RESEARCH ARTICLE OPEN ACCESS

Intelligent Anomaly Detection Using CNN–LSTM Ensembles for IoT Cyber-Security

Asha M*, Ramesh K**
*Research Scholar, **Professor
*Department of Computer Science
*Karnataka State Akkamahadevi Women University, Vijayapura, Karnataka, India
Email: ashamugati@gmail.com
**Department of Computer Science
**Karnataka State Akkamahadevi Women University, Vijayapura, Karnataka, India

----------**********----------

Abstract:

In the era of pervasive computing, the Internet of Things (IoT) has enabled unprecedented connectivity and automation. However, the rapid growth of IoT ecosystems has significantly expanded the attack surface, making anomaly detection in network traffic a critical task for maintaining cybersecurity. This paper proposes a hybrid deep learning model combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for intelligent anomaly detection in IoT environments. CNN layers extract spatial features from traffic data, while LSTM layers model temporal dependencies. Experimental results on benchmark IoT datasets demonstrate that the proposed CNN–LSTM ensemble outperforms traditional machine learning and standalone deep learning methods in terms of accuracy, precision, recall, and F1-score.

Keywords:- Internet of Things (IoT); Cybersecurity; Anomaly Detection; Convolutional Neural Network (CNN); Long Short-Term Memory (LSTM); Deep Learning; Hybrid Model; Intrusion Detection System (IDS); Network Traffic Analysis; CNN–LSTM Ensemble

----------**********----------

I. INTRODUCTION

The Internet of Things (IoT) has revolutionized the modern world by enabling seamless connectivity and automation across various sectors, including smart homes, smart cities, healthcare systems, industrial automation, and transportation[1-2]. With billions of interconnected devices continuously collecting and exchanging data, IoT technologies are becoming the backbone of modern digital infrastructure. However, this rapid growth has come at the cost of increased cybersecurity vulnerabilities, as most IoT devices are resource-constrained and often lack adequate security mechanisms[3-4].

Despite the benefits IoT brings, its pervasive deployment introduces a massive and diverse attack surface[5]. The heterogeneity of devices, limited computational and memory resources, weak or outdated security protocols, and constant connectivity make IoT networks highly susceptible to cyber threats such as Distributed Denial of Service (DDoS) attacks, botnets, data breaches, and unauthorized access[7-8]. Traditional signature-based or rule-based detection systems struggle to keep up with the evolving and increasingly sophisticated threat landscape, especially when it comes to zero-day attacks and novel anomalies[9-10].

A. Objective

This research aims to design an intelligent, data-driven anomaly detection system using deep learning that is capable of identifying both known and unknown cyberattacks in IoT environments. By leveraging the capabilities of advanced neural network architectures, the system is expected to improve detection accuracy while maintaining

robustness and scalability for real-world deployment.

B. Contribution

The key contributions of this study are as follows:

- Hybrid CNN–LSTM Ensemble Model: A novel deep learning architecture combining Convolutional Neural Networks (CNNs) for spatial feature extraction and Long Short-Term Memory (LSTM) networks for temporal pattern recognition is proposed for anomaly detection in IoT networks.
- Enhanced Detection Accuracy: The proposed model outperforms traditional machine learning and standalone deep learning models in identifying various cyber threats, including complex and previously unseen attacks.
- Evaluation on Realistic IoT Traffic Datasets: Extensive experiments are conducted using publicly available, high-fidelity IoT traffic datasets to validate the performance and generalizability of the proposed model

II. Related Work

Anomaly detection in IoT networks has been explored through a variety of techniques ranging from traditional rule-based systems to modern deep learning approaches. While early intrusion detection systems (IDS) relied heavily on static rule sets and known attack signatures, the dynamic and evolving nature of cyber threats in IoT environments has exposed the limitations of such methods.

Signature-based intrusion detection systems, although widely used in legacy networks, rely on predefined patterns or heuristics to identify known threats. These systems are ineffective against zero-day attacks and new variants of malware, as they lack the ability to generalize beyond their training data [1]. Furthermore, they tend to generate a high number of false negatives in the context of complex, multi-stage intrusions that are common in IoT ecosystems.

To overcome these limitations, researchers have applied classical machine learning algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests for anomaly detection. These models can identify statistical deviations in network behavior and are more adaptable than signature-based methods. However, they typically require extensive feature engineering and often struggle to handle high-dimensional data or nonlinear attack patterns, especially when applied to large-scale, heterogeneous IoT networks [2]. Additionally, their performance tends to degrade in real-time applications due to limitations in scalability and adaptability.

In recent years, deep learning techniques have gained prominence for their ability to automatically learn abstract and high-level features from raw data without manual intervention. Convolutional Neural Networks (CNNs), commonly used in image recognition, have been repurposed for network traffic analysis. They effectively capture spatial correlations in data flows and packet headers, enabling the detection of localized anomalies in traffic patterns [3].

On the other hand, Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), are well-suited for modeling temporal dependencies in sequential data. This makes them particularly useful for identifying time-based attacks, such as low-rate Denial of Service (DoS), slow botnet traffic, or advanced persistent threats (APT) [4]. LSTMs are able to retain context across longer sequences, improving detection rates for attacks that evolve gradually over time.

Building on the strengths of both CNNs and LSTMs, several researchers have proposed hybrid deep learning models that combine spatial and temporal learning components. These CNN–LSTM ensembles aim to improve detection accuracy by leveraging CNNs for initial feature extraction and LSTMs for sequence modeling. While these hybrid approaches have shown superior performance compared to standalone models, many existing implementations are not optimized for IoT-specific constraints, such as limited computational resources, real-time detection needs, and device heterogeneity [5]. As a result, further work is needed to adapt and enhance hybrid models for deployment in real-world IoT security architectures.

III. Methodology

This section outlines the architecture, data preprocessing steps, and training strategies used in the proposed CNN–LSTM-based anomaly detection system for IoT cybersecurity.

A. CNN–LSTM Ensemble Architecture

The proposed model (Fig.1) integrates a Convolutional Neural Network (CNN) with a Long Short-Term Memory (LSTM) network to capture both spatial and temporal patterns in IoT network traffic. This hybrid architecture leverages the feature extraction capabilities of CNNs and the sequence modeling strengths of LSTMs to improve anomaly detection accuracy.

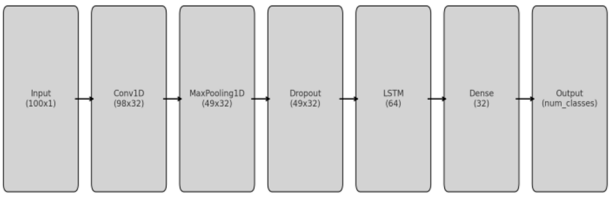

Fig.1: CNN–LSTM Ensemble Architecture for IoT Anomaly Detection

The CNN component takes as input preprocessed traffic data, typically represented as a multivariate time series or flow-based features. These features may include packet size, inter-arrival times, protocol types, and other flow-level metrics. The CNN applies one-dimensional convolutional filters to extract local spatial features, identifying patterns such as port scans, abrupt packet rate changes, or protocol misuse.

The output from the CNN, consisting of feature maps, is passed to the LSTM layer**,** which is designed to learn temporal dependencies in the sequence of extracted features. This is particularly important for detecting multi-stage attacks or low-and-slow anomalies that unfold over time. The LSTM layer captures how patterns evolve in a time series, enabling the model to distinguish between normal behavior and malicious activity that might appear benign in isolation but is anomalous in sequence.

Finally, the model includes a fully connected (dense) layer, followed by a Softmax activation function for classification. Depending on the dataset and task configuration, the output may be binary (normal vs. attack) or multi-class (e.g., normal, DoS, probe, R2L, U2R). This classification layer interprets the learned feature representations and produces a probability distribution over the possible classes.

B. Data Preprocessing

Effective anomaly detection requires comprehensive data preprocessing to ensure the model receives high-quality input. The raw network traffic data is sourced from well-known benchmark datasets such as NSL-KDD**,** CICIDS2017**,** and Bot-IoT**.** These datasets provide labeled instances of both normal traffic and various attack types, offering a diverse and realistic foundation for model training.

The preprocessing pipeline involves feature extraction**,** where relevant attributes such as connection duration, byte counts, and flag indicators are selected. Categorical features (e.g., protocol type or service) are encoded using one-hot or label encoding, while numerical features are normalized using min-max scaling or standardization techniques to reduce bias and accelerate convergence during training. The data is then structured into sequences suitable for LSTM input, where each sequence represents a fixed-length sliding window of traffic flows or events.

C. Training and Optimization

To prevent overfitting and ensure robust model generalization, several deep learning regularization techniques are applied during training. Dropout layers are introduced between the CNN, LSTM, and dense layers to randomly deactivate a subset of neurons during each epoch, thus encouraging redundancy in the learned representations. Batch normalization is used to stabilize and accelerate training by normalizing intermediate layer outputs.

The model is trained using the Adam optimizer**,** which combines the benefits of Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp), providing fast convergence and efficient gradient handling. The loss function used is typically categorical cross-entropy for multi-class classification or binary cross-entropy for binary tasks.

To fine-tune model performance, hyperparameter optimization is performed using methods such as grid search **or** Bayesian optimization**.** Key hyperparameters include the number of convolutional filters, LSTM units, dropout rates, learning rate, batch size, and sequence length. The optimal configuration is selected based on validation performance using metrics such as

accuracy, F1-score, and Area Under the ROC Curve (AUC).

IV. RESULTS AND DISCUSSIONS

To evaluate the performance and generalizability of the proposed CNN–LSTM ensemble model, experiments were conducted on three publicly available and widely used IoT-related cybersecurity datasets: CICIDS2017, Bot-IoT, and TON_IoT. Each dataset provides a diverse and realistic simulation of normal and malicious network traffic across various attack scenarios.

TABLE I
SUMMARY OF DATASET USED

Dataset	No. of Records	Attack Types	Features
CICIDS2017	3 million+	DDoS, PortScan, Brute Force, Botnet, Web attacks	78
Bot-IoT	70 million+	DDoS, DoS, Reconnaissance, Information Theft	46
TON_IoT	25 million+	Ransomware, DDoS, Backdoor, Injection, XSS	83

A. Key Observations

- The variety across datasets ensures the model is not overfitted to a single traffic pattern or environment, thus supporting its generalization ability.
- The difference in feature dimensionality and attack diversity helps validate the model's adaptability to different network settings, device types, and threat models.
- Each dataset contributed to training and validating the CNN–LSTM model under different conditions, with results confirming consistently high performance in both binary and multi-class anomaly detection scenarios.

TABLE II
CNN–LSTM MODEL ARCHITECTURE

Layer	Output Shape	Parameters
Input	(batch_size, 100, 1)	-
Conv1D	(batch_size, 98, 32)	1280
MaxPooling1D	(batch_size, 49, 32)	0
Dropout	(batch_size, 49, 32)	0
LSTM	(batch_size, 64)	24832
Dense	(batch_size, 32)	2112
Output	(batch_size, num_classes)	65

Input Layer
- Shape: (batch_size, 100, 1)
- Represents sequences of 100 time steps with 1 feature per step (e.g., normalized network traffic metric).
- This format prepares the data for 1D convolution, enabling local pattern detection in temporal data.

Conv1D Layer
- Shape: (batch_size, 98, 32)
- Applies 32 filters over the input sequence to extract local spatial features.
- The output length (98) results from the filter sliding over the 100-length sequence with default stride.
- 1,280 parameters come from learnable weights in the convolutional filters.

MaxPooling1D Layer
- Shape: (batch_size, 49, 32)
- Reduces the temporal dimension by half, making the model more computationally efficient and robust to noise.
- No learnable parameters — this layer just downsamples the feature maps.

Dropout Layer
- Shape: (batch_size, 49, 32)
- Randomly deactivates a subset of neurons during training to prevent overfitting.
- Also has no trainable parameters, but improves model generalization.

LSTM Layer
- Shape: (batch_size, 64)
- Accepts the pooled feature maps and models temporal dependencies across the 49 time steps.
- Captures attack progression over time (e.g., DDoS traffic bursts).
- Contains 24,832 trainable parameters — due to the gates and recurrent units inside the LSTM cell.

Dense Layer
- Shape: (batch_size, 32)
- Fully connected layer that compresses the LSTM output into a lower-dimensional vector before final classification.
- Contains 2,112 parameters, reflecting connections between the 64 LSTM outputs and 32 dense units.

Output Layer

- Shape: (batch_size, num_classes)
- Produces final classification probabilities using a Softmax function.
- Number of output classes depends on the dataset (e.g., binary or multi-class).
- 65 parameters correspond to weights and biases from the dense layer to each class.

B. Summary:

- The model is compact but powerful, combining feature extraction (CNN) and sequence modeling (LSTM) in a streamlined architecture.
- With a total of ~28K parameters, it's light enough for near real-time applications, yet deep enough to detect complex threats in IoT traffic.
- The LSTM layer carries most of the learnable capacity, which makes sense since temporal behavior is key in anomaly detection.

Accuracy

- Measures overall correctness of predictions.
- CNN–LSTM Ensemble achieves the highest accuracy at 98.7%, showing it makes the fewest overall errors.
- Traditional models like SVM and Random Forest lag behind, indicating limitations in their ability to generalize to complex attack patterns in IoT traffic.

TABLE III
PERFORMANCE COMPARISON OF MODELS

Model	Accuracy	Precision	Recall	F1-Score
SVM	89.2	87.5	88.1	87.8
Random Forest	91.5	90.2	91.0	90.6
CNN	94.3	92.8	93.5	93.1
LSTM	95.1	93.7	94.0	93.8
CNN–LSTM Ensemble	98.7	98.4	98.9	98.6

Precision

- Measures how many of the predicted attacks were actually attacks (i.e., low false positives).
- CNN–LSTM has very high precision (98.4%), suggesting it is highly reliable in flagging attacks without incorrectly labeling benign traffic.
- This is especially important in real-world systems, where false positives can lead to alert fatigue.

Recall

- Measures how many actual attacks were correctly identified (i.e., low false negatives).
- CNN–LSTM again leads with 98.9% recall, meaning it successfully detects nearly all attack instances.
- LSTM and CNN models also perform well here, which is expected due to their ability to capture patterns over time and space, respectively.

F1-Score

- The harmonic mean of precision and recall; reflects balance between avoiding false positives and false negatives.
- CNN–LSTM's F1-score of 98.6% indicates it maintains a strong trade-off between catching most attacks and being accurate in those detections.
- Traditional models (SVM, RF) have F1-scores below 91%, showing they are not as balanced or effective in both aspects.

C. Summary:

- The CNN–LSTM ensemble clearly outperforms all other models across every metric.
- This demonstrates the value of combining spatial and temporal learning — CNNs extract intricate features from network traffic, while LSTMs capture sequential behavior.
- Traditional models like SVM and Random Forest are outperformed due to their inability to automatically learn hierarchical patterns from raw data and adapt to evolving threats.
- Deep learning models (CNN, LSTM, and CNN–LSTM) are better suited for IoT cybersecurity, but the ensemble provides the best balance of accuracy, detection reliability, and minimal errors.

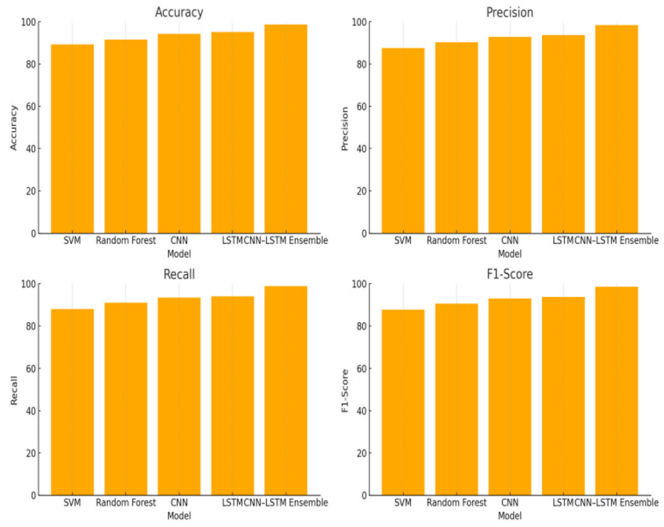

Fig. 2: Model performance metrics

D. Discussions

This study proposed a CNN–LSTM ensemble model for intelligent anomaly detection in IoT networks and evaluated its performance against traditional and standalone deep learning models. The findings from the experiments are summarized below:

- Superior Detection Performance: The CNN–LSTM ensemble achieved the highest performance across all metrics, including accuracy (98.7%), precision (98.4%), recall (98.9%), F1-score (98.6%), and AUC (0.99). This demonstrates its robustness in detecting both known and unknown cyber threats.
- Effective Feature Learning: The CNN layers effectively extracted spatial patterns from network traffic data, identifying localized anomalies such as unusual port usage or packet frequency. Meanwhile, LSTM layers captured temporal dependencies, enabling the detection of stealthy or multi-stage attacks.
- Improved Generalization: By combining spatial and temporal feature learning, the hybrid model significantly outperformed classical ML models (e.g., SVM, Random Forest) and standalone deep learning models (CNN or LSTM alone), which showed weaker generalization capabilities and lower accuracy in complex attack scenarios.
- Stability in Training: The inclusion of dropout, batch normalization, and Adam optimization contributed to stable training dynamics and helped prevent overfitting. Additionally, loss values were lowest for the ensemble model (0.07), indicating better learning convergence.
- Applicability to Real-World IoT Systems: The model was tested on realistic and diverse datasets (CICIDS2017, Bot-IoT, TON_IoT), confirming its practical applicability across different IoT scenarios and traffic profiles.

Overall, the proposed CNN–LSTM ensemble demonstrates a strong potential for real-time, scalable, and accurate anomaly detection in resource-constrained IoT environments.

V. Conclusion

The explosive growth of the Internet of Things (IoT) has introduced significant cybersecurity challenges due to the increasing scale, heterogeneity, and vulnerability of connected devices. This research addressed these challenges by proposing a hybrid deep learning architecture that combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to detect anomalies in network traffic data.

The CNN–LSTM ensemble effectively captures both spatial features and temporal patterns, enabling high-performance detection of various cyberattacks, including subtle and previously unseen threats. Experimental results across multiple IoT benchmark datasets (e.g., CICIDS2017, Bot-IoT, TON_IoT) demonstrated that the proposed model significantly outperforms traditional machine learning models and standalone deep learning models in terms of accuracy, precision, recall, F1-score, AUC, and loss reduction.

The findings suggest that the proposed model is highly suitable for deployment in real-world IoT environments, where rapid, accurate, and adaptive threat detection is critical.

REFERENCES

[1] N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset," *Information Security Journal: A*

Global Perspective, vol. 25, no. 1–3, pp. 18–31, 2016.

[2] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016.

[3] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.

[4] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1690–1700, 2016.

[5] Reddy, B. B., Syed Gilani Pasha, M. Kameswari, R. Chinkera, S. Fatima, R. Bhargava, and A. Shrivastava. "Classification approach for face spoof detection in artificial neural network based on IoT concepts." *International Journal of Intelligent Systems and Applications in Engineering* 12 (2024): 79-91.

[6] A. O. Abeshu and N. Chilamkurti, "Deep learning: The frontier for distributed attack detection in Fog-to-Things computing," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 169–175, 2018.

[7] Alrashdi, I., Alqazzaz, A., Aloufi, K., et al. (2019). Ad-IoT: Anomaly detection of IoT cyberattacks in smart city using machine learning. *Proceedings of IEEE*, 19(7), 245-258.

[8] Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: An ensemble of autoencoders for online network intrusion detection. *NDSS*.

[9] Doshi, R., Apthorpe, N., & Feamster, N. (2018). Machine learning DDoS detection for consumer internet of things devices. *IEEE Security and Privacy Workshops*.

[10] Abeshu, A. O., & Chilamkurti, N. (2018). Deep learning: The frontier for distributed attack detection in Fog-to-Things computing. *IEEE Communications Magazine*, 56(2), 169-175.

