222093128.
- [16] Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. *Front Public Health*. 2023 Nov 7;11:1273253. doi: 10.3389/fpubh.2023.1273253. PMID: 38026291; PMCID: PMC10662291.
- [17] Arti Tiwari and Millie Pant. 2022. Optimized Deep-Neural Network for Content-based Medical Image Retrieval in a Brownfield IoMT Network. *ACM Trans. Multimedia Comput. Commun. Appl.* 18, 2s, Article 125 (June 2022), 26 pages. <https://doi.org/10.1145/3546194>
- [18] Balaji, P., Sri Revathi, B., Gobinathan, P., Shamsudheen, S., & Vaiyapuri, T. (2022). Optimal IoT based improved deep learning model for medical image classification. *Computers, Materials & Continua*. <https://doi.org/10.32604/cmc.2022.028560>
- [19] T. Zhang et al., "A Joint Deep Learning and Internet of Medical Things Driven Framework for Elderly Patients," in *IEEE Access*, vol. 8, pp. 75822-75832, 2020, doi: 10.1109/ACCESS.2020.2989143
- [20] Ogundokun, Roseline & Misra, Sanjay & Douglas, Mychal & Damaševičius, Robertas & Maskeliunas, Rytis. (2022). Medical Internet-of-Things Based Breast Cancer Diagnosis Using Hyperparameter-Optimized Neural Networks. *Future Internet*. 14. 153. 10.3390/fi14050153.
- [21] Cooney, Ciaran & Korik, Attila & Folli, Raffaella & Coyle, Damien. (2020). Evaluation of Hyperparameter Optimization in Machine and Deep Learning Methods for Decoding Imagined Speech EEG. *Sensors*. 20. 4629. 10.3390/s20164629.
- [22] Memon, Muhammad & Li, Jian & Haq, Amin & Memon, Muhammad & Zhou, Wang. (2019). Breast Cancer Detection in the IOT Health Environment Using Modified Recursive Feature Selection. *International Journal of Wireless and Mobile Computing*. 2019. 10.1155/2019/5176705.