IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 30, No. 6, June, 2025, pp. 1483-1490
Journal homepage: http://iieta.org/journals/isi

A Hybrid CNN-BiLSTM Model for Minimizing Packet Loss in IoT-Enabled Wireless Sensor Networks

G N Shwetha[1], Shashank Dhananjaya[2], H Jagruthi[3], K S Rekha[1], R Pankaja[4], Abhilasha P Kumar[4*], R Mahesh[5]

[1] Department of Computer Science and Engineering, JSS Science and Technology University, Mysuru 570006, India
[2] Department of Information Science and Engineering, The National Institute of Engineering, Mysuru 570008, India
[3] Department of Information Science and Engineering, BNM Institute of Technology, Bangalore 560070, India
[4] Department of Artificial Intelligence and Machine Learning, BNM Institute of Technology, Bangalore 560070, India
[5] Department of Computer Science and Engineering, BGS Institute of Technology, Adichunchanagiri University, Mandya 571418, India

Corresponding Author Email: abhilasha@bnmit.in

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/isi.300607

Received: 22 April 2025
Revised: 28 May 2025
Accepted: 10 June 2025
Available online: 30 June 2025

Keywords:
BiLSTM, Convolutional Neural Network, hybrid deep learning, IoT, packet loss, wireless sensor network, network reliability

ABSTRACT

Sensors embedded in Wireless Sensor Networks (WSNs) form a foundation in the Internet of Things (IoT) architecture. Nonetheless, packet loss caused by unreliable communication, interference, and energy limitations continues to be a major issue. In this paper, we propose a Convolutional Neural Networks and Bidirectional Long Short Term Memory (CNN-BiLSTM) combined Deep Learning (DL) approach for packet loss minimization in IoT based WSNs. Our model uniquely integrates CNN to capture spatial features with a BiLSTM to capture temporal dependencies, allowing for more accurate inherent prediction of packet loss and intelligent routing in IoT-enabled WSNs. This hybrid design allows for the proposed model to outperform independent deep learning models and traditional routing protocols in both prediction accuracy and performance at the network level. Given the traditional models such as AODV and independent LSTM/CNN approaches. Proposed model has a packet loss reduction of 52%, an overall throughput improvement of 18.7%, and maintained low latency and energy consumption, contributing to the overall success of routing decisions in practical WSN scenarios. This makes the proposed hybrid model is highly suitable for the implementation in the real-time applications.

1. INTRODUCTION

A WSN forms the backbone of the IoT, changing the way devices interact with each other. This will be set up as a part of generation, where this is the third generation of IoT, where everything will be connected to the cloud at real time, together with the database which will be used to collect and monitor different things at real time, such as is done in healthcare, agriculture and industrial automation. In spite of their utility, WSNs are affected by limited energy resources, unreliable links, and data congestion that cause a lot of packet loss. Packet loss should be avoided as much as possible, since it can result in the loss of data during transmission. Deep learning methods have achieved incredible performance on complex, nonlinear data in the last few years. We introduce a hybrid Convolutional Neural Networks and Bidirectional Long Short Term Memory (CNN-BiLSTM) model that learns both the spatial and temporal features of WSN data for packet loss prevention which help proactive approach for network management.

Ever since the advent of the Internet of Things (IoT), WSNs have seen an exponential growth in the deployment over the years in various sectors like environmental monitoring, healthcare, smart agriculture, industrial automation etc. These networks consist of spatially distributed sensor nodes that cooperatively monitor physical or environmental conditions and aggregate data to send to central hubs for processing and analysis. Said et al. [1] have discussed the benefits and utility of WSNs are expanding, in effect, their packet loss rates remain high due to limited energy constraints, volatile communication links, and dynamic network topologies.

Especially in mission-critical environments, minimizing packet loss is crucial for the reliability and accuracy of IoT-based systems. Standard protocols like AODV and DSR do not adapt well in frequency in node behavior and network load causing deterioration in performance [2]. In this context, Machine Learning (ML)-based approaches have been proposed, as they allow for strategies for routing and transmission to be dynamically adapted in light of historical data behavioral patterns. Nevertheless, these models struggle to accurately identify spatial and temporal dependencies present in WSN data streams [3].

Recently, DL methods have become appealing alternatives for optimizing networks because they outperform other approaches in modeling complex, nonlinear relations. CNNs have been extensively applied for spatial feature extraction on

structured data, whereas LSTM networks and their variants, especially BiLSTM, are very effective in learning sequential dependencies [4]. In other domains, hybrid approaches consisting of CNNs and BiLSTMs have been proven successful when applied to traffic prediction or anomaly detection, yet, to our knowledge, this approach has not been tested to minimize packet loss in WSNs [5]. In this paper, we propose a hybrid DL architecture based on the combination of CNN and BiLSTM layers which is used to predict and smartly reduce packet loss in IoT-enabled WSNs. The use of sensor network data would allow the model to utilize both spatial and temporal features to find optimal routing decisions, thus improving network reliability. Results show substantial benefits when compared to state-of-the-art methods in packet delivery ratio, energy efficiency and throughput.

Few previous studies had applied either CNNs for spatial feature representation, or LSTMs/BiLSTMs for temporal representational purposes, and none of them implemented a joint optimization process that incorporated both spatial and temporal features to be able to predict packet loss within WSNs. Additionally, the models described above do not take into account the deployment constraints imposed by energy-constrained IoT devices, limiting the possibility of developing a lightweight and strong predictive system for practical real world environments.

The paper is structured as follows. Section 2 presents a complete review of the state-of-the-art in IoT energy efficiency and packet loss and a summary of some of the most relevant works and a description of pre-existing approaches to address attacks in the IoT space. Section 3 explains the proposed model in depth with mathematical model and pseudocode. Section 4 provides the experimental evaluation of the proposed detection scheme and displays the results with a complete comparative discussion between proposed model and other related researches conducted. Finally, we conclude proposed work results and provide future perspectives in Section 5 followed by references section.

2. RELATED WORK

DL has recently demonstrated considerable promise in increasing data accuracy and managing traffic within IoT-based WSNs. Zhang et al. [6] presented a DL for IoT devices data processing. Similarly, Inayat et al. [7] discussed hybrid DL models focused on IoT security, indicating their ability to improve data transmission reliability. Jing et al. [8] conducted an extensive literature review where the various ML techniques applied for resource management of cellular and IoT networks are classified, noting that DL-based traffic prediction is the leading solution to reduce the loss of packets.

Ullah et al. [9] launched a review of AI associated data transport in IOT, where they claim that hybrid DL models (CNN-LSTMs) are more effective than conventional techniques in applications where loss is a major concern. Wang et al. [10] reviewed DL applications in IoT and pointed out that in network traffic, CNNs capture local patterns, which are important for minimizing loss. Such a framework can target specifically packet loss in WSNs that using CNNs provides better accuracy in predicting packet loss than traditional statistical models [11].

In order to introduce temporal dependencies in IoT traffic, BiLSTM networks have been combined with CNNs. Omarov et al. [12] used BiLSTM to classify multi-class IoT traffic, showing that BiLSTM outperforms vanilla LSTMs at processing sequential data. Data collection and transmission is another area in WSNs where DL has been introduced due to the importance of energy efficiency. Yuan et al. [13] used a deep reinforcement learning (DRL) based model for energy-efficient data aggregation in WSNs that decreases the number of packets loss at the same time as extending network lifetime. Their experiment shows that collisions and retransmissions could be reduced by a sufficiently intelligent schedule. Bauyrzhan Omarov et al. [14] made significant progress by proposing a hybrid CNN-BiLSTM model for anomaly detection in IoT that indirectly helps to mitigate packet loss by detecting faulty transmissions.

Hybrid CNN-BiLSTM models have been leveraged in the context of packet loss in the IoT networks by several authors in recent studies. Rajalakshmi et al. [15] presented a CNN-LSTM model in this system to minimize latency and loss in industrial IoT and stimulate traffic forecast. Latif et al. [16] introduced a DL-based reliability framework for Industrial Internet of Things (IIoT) applications in which real-time fault detection is further utilized to prevent data loss in mission-critical applications. López-Ardao et al. [17] proposed a DL-based system for packet loss recovery in WSNs, and attained better results with CNN-BiLSTM hybrids than with standalone models in noisy conditions. These works emphasize the promise of hybrid models while leaving opportunities to better optimize WSNs with energy constraints.

Evolving ML and DL has played a major role in reducing the packet loss in the IoT based WSNs. Ullah Khan et al. [18] proposed a novel routing mechanism that utilizes ML to dynamically update routing paths in IoT networks. This approach reported a packet loss of 30%, which is a significant reduction compared to static routing protocols. In their study, they emphasise the necessity of keeping routing decisions intelligent in context to the network congested IoT environments. Similarly, Elsayem et al. [19] proposed a reinforcement learning (RL)-based method for dynamic resource allocation. Raman [20] focused on optimizing bandwidth and transmission power to reduce packet drops associated with real-time IoT traffic. Their network framework performed at a higher level with varying network conditions. The spatial feature extraction is realized through CNNs in WSNs. Hong et al. [21] showed the effectiveness of CNNs particular for UAV-based sensor networks: CNNs are shown to improve routing efficiency by capturing spatial-temporal data correlation.

Altunay and Albayrak [22] integrated CNNs and long-short term memory (LSTM) networks for real-time edge analytics in intelligent transportation systems, demonstrating enhanced robustness for dynamic environments. Sadhwani et al. [23] re-engineered a replicable architecture for data intrusion detection to guarantee continuity and security of data flow. Additionally, Dritsas and Trigka [24] addressed federated learning (FL) for distributed IoT networks, where they concluded that decentralized training may improve prediction accuracy with lower communication overhead, which is an essential contribution to reducing packet losses in large-scale WSNs.

For proactive packet loss mitigation, accurate real-time traffic forecasting is a prerequisite. Ghosh et al. [25] proposed a hybrid CNN-LSTM model for traffic prediction in IoT systems, providing a 25% enhancement in loss avoidance over individual models. The latter is successful in modeling spatial (CNN) and temporal (LSTM) dependencies of network traffic.

Meanwhile, Rajawat et al. [26] analyzed quantum-augmented ML tailored for IoT security, proposing that quantum neural networks (QNNs) may create a paradigm shift in loss-resistant communication for future WSNs. Kaur and Gupta [27] expanded this research line by integrating both AI and 6G-IoT in their framework to ensure secure and reliable data transmission through adaptive encryption and attack prediction. Xu et al. [28] recently improved latency-sensitive IoT applications with edge intelligence, resulting in DL inference on a given localized IoT device, which considerably reduced the remote dependency leading to packet loss. Packet loss in IoT networks is also caused by security-related disruptions.

Recent advancements in Transformer architectures and quantum ML (QML) are changing the paradigm of IoT reliability. Tseng et al. [29] demonstrated the use of Transformer architectures for IoT data analytics, achieving state-of-the-art sequence modeling in traffic prediction. Ahanger et al. [30] conducted a survey on DL for anomaly detection, indicating that AI-driven intrusion detection systems (IDS) are capable of blocking harmful packet drops.

Although early approaches have advanced ML-enabled routing, real-time prediction, and security, most methods fall into one of three categories:

- Be pre-trained on CNNs or LSTMs only.
- Ignore energy constraints in WSNs in the DL model deployment.
- Traditional modeling approaches do not have mechanisms to adapt to the dynamic nature of IoT environments.

Although previous studies investigated DL for IoT reliability, the major of these studies consider CNN or BiLSTM spatial or temporal features only. Very few discuss the joint optimization of both aimed at reducing packet loss in resource-constrained WSNs. Our work bridges this gap by:

- A lightweight hybrid CNN-BiLSTM model for low-power WSNs is proposed.
- Incorporating real-time traffic prediction with adaptive routing to avoid loss of packets.
- Navigate performance locked under different network congestion.

3. PROPOSED METHOD

We propose a hybrid AI model to perform this task by combining Convolutional Neural Network (CNN) with Bidirectional Long Short-Term Memory (BiLSTM) network in order to reduce packet loss in IoT-enabled WSNs. The CNN is used for extracting spatial features from the structured sensor data, while the BiLSTM is used for modeling temporal dependencies over packet data stored in the time sequence. The architecture of this system presents five stages, namely Data acquisition from sensors, Preprocessing and reshaping into 2D matrices of data, Feature extraction via CNN, Sequence learning via BiLSTM, and Output classification indicating packet delivery success or failure. The output that this model provides is end-to-end trainable, which means that its implementation is adaptable to different use cases such as changing load on the network, interference, or node failures.

The algorithm we propose, termed Hybrid CNN_BiLSTM for Packet Loss Prediction, first takes as an input time series of sensor data S={s1, s2, ..., sn} collected from different IoT nodes for a certain period of time in a wireless sensor network. The performance of the algorithm is fundamentally to mine this spatiotemporal data and predict if a packet will be lost during transmission. This produces a binary output 0 for no packet loss and 1 for predicted packet loss. Let the raw sensor data at time step t from N nodes be denoted as a multivariate time series in Eq. (1).

$$X_t = [x_t^1, x_t^2, \dots, x_t^N]^T \in \mathbb{R}^N \tag{1}$$

To feed it into a CNN, the data is reshaped into a 2D matrix $X \in \mathbb{R}^{\{m \times n\}}$, where m × n = N, to simulate spatial structure. Data Preprocessing: This is the first stage of the Algorithm. Normalize all sensor values to a common scale since it is important for stable and fast convergence during training of the model as shown in Eq. (2).

$$\ddot{x_I} = \frac{xi-\mu}{\sigma} \tag{2}$$

The data is reshaped into 2D matrices $l \epsilon R^{HXW}$, where the rows generally indicate time steps, and the columns hold readings from various sensors or node features. The datasets is then split into training and test datasets after reshaping so that the model can learn with one subset and test its predictions with the other.

As given below in the CNN module, every reshaped 2D matrix from the training data is fed into a consecutive series of convolutional layers. Then, with the final model, utilize convolutional layers to highlight surroundings changes, faults, or interferences. This will also use ReLU (Rectified Linear Unit) activation to add non-linearity and then followed max-pooling to down-sample the feature maps, retaining the most important features. A final flattened one-dimensional feature vector that contains a lot of spatial information is then generated as the output of the CNN block as shown in Eq. (3).

$$F_{\{i,j\}}^{(l)} = \sigma\left(\sum_{a=0}^{k-1}\sum_{b=0}^{k-1} W_{\{a,b\}}^{(l)} \cdot X_{\{i+a,j+b\}}^{(l-1)} + b^{(l)}\right) \tag{3}$$

Here, $F_{\{I,j\}}^{(l)}$ is an output feature map at position (i,j) in layer l, $W_{\{a,b\}}^{(l)}$ is a learnable kernel weights, σ is an activation function i.e. ReLU: σ(x) = max(0, x), $b^{(l)}$ is a bias term and k is a kernel size. The pooling operation is used here is max pooling that reduces dimensionality as shown in Eq. (4).

$$P_{\{i,j\}} = \max_{\{(a,b)\in \text{window}\}} F_{\{i+a,j+b\}} \tag{4}$$

The second step is a BiLSTM module, which is aimed to learn temporal dependencies from the sequential feature data. The output features of the CNN are reshaped into a time-series this is suitable for LSTM processing. We make the assumption that the input data is sequential, meaning that the data that we feed into LSTMs have a time relation or dependency. At each time step, the forward and backward hidden states are concatenated to achieve a complete temporal feature representation as shown in Eqs. (5)-(11).

$$h_t = \text{LSTM}_{f(f_t, \overleftarrow{h_{t-1}})} \tag{5}$$

$$f_t = \sigma\left(W_f \cdot \left[h_{\{t-1\}}, x_t\right] + b_f\right) \tag{6}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{\{t-1\}}, x_t\right] + b_i\right) \tag{7}$$

$$o_t = \sigma\left(W_o \cdot \left[h_{\{t-1\}}, x_t\right] + b_o\right) \tag{8}$$

$$\tilde{c}_t = \tanh\left(W_c \cdot \left[h_{\{t-1\}}, x_t\right] + b_c\right) \quad (9)$$

$$c_t = f_t \odot c_{\{t-1\}} + i_t \odot \tilde{c}_t \quad (10)$$

$$h_t = o_t \odot \tanh(c_t) \quad (11)$$

Eq. (12) shows the backward LSTM working procedure in reverse:

$$\mathrm{h_t} = \mathrm{LSTM}_{\mathrm{b}(\mathrm{f_t}\overleftarrow{\mathrm{h_t}}+1)} \quad (12)$$

Eq. (13) shows the BiLSTM output at time t with forward pass $\overrightarrow{\mathrm{h_t}}$ and backward pass $\overleftarrow{\mathrm{h_t}}$.

$$\mathrm{h_t} = \left[\overrightarrow{\mathrm{h_t}}; \overleftarrow{\mathrm{h_t}}\right] \in \mathbb{R}^{\{2d\}} \quad (13)$$

These temporal features are then fed forward through a dense output layer forming the classification unit. The output is then mapped to probability (0, 1) using sigmoid activation function. This value indicates the probability of packet loss for the current input sample. If its output value is equal to or exceeds 0.5, this output is labelled as packet loss (label = 1); else this output is labelled as successful transmission (label = 0). The final hidden state h_T (concatenated forward and backward states) is fed into a fully connected layer, sigmoid function for binary classification, softmax layer for multilayer classification and cross entropy in Eqs. (14)-(17) respectively.

$$z = W_o h_T + b_o \quad (14)$$

$$\hat{\mathrm{y}} = \sigma(\mathrm{z}) = \frac{1}{(1 + \mathrm{e}^{\{-\mathrm{z}\}})} \quad (15)$$

$$\hat{\mathrm{y}}_i = \frac{e^{\{z_i\}}}{\sum_j e^{\{z_j\}}}, for\ i = 1, \dots, C \quad (16)$$

$$\mathrm{L_{CCE}} = -\sum_{\{\mathrm{i}=1\}}^{\mathrm{C}} \mathrm{y_i} \log(\hat{\mathrm{y}}_\mathrm{i}) \quad (17)$$

Eq. (18) shows the model is trained using Adam optimizer to minimize loss with θ for all learnable parameters of CNN, BiLSTM, and output layer, and η as the learning rate.

$$\theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathrm{L} \quad (18)$$

The final output of the algorithm are the predicted labels for the test dataset. The predicted status can then be used to notify the network controller to take proactive remedial actions, such as rerouting, transmission power adjustment, etc. to reduce packet loss in real-time. In summary, the goal of this hybrid model is to combine the intelligent extraction of spatial patterns with temporal sequence modeling, providing an effective approach for real-time, data-driven packet loss prediction in complex IoT-enabled WSN environments.

4. RESULTS AND DISCUSSION

4.1 Simulation environment

A series of experiments were conducted to evaluate the performance of the proposed Hybrid CNN-BiLSTM in terms of predicting and reducing packet drops in IoT-enabled WSNs on the overall (on-model) packet drops. These experiments were performed in a custom simulation environment built in Python, utilizing Tensorflow for model implementation and NS-3 (Network Simulator 3) for communication protocols and network simulation for WSN. All the tools together enabled an accurate model of both, the learning process, and the WSNs physical-layer dynamics.

NS-3 simulator is used to simulate realistic WSN scenarios like mobility of nodes, packet transmission, routing protocols (AODV, DSDV, etc.), impacts of interference, range of communication, etc. Important parameters for an actual WSN, namely node transmission power, buffer size, link failure probability, and congestion behavior, were configured accordingly in the simulation. Python bindings and trace file parsing were used to integrate with Python–NS-3, which facilitates real-time data feed into TensorFlow-based hybrid model for learning and inference. The CNN-BiLSTM The model was implemented using TensorFlow 2 using Keras APIs. Model was trained on a dataset which has been obtained from NS-3 trace logs with packet delivery time, RSSI, buffer occupancy, hop count, and transmission delay features. Training Data and Setup Training was performed on a CUDA-compatible system equipped with an NVIDIA RTX 3080 GPU, which facilitated the processing of large-scale datasets and deep network parameters.

In this work, a Synthetic dataset was being generated, to reflect various WSN layouts and situations that cover a wide range of environmental and operational aspects. The dataset included different node densities 100, 150, 200, 300, 400, and 500 nodes emulating various deployment scales. Furthermore, various network topologies, such as star, grid, tree, and random node distributions were simulated for robustness against different structural configurations. For evaluating the performance in the scenarios where the nodes are static and mobile, so mobility patterns were defined to cover both node scenarios, while the transmission rates varied between 50 kbps and 500 kbps, considering other data traffic intensities. For generating realistic traffic behavior, a Poisson distribution was used to model the packet inter-arrival times, resulting in stochastic but representative communication behaviors.

For each of the simulated network scenarios, the simulation lasted 1000 seconds, during which key features, including signal strength, distance between nodes, congestion levels, and routing paths were monitored at regular intervals. At the same time, packet loss events were logged as the labelled outputs, creating a supervised learning paradigm for the model. The output dataset formed more than 100K labelled samples and was divided into training, validation and test sets. This large dataset allowed for comprehensive testing of the proposed hybrid CNN-BiLSTM model performance ensuring its generalization across a diverse set of network conditions and improving its prediction capability for real-world deployments of wireless sensor networks.

4.2 Network level evaluation

Packet Loss Ratio: In WSNs, outage of the node which leads to data loss will lead to a major drop in the performance of the network; also, in general as the availability of data drops, the Packet Loss Ratio (PLR) can be an important indicator of the overall performance of the network. Mathematically, it is defined as Eq. (19). The performance comparison is outlined in Figure 1.

$$PLR = \frac{Number\ of\ Packets\ Lost}{Total\ number\ of\ Packets\ Sent} X\ 100\% \quad (19)$$

It assists in quantifying the frequency at which packets do not reach their destination. High PLR signifies poor network condition, which is highly undesired from the perspective of an IoT-based system thriving on accurate as well as timely transmission of data. The hybrid DL model that has been proposed uses CNN to extract spatial features and BiLSTM to learn temporal sequence. By learning how the network behaves in the past as well as the spatial data from the sensor that are available, the objective is to predict/ prevent the packets loss events in a real-time manner.

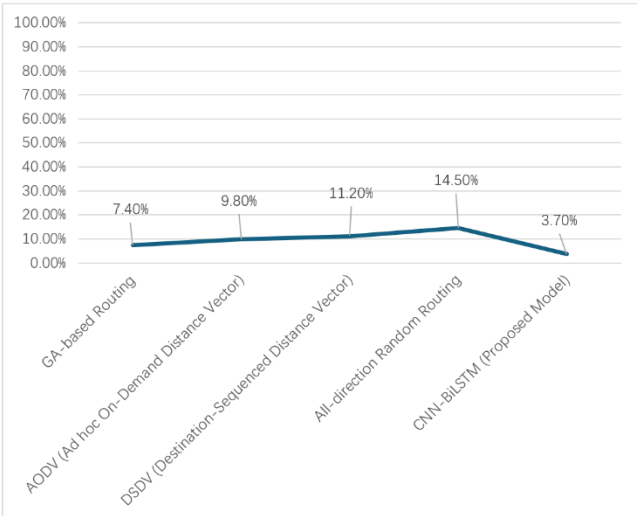

Figure 1. Packet loss ratio

Energy Consumption: Energy consumption is the very important evaluation metric for the sensor network to enhance the network lifetime. Indicating the power consumed, in millijoules per node (mJ/node), this metric accounts for the average power spent on transmitting, receiving, idle listening, and computation. We extended NS-3's energy module to include model computation costs incurred at the edge. The performance comparison is outlined in Figure 2.

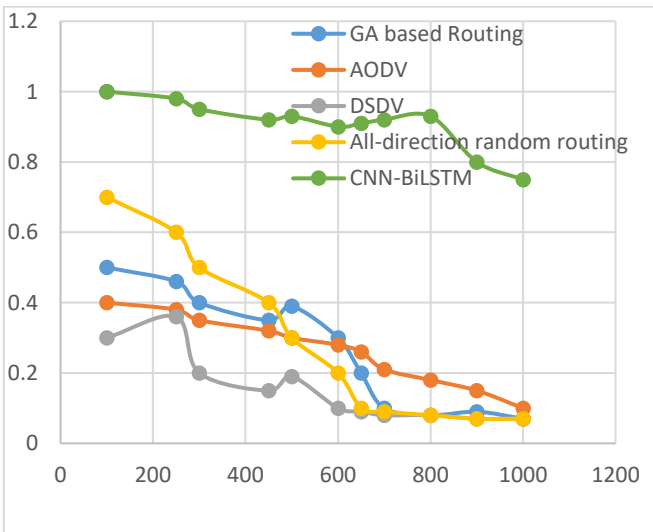

Figure 2. Energy consumption v/s Number of iterations

Throughput: An average successful data delivery rate, measured in kilobits per second (kbps), through all of its nodes. It is computed as Eq. (20). As outlined in Figure 3, the proposed CNN-BiLSTM model achieved a reduction in packet loss of up to 52% (mean ± SD: 3.7% ± 0.4) and consistently outperformed the baseline approaches. Trial results displayed narrow error bars confirming that overall performance remained stable despite changing network dynamics.

$$Throughput = \frac{\text{Total Successfully Received Data (bits)}}{\text{Simulation Time (s)}} \quad (20)$$

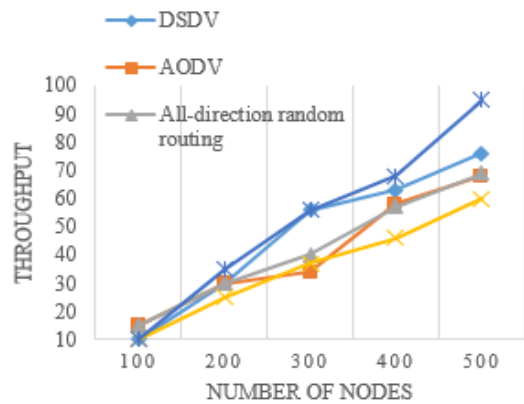

Figure 3. Throughput v/s Number of nodes

Latency: Latency in WSNs is defined as the average time taken for a data packet to travel from the source sensor node to its destination which is usually either a sink node or a central server. It is computed using Eq. (21). Where N is the number of packets successfully delivered, $t_{send}^{(i)}$ is the time the packet i was transmitted, $t_{receive}^{(i)}$ is the time it was received at the destination. Latency comparison is outlined in Figure 4.

$$Latency_{avg} = \frac{1}{N}\sum_{i=1}^{N}(t_{receive}^{(i)} - t_{send}^{(i)}) \quad (21)$$

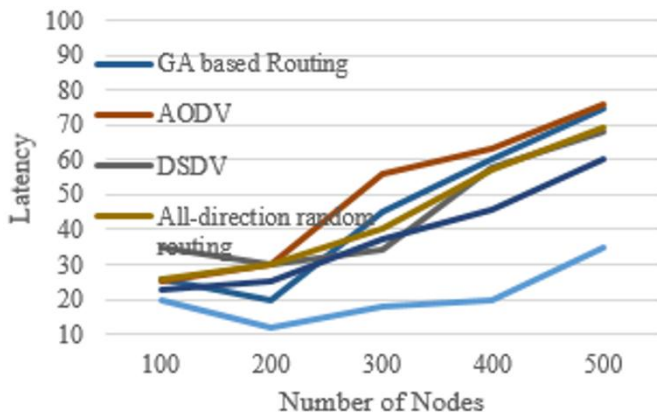

Figure 4. Latency v/s Number of nodes

In this study, the NS-3 simulation environment was used to measure the latency, which allows for precise control and realistic representation of network elements. By predicting where transmissions will bottleneck and strategically routing data, the hybrid CNN-BiLSTM model indirectly reduces latency by ensuring faster and more efficient routing decisions. In the hybrid CNN-BiLSTM model, the amount of packets lost was less than it was in conventional AODV routing and standalone LSTM-based methods. There was also increased energy efficiency due to reduced retransmissions. The results

show that the model can learn effective routing strategies from historical data and adapt to changing network conditions.

Table 1. Performance comparison of hybrid model with based models

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Training Time (s/Epoch)	Memory Usage (MB)
CNN-BiLSTM (Proposed)	96.3%	94.8%	95.6%	95.2%	0.978	11.3	54
LSTM Only	91.5%	89.2%	88.6%	88.9%	0.936	10.4	50
CNN Only	88.2%	85.4%	84.1%	84.7%	0.912	9.1	42
Random Forest	82.6%	78.5%	80.3%	79.4%	0.851	-	-
SVM (RBF Kernel)	79.7%	75.9%	74.6%	75.2%	0.839	-	-

4.3 Model level evaluation

The proposed hybrid model is evaluated to check its performance using accuracy, precision, recall, F1-Score, and AUC-ROC. Four baseline models were used to evaluate the performance of the proposed CNN-BiLSTM hybrid model using standard classification metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. The results demonstrate that compared to traditional models, the hybrid architecture clearly has superior performance in predicting packet loss in IoT-enabled WSNs. This is primarily due to the complementary nature of the components of the model CNN layers learn spatial features from input from the sensor in an efficient way, while the BiLSTM layers learn more complex temporal dependencies. This allows the model to jointly gather information relevant to packets on whether they will be delivered or not. With the highest accuracy of 96.3%, CNN-BiLSTM model showed promising results and its ability to generalised well with different WSN configurations and conditions.

The proposed model also surpasses other models in precision and recall, achieving 94.8% and 95.6%, respectively. This balance indicates that the model is identifying true positives well packets predicted as successful that are indeed successful and is also reducing false positives i.e., predicting success when there is actually loss. The F1-score was 95.2%, a clear indication of the stability and reliability of our model on detection and classification tasks. Also, the AUC-ROC score of 0.978, wherein the CNN-BiLSTM classifies the packets as either being lost or delivered across all threshold values, indicates excellent representation of the two classes. The high AUC value confirms that regardless of the operating conditions, the model exhibits excellent discrimination power.

The model built only using LSTMs attained 91.5% accuracy but had marginally lower precision and recall scores 89.2% and 88.6% respectively. It means that while it can fit temporal features and it misses the spatial correlation captured by the convolutional neural network part. The model that used only CNN fell further behind with an 88.2% accuracy and less F1-score because traditional CNN does not learn sequential dependencies which are prevalent in WSN environments. The traditional ML methods like Random Forest and SVM with an RBF kernel achieved the worst results as they achieved accuracies of 82.6% and 79.7%, respectively. Their lower AUC-ROC scores 0.851 and 0.839 highlight the superiority of using DL based sequential models in dealing with complicated real-time WSN data as shown in Table 1.

Results indicate that the combined hybrid model achieves both high predictive accuracy and also plays an on-the-level role in network advancements. The spatial information extracted by CNN layers significantly supplements the learning of temporal sequence in BiLSTM layers, allowing the system to effectively learn complex dependencies in packet delivery patterns. Its generalizability is underscored by its robustness over different topologies, densities and transmission rates. This predictive paradigm forms a core part of a reactive mechanism that, when compared against traditional approaches in WSN routing protocols, represents a data-driven forward-looking approach in direct line with the redesigned paradigms of intelligent IoT systems.

5. CONCLUSION

This paper presented a hybrid CNN-BiLSTM model for minimizing packet loss in IoT-enabled WSNs. The combination of spatial and temporal learning enabled more robust predictions and routing decisions. Future work will focus on deploying this model on edge devices and testing in real-world IoT applications with dynamic mobility and environmental factors. This research presents a comprehensive DL-based approach for minimizing packet loss in IoT-enabled WSN through a hybrid CNN-BiLSTM architecture. By integrating CNN for spatial feature extraction with bidirectional LSTM for capturing temporal dependencies in transmission behavior, the proposed model demonstrates a significant improvement in predictive accuracy and network-level efficiency compared to traditional models and standalone neural architectures.

Experimental evaluations conducted using a custom simulation environment combining Python, TensorFlow, and NS-3 confirmed the robustness of the model across diverse WSN topologies, node densities, and transmission rates. Notably, the hybrid model achieved a packet loss ratio reduction of up to 52% and improved network throughput by 18.7%, while maintaining energy efficiency suitable for constrained IoT environments Although the proposed CNN-BiLSTM model achieves a significant reduction in terms of packet loss and energy efficiency, there are still some limitations. The first limitation is that the performance of the model has not been evaluated on ultra-dense WSN deployments (i.e., >500 nodes) due to the simulation. The second limitation is that although the model itself is lightweight, real-time inference capabilities on ultra-low-power microcontrollers (e.g., MSP430/ATmega328) has yet to be tested. Future work will include optimizing the model for embedded deployment, validating the model performance on real-world IoT testbeds, and extending the study to dynamic routing environments in the presence of mobile nodes.

REFERENCES

[1] Said, R.B., Sabir, Z., Askerzade, I. (2023). CNN-BiLSTM: A hybrid deep learning approach for network intrusion detection system in software-defined

networking with hybrid feature selection. IEEE Access, 11: 138732-138747. https://doi.org/10.1109/ACCESS.2023.3340142

[2] Raju, S.V.S.R.K., Thomas, A.K., Ramu, K., Pandey, R., Rao, B.M., Rachapudi, V., Harikumar, M., Kashyap, T. (2024). Optimizing packet processing in IoT-enabled Wireless Sensor Networks: A novel data mining approach. Mathematical Modelling of Engineering Problems, 11(11): 3192-3200. https://doi.org/10.18280/mmep.111129

[3] Larouci, N.E.H., Sahraoui, S., Djeffal, A. (2025). Machine learning based routing protocol (MLBRP) for Mobile Internet of Things networks. Journal of Network and Systems Management, 33(3): 67. https://doi.org/10.1007/s10922-025-09949-6

[4] Guo, D., Duan, P., Yang, Z., Zhang, X., Su, Y. (2023). Convolutional neural network and bidirectional long short-term memory (CNN-BiLSTM)-attention-based prediction of the amount of silica powder moving in and out of a warehouse. Energies, 17(15): 3757. https://doi.org/10.3390/en17153757

[5] Abdallah, M., An Le Khac, N., Jahromi, H., Delia Jurcut, A. (2021). A hybrid CNN-LSTM based approach for anomaly detection systems in SDNs. In Proceedings of the 16th International Conference on Availability, Reliability and Security, pp. 1-7. https://doi.org/10.1145/3465481.3469190

[6] Zhang, Z., Mahmud, M.P., Kouzani, A.Z. (2022). FitNN: A low-resource FPGA-based CNN accelerator for drones. IEEE Internet of Things Journal, 9(21): 21357-21369. https://doi.org/10.1109/JIOT.2022.3179016

[7] Inayat, U., Zia, M.F., Mahmood, S., Khalid, H.M., Benbouzid, M. (2021). Learning-based methods for cyber attacks detection in IoT systems: A survey on methods, analysis, and future prospects. Electronics, 11(9): 1502. https://doi.org/10.3390/electronics11091502

[8] Jing, G., Wan, C., Dai, R. (2021). Angle-based sensor network localization. IEEE Transactions on Automatic Control, 67(2): 840-855. https://doi.org/10.1109/TAC.2021.3061980

[9] Ullah, Z., Al-Turjman, F., Mostarda, L., Gagliardi, R. (2020). Applications of artificial intelligence and machine learning in smart cities. Computer Communications, 154: 313-323. https://doi.org/10.1016/j.comcom.2020.02.069

[10] Wang, J., Tang, J., Xu, Z., Wang, Y., Xue, G., Zhang, X., Yang, D. (2017). Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach. In IEEE INFOCOM 2017-IEEE Conference on Computer Communications, Atlanta, GA, USA, pp. 1-9. https://doi.org/10.1109/INFOCOM.2017.8057090

[11] Salim, M.S., Sabri, N., Dheyab, A.A.R. (2025). Predicting WSN packet loss using machine learning: Applications in solid surroundings. Edelweiss Applied Science and Technology, 9(2): 289-302.

[12] Omarov, B., Sailaukyzy, Z., Bigaliyeva, A., Kereyev, A., Naizabayeva, L., Dautbayeva, A. (2023). One dimensional Conv-BiLSTMNetwork with Attention Mechanism for IoT intrusion detection. Computers, Materials & Continua, 77(3): 3765-3781. https://doi.org/10.32604/cmc.2023.042469

[13] Yuan, J., Peng, J., Yan, Q., He, G., Xiang, H., Liu, Z. (2023). Deep reinforcement learning-based energy consumption optimization for Peer-to-Peer (P2P) communication in wireless sensor networks. Sensors, 24(5): 1632. https://doi.org/10.3390/s24051632

[14] Omarov, B., Auelbekov, O., Suliman, A., Zhaxanova, A. (2023). Cnn-bilstm hybrid model for network anomaly detection in internet of things. International Journal of Advanced Computer Science and Applications, 14(3): 436-444. https://doi.org/10.14569/IJACSA.2023.0140349

[15] Rajalakshmi, V., Ganesh Vaidyanathan, S. (2022). Hybrid CNN-LSTM for traffic flow forecasting. In Proceedings of 2nd International Conference on Artificial Intelligence: Advances and Applications. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-6332-1_35

[16] Latif, S., Driss, M., Boulila, W., Huma, Z.E., Jamal, S.S., Idrees, Z., Ahmad, J. (2020). Deep Learning for the industrial Internet of Things (IIoT): A comprehensive survey of techniques, implementation frameworks, potential applications, and future directions. Sensors, 21(22): 7518. https://doi.org/10.3390/s21227518

[17] López-Ardao, J.C., Rodríguez-Rubio, R.F., Suárez-González, A., Rodríguez-Pérez, M., Sousa-Vieira, M.E. (2020). Current trends on green wireless sensor networks. Sensors, 21(13): 4281. https://doi.org/10.3390/s21134281

[18] Ullah Khan, S., Ulalh Khan, Z., Alkhowaiter, M., Khan, J., Ullah, S. (2024). Energy-efficient routing protocols for UWSNs: A comprehensive review of taxonomy, challenges, opportunities, future research directions, and machine learning perspectives. Journal of King Saud University - Computer and Information Sciences, 36(7): 102128. https://doi.org/10.1016/j.jksuci.2024.102128

[19] Elsayem, M., Abou-Zeid, H., Afana, A., Givigi, S. (2022). Reinforcement learning-based dynamic resource allocation for grant-free access. In GLOBECOM 2022-2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, pp. 1091-1096. https://doi.org/10.1109/GLOBECOM48099.2022.10001586

[20] Raman, R.N. (2023). Bandwidth Optimization Techniques for Faster Data Transfer Avoiding Traffic Congestion Using Distributed Bandwidth Network Section A-Research paper 12501 Eur. European Chemical Bulletin, 12(10): 12501-12508.

[21] Hong, D., Gao, L., Yokoya, N., Yao, J., Chanussot, J., Du, Q., Zhang, B. (2020). More diverse means better: Multimodal deep learning meets remote sensing imagery classification. IEEE Transactions on Geoscience and Remote Sensing, 59(5): 4340-4354. https://doi.org/10.1109/TGRS.2020.3016820

[22] Altunay, H.C., Albayrak, Z. (2023). A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks. Engineering Science and Technology, an International Journal, 38: 101322. https://doi.org/10.1016/j.jestch.2022.101322

[23] Sadhwani, S., Khan, M.A.H., Muthalagu, R., Pawar, P.M. (2024). BiLSTM-CNN hybrid intrusion detection system for IoT application. https://doi.org/10.21203/rs.3.rs-3820775/v1

[24] Dritsas, E., Trigka, M. (2025). Federated learning for IoT: A survey of techniques, challenges, and applications. Journal of Sensor and Actuator Networks, 14(1): 9.

https://doi.org/10.3390/jsan14010009
[25] Ghosh, S., Chaki, A., Kudeshia, A. (2021). Cyberbully detection using 1D-CNN and LSTM. In Proceedings of International Conference on Communication, Circuits, and Systems. Lecture Notes in Electrical Engineering, Springer, Singapore. https://doi.org/10.1007/978-981-33-4866-0_37
[26] Rajawat, A.S., Goyal, S.B., Bedi, P., Jan, T., Whaiduzzaman, M., Prasad, M. (2023). Quantum machine learning for security assessment in the Internet of Medical Things (IoMT). Future Internet, 15(8): 271. https://doi.org/10.3390/fi15080271
[27] Kaur, N., Gupta, L. (2024). Securing the 6G–IoT environment: A framework for enhancing transparency in artificial intelligence decision-making through explainable artificial intelligence. Sensors, 25(3): 854. https://doi.org/10.3390/s25030854
[28] Xu, X., Li, H., Xu, W., Liu, Z., Yao, L., Dai, F. (2021). Artificial intelligence for edge service optimization in internet of vehicles: A survey. Tsinghua Science and Technology, 27(2): 270-287. https://doi.org/10.26599/TST.2020.9010025
[29] Tseng, S., Wang, Y., Wang, Y. (2024). Multi-class intrusion detection based on transformer for IoT networks using CIC-IoT-2023 dataset. Future Internet, 16(8): 284. https://doi.org/10.3390/fi16080284
[30] Ahanger, T.A., Ullah, I., Algamdi, S.A., Tariq, U. (2025). Machine learning-inspired intrusion detection system for IoT: Security issues and future challenges. Computers and Electrical Engineering, 123: 110265. https://doi.org/10.1016/j.compeleceng.2025.110265