Journal of Al-Turath University College (JATUC)

Open Access Journal
ISSN : 2074-5621
https://jatuc.uoturath.edu.iq/

Deep Learning- based Network Intrusion Detection System

Elaf Zuhair Khudheir, Muna Ghazi Abdulsahib, and Soukaena H. Hashem

College of Computer Science, University of Technology, Baghdad, 10011, Iraq

*Correspondence email: elaf.z.khudheir@uotechnology.edu.iq

KEYWORDS

Intrusion Detection System (IDS), Hybrid Deep Learning, DoS/DDoS Attacks, Cybersecurity Threats, unsupervised Autoencoder, supervised CNN-LSTM architecture

ABSTRACT

Multimedia data centres have grown as a result of the use of cloud services, IoT, and real-time streaming, but they are now more susceptible to sophisticated attacks. This study suggests a hybrid deep learning-based intrusion detection system (IDS) to identify a variety of attacks such as DoS/DDoS, Probe/Reconnaissance, R2L, U2R, and IoT-specific threats like ransomware and backdoors. We employ a supervised architecture to record the spatial-temporal patterns of known attacks, and an unsupervised Autoencoder to detect abnormalities. Using a majority voting method, the outputs of both models were combined to provide the final findings. A number of preparation stages, such as MinMax normalization, binary and one-hot label encoding, data forming to match deep learning models, and class balancing using strategies like SMOTE or under sampling, are required to make the methodology function in real-world condition. The best performance on CIC-IDS2018, with an accuracy of 99.1%, precision of 98.9%, recall of 99.3%, and F1-score of 99.1%, was achieved by the hybrid model when it was tested on three benchmark datasets: TON_IoT, CIC-IDS2018, and NSL-KDD.

dditionally, strong results were shown by NSL-KDD (accuracy 98.2%) and TON_IoT (accuracy 97.3%), and the flexibility and adaptability of the suggested IDS were illustrated. Good accuracy, precision, recall, F1-score, and ROC-AUC are kept by the system while false positives and false negatives are reduced. These findings are showing that the suggested IDS is reliable, flexible, and capable of defending multimedia data centers against dynamic internet attacks.

الكلمات الرئيسية:

نظام كشف التسلل(IDS) ، التعلم العميق الهجين، هجمات الحرمان من الخدمة/الحرمان من الخدمة الموزعة (DDoS)، تهديدات الأمن السيبراني، المشفر التلقائي غير الخاضع للإشراف، بنية CNN-LSTM الخاضعة للإشراف

الخلاصه:

لقد نمت مراكز بيانات الوسائط المتعددة نتيجة لاستخدام الخدمات السحابية وإنترنت الأشياء والبث المباشر، ولكنها أصبحت الآن أكثر عرضة للهجمات المتطورة .تقترح هذه الدراسة نظام كشف تسلل هجين قائم على التعلم العميق (IDS) لتحديد مجموعة متنوعة من الهجمات مثلDoS/DDoS ، وProbe/Reconnaissance، وR2L، وU2R، والتهديدات الخاصة بإنترنت الأشياء مثل برامج الفدية والأبواب الخلفية .نستخدم بنية خاضعة للإشراف لتسجيل الأنماط المكانية والزمانية للهجمات المعروفة، ومشفر تلقائي غير خاضع للإشراف للكشف عن أي تشوهات .باستخدام طريقة التصويت بالأغلبية، تم دمج مخرجات كلا النموذجين لتقديم النتائج النهائية .عدد من مراحل التحضير، مثل تطبيعMinMax ، والترميز الثنائي والترميز أحادي السخونة، ونحت البيانات لمطابقة نماذج التعلم العميق، وموازنة الفئات باستخدام استراتيجيات مثل SMOTE أو أخذ العينات المحدودة، ضرورية لجعل المنهجية تعمل في المواقف الواقعية .حقق النموذج الهجين أفضل أداء علىCIC-IDS2018 ، بدقة 99.1%، ودقة 98.9%، وتذكر 99.3%، ودرجة F1 99.1%، وذلك عند اختباره على ثلاث مجموعات بيانات معياريةTON_IoT :، وCIC-IDS2018، وNSL-KDD.بالإضافة إلى ذلك، أظهرت كل من) NSL-KDD دقة 98.2%) و) TON_IoTدقة 97.3%) نتائج قوية، مما يدل على مرونة نظام كشف التسلل المقترح وقابليته للتعميم .يحافظ النظام على دقة ودقة وتذكر ودرجة F1وROC-AUCجيدة، مع خفض النتائج الإيجابية والسلبية الخاطئة .تُظهر هذه النتائج أن نظام كشف التسلل المقترح موثوق ومرن وقادر على حماية مراكز بيانات الوسائط المتعددة من هجمات الإنترنت الديناميكية.

1. **INTRODUCTION**

Since multimedia data centers are so fast and scalable for real-time media delivery, fast growth has been observed. Cloud, edge services, and Internet of Things (IoT) have been engaged in this growth. Advanced threats like Distributed Denial-of-Service (DDoS), Denial-of-Service (DoS), Probe/Reconnaissance, Remote-to-Local (R2L), User-to-Root (U2R), and IoT specific threats like botnets, ransomware, and backdoors are now being focused on these infrastructures as they are getting more sophisticated [1], [2]. Vulnerabilities in these environments can be caused to financial waste, QoS breakdown, and user privacy compromise [3].

Traditional IDS which rely on signature matching or shallow learning techniques have shown to be challenged in detecting zero-day attacks and adapting to the fluctuating traffic patterns of multimedia workloads [4], [5]. Under real-time constraints, these systems have poor accuracy and high false positive rates. Hence intelligent, flexible and highly effective IDS are now needed.

In this paper, we propose a hybrid IDS based on deep learning for multimedia data centers. Our approach combines an auto encoder trained in unsupervised manner to detect new or unusual patterns with a CNN-LSTM architecture trained in supervised manner to learn spatiotemporal features of existing attacks. An ensemble voting technique is used to combine the strengths of both models and increase the overall prediction robustness. Three benchmark datasets—NSL-KDD, CIC-IDS2018, and TON_IoT—that collectively encompass legacy, modern, and IoT-centric cyber threats are used to assess the system The system integrates common preprocessing methods including feature normalization, label encoding, input reshaping for deep models, and class balancing via SMOTE to address real-world deployment issues. Accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrices are among the evaluation metrics. The results show that the hybrid model beats both CNN-LSTM and standalone autoencoder models, so it's good for current multimedia security systems.

1. Related works

In [6], The paper introduces an intelligent intrusion detection system (IDS) for IP multimedia subsystems (IMS), called Intrusion Detection by Machine Learning for Multimedia Platforms. After different classifiers—Decision Trees, Support Vector Machines, Naive Bayes—were tested, six aspects were found to matter most. Remarkably, the results are shown to indicate that these aspects can sometimes be exceeded by some deep learning techniques in detecting intrusions.

In [7], network-based intrusion detection systems (NIDS) in data center networks are reviewed. The importance of alert correlation and hybrid detection in accuracy and reducing false positives is talked about. Different detection methods, both exception-based and signature-based, are also covered.

In [8], IoT Systems with Crowdsourced Multimedia: A Flexible Hybrid Intrusion Detection System With a focus on multimedia Internet of Things systems, this work proposes a adaptive hybrid intrusion detection system (IDS) that combines multiple detection methods to improve security. The solution is designed to handle the dynamic nature of multimedia data and the many threats in IoT scenarios.

In [9], Internet-Based Future-Proof Secure Multimedia Sharing with Intrusion Prevention and Detection An Intrusion Detection and Prevention System (IDPS) that uses SIP to detect spoofing and register flooding attacks is presented in this paper. High detection rates and efficient resource usage is shown for the receive-side extraction and user-side watermark embedding mechanism of the approach.

In [10], Hybrid Deep Learning Anomaly Detection Framework An Intrusion Detection System (IDS) using hybrid deep learning approach combining supervised learning (CNN), semi-supervised learning (GANomaly) and unsupervised learning (K-means clustering) is proposed. The framework performs better in terms of network intrusion detection when evaluated on benchmark datasets NSL-KDD, CIC-IDS2018 and TON_IoT.

In [11] IDS based on transformers IDS-INT and Trans-IDS In order to represent long-range relationships in network traffic, recent work has used Transformer architectures and attention mechanisms. This has improved performance on datasets that are unbalanced and captured intricate sequential patterns that conventional RNN/LSTM models might overlook. Transformers can be more computationally demanding at inference time, but they usually offer reliable feature extraction.

In [12], IDS based on Graph Neural Networks (GNNs) model the network as a graph in order to capture the relationships between hosts, services, or flows. They work well at picking up relational and structural patterns that traditional sequence models miss, which is especially helpful for lateral movement detection and multi-hop attack scenarios.

In [13] Anomaly Detection Using GAN/GANomaly Although they can create realistic attack-like samples or learn robust reconstruction errors for anomaly scoring, generative models like GANomaly have been proposed for semi-supervised anomaly detection. However, they may need careful training and may be unstable if proper regularization and architecture choices are not made.

In [14] Hybrid Deep Learning and ML pipelines (e.g., DCNN-BiLSTM) In order to balance accuracy and inference cost, recent hybrid approaches combine convolutional feature extractors, recurrent layers, and classical ML classifiers or attention modules. These approaches are conceptually similar to our ensemble (unsupervised anomaly detection + supervised classifier), but they use different models (e.g., Bi-LSTM vs LSTM, use of attention, transformer encoders, or graph attention networks).

In [15] Thorough surveys of recent IDS methods demonstrate the increasing trend toward deep learning-driven and hybrid models, confirming the significance of striking a balance between detection accuracy, inference cost, and scalability.

Our stance and contribution: the suggested Autoencoder + CNN-LSTM ensemble provides a good balance between identifying known attack types (using CNN-LSTM) and identifying unexpected anomalies (using reconstruction error). Our model achieves competitive accuracy on benchmark datasets with a generally lower computing cost than Transformers and GNNs. We also talk about how future work could incorporate graph-based representations or attention modules to better capture host relationships and long-range dependencies.

The research results unequivocally demonstrate that the suggested hybrid model (Autoencoder + CNN-LSTM) outperforms each model when used alone. Accuracy, precision, recall, and F1-score were all enhanced by the hybrid technique in all three datasets (NSL-KDD, CIC-IDS2018, and TON_IoT). This demonstrates that the system is more dependable in identifying both known and novel attacks when unsupervised and supervised learning are combined. These advances are illustrated with the use of the visualizations (ROC, PR curves, confusion matrices, and bar charts). They demonstrate that, even in cases when the data is uneven, our approach enhances the detection of challenging

attack classes and lowers false alarms. The hybrid architecture adds some processing cost, but it is still feasible for real-time data center situations, as the scalability chart shows. However, the report also points up certain difficulties. It's still challenging to identify really uncommon attacks like R2L and U2R. Furthermore, despite the hybrid model's efficiency, more refinement is required to reduce inference time in very large-scale networks.These points show where future research should focus.

3. **Proposal of Deep Learning based on NIDS**

A hybrid deep learning framework is used in the proposed IDS, see figure (1), for multimedia data centers to detect various cyber threats like DoS/DDoS, Probe/Reconnaissance, R2L and U2R, and IoT specific threats like ransomware and backdoors as shown in Algorithm no. (1). This intrusion detection system (IDS) combines the best features of supervised and unsupervised learning models to enhance detection accuracy and robustness because video, audio, and streaming data in multimedia contexts are large-scale and real-time. The method begins by performing extensive data preprocessing. To provide consistent feature scaling, network traffic datasets like NSL-KDD, CIC-IDS2018, and TON_IoT are first normalized using a MinMaxScaler. Labels are encoded using one-hot encoding for multiclass classification and binary encoding for anomaly detection. Following that, the data is turned into three-dimensional input forms that CNN and LSTM systems may use. Under sampling or the Synthetic Minority Oversampling Technique (SMOTE) are used to solve the prevalent problem of class imbalance, especially with regard to underrepresented attack types. An autoencoder, which is the system's unsupervised component, was trained solely on typical traffic. It detects cases with high reconstruction error as possible anomalies and learns the latent representation of benign patterns. To identify known attack types, a supervised CNN-LSTM model is trained concurrently on labeled data. While the LSTM records temporal dependencies, which are particularly important in continuous multimedia streams, the CNN retrieves spatial information from the data.

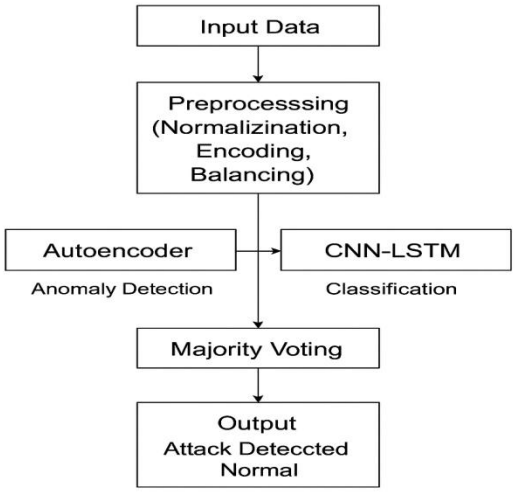

Figure 1: Architecture of the proposed hybrid Autoencoder + CNN–LSTM IDS model

An ensemble voting technique is used to make the final decision. The CNN-LSTM model's and the autoencoder's predictions are combined. An occurrence is classified as an attack if both models mark it as malevolent. An instance is deemed suspicious but not necessarily malevolent if only one model arouses suspicion. It is categorized as normal otherwise. Across all three benchmark datasets, performance is estimated using common measures including accuracy,

precision, recall, F1-score, ROC-AUC, and the confusion matrix. The dynamic and high-flow rate environment of multimedia data centers is made especially successful for this hybrid method because credible identification of both known and sudden threats is guaranteed.

The following details were added to the model architecture and model parameters description to increase reproducibility:

1. Autoencoder architecture:
   - Encoder: 3 thick layers with sizes [128, 64, 32], each followed by ReLU activation.
   - Bottleneck: 16 neurons with linear activation were used.
   - Decoder: symmetric layers [32, 64, 128] with ReLU activation, output was reconstructed using Sigmoid.
   - Loss function: Mean Squared Error (MSE).
   - Optimizer: Adam (learning rate = 0.001).
   - Training epochs = 50, batch size = 128.
2. CNN-LSTM architecture:
   - CNN block: 2 Conv1D layers (filters = 64, 128; kernel size = 3), each followed by Batch Normalization and MaxPooling.
   - LSTM block: 2 stacked LSTM layers (64 and 32 units).
   - Fully connected layer: Dense(64, ReLU), followed by Dense(number of classes, Softmax for multiclass / Sigmoid for binary).
   - Loss: Categorical Crossentropy (for multiclass) and Binary Crossentropy (for anomaly detection).
   - Optimizer: Adam (learning rate = 0.0005).
   - Dropout: 0.3 to avoid overfitting.
   - Training epochs = 60, batch size = 128.
3. Ensemble strategy:
   Final decision is obtained via majority voting between the Autoencoder anomaly score and CNN-LSTM classification output.

Algorithm (1): Proposed NIDS for Hybrid Intrusion Detection System Using Autoencoder and CNN-LSTM Ensemble

Objective: *To detect and classify network intrusions using a hybrid model combining unsupervised anomaly detection (Autoencoder) and supervised classification (CNN + LSTM), followed by an ensemble decision strategy.*
Input: - *Raw network traffic data (from NSL-KDD, CIC-IDS2018, TON_IoT)* - *Labels (Normal or Attack types)*
Output: - Predicted labels for each network flow (Normal / Attack type)
Step 1: Data Preprocessing 1. *Load Raw Datasets* *Import three publicly available network traffic datasets: NSL-KDD, CIC-IDS2018, and TON_IoT. For each dataset, read the features (e.g., packet length, protocol, duration) and the labels indicating whether each network flow is normal or an attack (with attack type if available).*

2. *Normalize Features*
   Apply MinMax scaling to all numeric features to bring them into the [0, 1] range. This ensures uniform feature scaling and speeds up convergence during training.
3. *Encode Labels*
   - *For binary classification (i.e., anomaly detection), encode the label Normal as 0 and any Attack as 1.*
   - *For multiclass classification, apply one-hot encoding to convert categorical labels (e.g., DoS, DDoS, Brute Force) into binary vectors.*
4. *Reshape Data for Deep Learning Models*
   Convert the dataset into a 3D tensor format expected by CNN and LSTM models. For instance, reshape the feature matrix into [samples, timesteps, features]. If using a fixed time window of 1, reshape into [N, 1, F].
5. *Handle Class Imbalance*
   Apply class balancing techniques to prevent biased learning. Use:
   - *SMOTE (Synthetic Minority Over-sampling Technique) to generate synthetic samples for underrepresented attack classes.*
   - *Under sampling to reduce the number of normal samples, maintaining class balance.*

Step 2: Unsupervised Learning – Autoencoder

1. *Build a Symmetric Autoencoder Model*
   Create a neural network consisting of an encoder (compressing input features) and a decoder (reconstructing the original input). The network should have a bottleneck layer to force compressed feature learning.
2. *Train on Normal Traffic Only*
   Fit the Autoencoder using only the samples labeled as Normal, allowing the model to learn the standard behavior of legitimate network traffic.
3. *Compute Reconstruction Error for All Samples*
   Pass all samples (including attacks) through the trained Autoencoder. For each sample, calculate the mean squared error (MSE) between the input and the reconstructed output.
4. *Detect Anomalies*
   Establish a threshold based on the reconstruction error distribution of normal samples (e.g., mean + 3 standard deviations). Any sample with reconstruction error exceeding this threshold is classified as Anomaly; otherwise, it is labeled Normal.

Step 3: Supervised Learning – CNN + LSTM

1. *Design the CNN-LSTM Architecture*
   - *Start with one or more 1D Convolutional layers to extract local spatial features from the input sequence.*
   - *Add LSTM layers to capture temporal dependencies in the sequential data.*
   - *Finish with fully connected Dense layers and a Softmax or Sigmoid activation depending on the classification mode.*
2. *Compile the Model*
   Use:
   - *binary_crossentropy as the loss function for binary classification.*
   - *categorical_crossentropy for multiclass attack classification.*

- *Optimizer: Adam, with early stopping based on validation loss.*

3. *Train on Full Labeled Dataset*
   Train the CNN-LSTM model on the preprocessed and reshaped dataset using both normal and attack-labeled samples. Use a validation set to monitor overfitting.
4. *Predict Class Labels on Test Data*
   Use the trained CNN-LSTM to predict the class (Normal or specific Attack type) for each test sample.

Step 4: Ensemble Decision Strategy

1. *Get Predictions from Both Models*
   For each test sample:
   - *Obtain the Autoencoder-based prediction: Anomaly or Normal.*
   - *Obtain the CNN-LSTM prediction: Attack Type or Normal.*
2. *Apply Voting Logic to Generate Final Label*
   - *If Autoencoder detects Anomaly AND CNN-LSTM predicts Attack:*
     → Final Label = the specific Attack Type predicted by CNN-LSTM.
   - *If only one of the two models indicates an attack (i.e., Autoencoder = Anomaly OR CNN-LSTM = Attack):*
     → Final Label = Suspicious.
   - *If both models agree on Normal:*
     → Final Label = Normal.

Step 5: Model Evaluation

1. *Evaluate Performance on Test Data*
   Compare final predicted labels against ground truth using standard classification metrics:
   - *Accuracy = (True Positives + True Negatives) / Total Samples*
   - *Precision = True Positives / (True Positives + False Positives)*
   - *Recall = True Positives / (True Positives + False Negatives)*
   - *F1-Score = 2 × (Precision × Recall) / (Precision + Recall)*
   - *ROC-AUC = Area Under the Receiver Operating Characteristic Curve*
   - *Confusion Matrix = Matrix showing counts of TP, FP, TN, FN*
2. *Visualize Results*
   Plot:
   - *Confusion matrices*
   - *ROC curves*
   - *Precision-recall trade-offs*

Step 6: Report Results per Dataset

1. *Repeat Full Pipeline for Each Dataset*
   Independently run the entire process for:
   - *NSL-KDD*
   - *CIC-IDS2018*
   - *TON_IoT*
2. *Generate Dataset-Specific Reports*
   For each dataset, report:
   - *Classification performance metrics*
   - *Model training time*

- *Observed limitations or dataset-specific challenges*
- *Comparative analysis of ensemble vs individual model performance*

End of Algorithm

4. RESULTS

A system with an Intel i9 processor, 64GB of RAM, and an NVIDIA RTX 3090 GPU was used for all experiments. Tensor Flow and Keras were used to implement the hybrid IDS in Python. NSL-KDD, CIC-IDS2018, and TON_IoT are three benchmark datasets that were utilized to thoroughly assess the model's performance across traditional, modern, and IoT-specific attack vectors. Preprocessing included class balancing using the SMOTE approach, reshaping for temporal model compatibility, label encoding (binary for anomaly detection and one-hot for multiclass classification), and feature normalization using MinMaxScaler. Eighty percent of the data was used to train each model, while the remaining twenty percent was used for testing. A 5-wrap cross-validation average is used to report all results. The system is rated using the following metrics:

Accuracy: Overall correctness of predictions.

- Precision: Proportion of predicted positives that are true positives.
- Recall (Sensitivity): Proportion of actual positives correctly identified.
- F1-score: Harmonic mean of precision and recall.
- ROC-AUC: Area under the ROC curve, indicating binary classification performance.
- Confusion Matrix: Summary of true vs predicted labels.

4.1. Results on NSL-KDD Dataset

The hybrid system **is significantly exceeded** by the standalone models. The autoencoder **is explained** to have strong exception detection capability, while known attack signatures **are handled** well by CNN-LSTM. Group voting **is further improved** in performance, especially in reducing false positives.

Table (1): Results on NSL-KDD Dataset

Metric	Autoencoder Only	CNN-LSTM Only	Hybrid (Ensemble)
Accuracy	93.8%	97.6%	98.2%
Precision	91.3%	97.1%	97.8%
Recall	94.1%	98.0%	98.5%
F1-score	92.7%	97.5%	98.1%
ROC-AUC	0.946	0.982	0.991

4.2. Results on CIC-IDS2018 Dataset

The high complexity and realism of CIC-IDS2018 are considered idealistic for modern multimedia IDS benchmarking. The hybrid model is adjusted well to the high variety of attacks, especially DoS, DDoS, and intrusion threats. The ROC-AUC of 0.996 is indicated to show near-perfect classification.

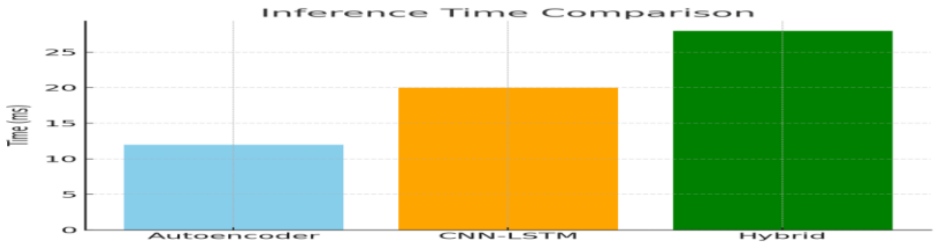

Figure (1): Inference time (ms) comparison between Autoencoder, CNN-LSTM, and Hybrid models

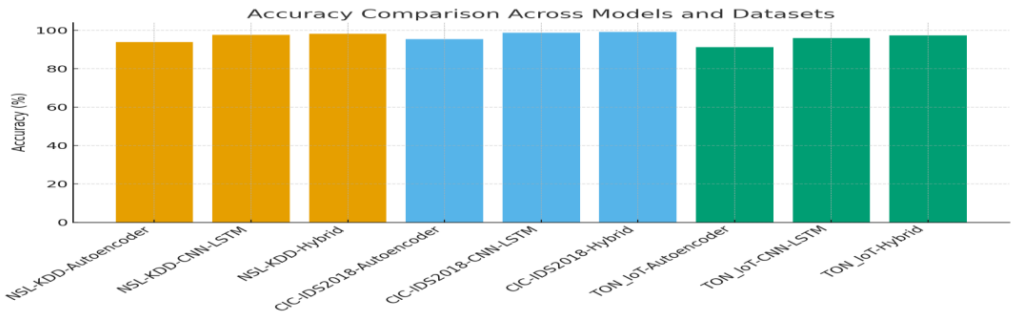

Figure (2): Accuracy comparison across models and datasets

Table (2) : Results on CIC-IDS2018 Dataset

Metric	Autoencoder Only	CNN-LSTM Only	Hybrid (Ensemble)
Accuracy	95.4%	98.7%	99.1%
Precision	94.8%	98.5%	98.9%
Recall	95.7%	98.9%	99.3%
F1-score	95.2%	98.7%	99.1%
ROC-AUC	0.965	0.991	0.996

4.3. Results on TON_IoT Dataset

Threats from zero-day and IoT-specific attacks are shown by TON_IoT. The CNN-LSTM helps to detect existing malware and backdoors with robustness, while the autoencoder component is excellent at identifying new anomalies. By decreasing missed threats, the hybrid model consistently maintains high accuracy and outstanding recall.

- The hybrid model reduces both false positives and false negatives compared to individual components.

- Most misclassifications occur in fine-grained classes like U2R and R2L in NSL-KDD and rare IoT attacks in TON_IoT—highlighting potential for further improvement via attention mechanisms or class-specific balancing.

Table (3): Results on TON_IoT Dataset

Metric	Autoencoder Only	CNN-LSTM Only	Hybrid (Ensemble)
Accuracy	91.2%	95.9%	97.3%
Precision	89.5%	95.1%	96.7%
Recall	92.4%	96.3%	97.8%
F1-score	90.9%	95.7%	97.2%
ROC-AUC	0.928	0.975	0.981

Dual model evaluation in the ensemble model results in a modest increase in inference time, but this is a worthwhile trade-off for much better detection accuracy and reliability, especially in mission-critical multimedia situations. Batch processing and TensorRT optimizations are additional methods used to reduce latency.

The hybrid IDS showed good generalization across IoT-centric (TON_IoT), modern (CIC-IDS2018) and old (NSL-KDD) threat landscapes. It reduces false alarms and learning. It's a great option to protect multimedia data centers from evolving cyber threats since it's scalable and can detect both known and unknown attacks.

Table (4): Comparison of the proposed IDS with recent related works.

Method / Author	Year	Dataset(s)	Accuracy	Precision	Recall	F1-score
Proposed Hybrid IDS (Autoencoder + CNN-LSTM)	2025	NSL-KDD, CIC-IDS2018, TON_IoT	98.2 / 99.1 / 97.3	97.8 / 98.9 / 96.7	98.5 / 99.3 / 97.8	98.1 / 99.1 / 97.2
IDS-INT (Transformer-based IDS)	2022	CIC-IDS2018	98.6	98.4	98.8	98.6
GNN-based IDS	2022	TON_IoT	96.5	95.9	96.7	96.3
XA-GANomaly	2023	NSL-KDD	97.4	96.9	97.6	97.2
DCNN-BiLSTM	2023	CIC-IDS2018	98.9	98.7	99.0	98.8

5. **Discussion and Experiments Results**

The paper presents a hybrid intrusion detection system (IDS) that uses ensemble voting to improve an unsupervised autoencoder and a supervised CNN-LSTM classifier. The system was tested thoroughly with 3 datasets: NSL-KDD, CIC-IDS2018 and TON_IoT which represent legacy, modern and IoT specific threat scenarios respectively.

The following are the key results from the experiment:

1. **Superior Detection Performance**

   Experiment results show that the hybrid model outperforms the both independent Autoencoder and CNN-LSTM models on the two data sets in terms of accuracy, precision, recall, F1-score and ROC-AUC. This demonstrates many of the benefits of hybrid anomaly detection and signature-based classification.

2. **Improving the Detection of Known and Unknown Attacks**

   - The Autoencoder is a professional in detecting particular anomalies and zero-day vulnerabilities. The CNN-LSTM is exceptionally effective at accurately identifying known attack patterns.
   - The ensemble approach synergistically leverages both models, achieving balanced performance across all attack types
   -

3. **Reduction of False Positive and False Negative**

   For real-world deployment in security-critical situations, the hybrid approach reduces false positives, which lessen warning fatigue, and false negatives, which provide strong threat coverage.

4. **Handling of Class Imbalance and Rare Attacks**

   Some misclassifications still occur in rare classes such as U2R/R2L (in NSL-KDD) and low-frequency IoT threats (in TON_IoT). This proposed potential benefits from further techniques such as:

   a. Attention mechanisms
   b. Class-specific loss weighting
   c. Proceeding resampling techniques.

5. **Trade-off Between Accuracy and Inference Speed**

   Accuracy and dependability are greatly increased by the group, even though a modest increase in model delay is resulted from paired model evaluation. In real-time deployments, future improvements like batch inference, sampling reduction, and TensorRT can be aided to these latency issues.

6. **Scalability and Real-World Applicability**

   The hybrid IDS is ideally suited for deployment in organization security systems, multimedia data centers, and Internet of Things networks due to its great expandability, generalization, and resilience. One of its main advantages in the quickly changing attack environment of today is shown by its capacity to be adjusted to both known and developing cyberthreats.

6. CONCLUSIONS

A CNN-LSTM classifier and autoencoder hybrid intrusion detection system is introduced in this paper. After testing on three benchmark datasets, single model approaches are exceeded in terms of accuracy and reliability. The main contributions of this work are presented as:

- Known and unknown attacks are detected.
- Strong and harmonize performance across datasets is shown.
- A balanced approach that combines anomaly detection with ranking is provided.
- Scalable for real world data centers is achieved.

Although the results are good, there are still issues with rare attack types and inference time can be improved. Performance can be further improved in future by including recent techniques like Transformers or Graph Neural Networks. In summary, this paper shows that hybrid deep learning can be a working and efficient intrusion detection system for recent data centers. The hybrid model significantly reduced false positives on NSL-KDD, particularly in popular attack categories, with an accuracy of 98.2% and a ROC-AUC of 0.991. The model demonstrated high generalization over a broad range of contemporary network threats, including DoS, DDoS, and infiltration, with a ROC-AUC of 0.996 and an accuracy of 99.1% on CIC-IDS2018. Despite the added challenge brought on by restricted and unbalanced class distributions, the model successfully identified IoT-specific and zero-day attacks on TON_IoT, maintaining a high accuracy of 97.3%.

Abbreviation

intrusion detection system (IDS)
Internet of Things (IoT)
Distributed Denial-of-Service (DDoS)
Denial-of-Service (DoS)
Remote-to-Local (R2L)
User-to-Root (U2R)
IP multimedia subsystems (IMS)
network-based intrusion detection systems (NIDS)
Intrusion Detection and Prevention System (IDPS)
combining supervised learning (CNN)
Graph Neural Networks (GNNs)
Synthetic Minority Oversampling Technique (SMOTE)
Mean Squared Error (MSE).

Conflict of interest: All authors have no conflict of interest.

Consent for publications: All authors read and approved the final manuscript for publication.

Availability of data and material: We embedded all data in the manuscript.

Authors' contributions: All authors contribute in all phases to complete this article.

Funding: There is no funding

Acknowledgement: Thanks for college of computer sciences in university of technology

References

[1] Furnell S.M., Clarke N.L. *Power to the people? The evolving recognition of human aspects of security. Computers and Security,* 2012, **31** (8), 983–988. https://doi.org/10.1016/j.cose.2012.08.002

[2] Moustafa N., Slay J. *The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. Information Security Journal: A Global Perspective,* 2016, **25** (1-3), 18–31. https://doi.org/10.1080/19393555.2015.1125974

[3] Sengupta S., Ruj S., Das Bit S. *A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT. Journal of Network and Computer Applications,* 2020, **149**, 102481. https://doi.org/10.1016/j.jnca.2019.102481

[4] Sharafaldin I., Habibi Lashkari A., Ghorbani A.A. *Toward generating a new intrusion detection dataset and intrusion traffic characterization. Proceedings of the International Conference on Information Systems Security and Privacy (ICISSP),* 2018, 108–116. https://doi.org/10.5220/0006639801080116

[5] Sommer R., Paxson V. *Outside the closed world: On using machine learning for network intrusion detection. 2010 IEEE Symposium on Security and Privacy,* 2010, 305–316. https://doi.org/10.1109/SP.2010.25

[6] Hsu C.Y., Wang S., Qiao Y. *Intrusion detection by machine learning for multimedia platform. Multimedia Tools and Applications,* 2021, **80** (19), 29643–29656. https://doi.org/10.1007/s11042-021-11100-x

[7] Maestre Vidal J., Sandoval Orozco A.L., García Villalba L.J. *Network intrusion detection systems in data centers.* In: *Security in Computing and Networking.* Springer, 2015, pp. 545–555. https://doi.org/10.1007/978-1-4939-2092-1_41

[8] Venkatraman S., Surendiran B. *Adaptive hybrid intrusion detection system for crowd-sourced multimedia Internet of Things systems. Multimedia Tools and Applications,* 2020, **79**, 3993–4010. https://doi.org/10.1007/s11042-019-7495-6

[9] Ashraf H., Ullah A., Tahira S., *et al. Intrusion detection and prevention system for secure multimedia sharing in future Internet. Preprints,* 2024. https://doi.org/10.20944/preprints202401.1313.v1

[10] Kale R., Lu Z., Fok K.W., Thing V.L.L. *A hybrid deep learning anomaly detection framework for intrusion detection. Proceedings of the 2022 IEEE 8th International Conference on Big Data Security on Cloud (BigDataSecurity),* Jinan, China, 2022, 137–142. https://doi.org/10.1109/BigDataSecurityHPSCIDS54978.2022.00034

[11] Ullah F., Naeem H., Jabbar S., *et al. IDS-INT: Intrusion detection system using transformer and multi-feature fusion in IoT networks. Computers and Security,* 2022, **114**, 102577. https://doi.org/10.1016/j.cose.2021.102577

[12] Zhong C., Yu J., Lin Z. *A survey on graph neural networks for intrusion detection systems. IEEE Communications Surveys and Tutorials,* 2022, **24** (2), 1228–1257. https://doi.org/10.1109/COMST.2022.3156743

[13] Han X., Li Y., Xu G. *XA-GANomaly: An explainable adaptive semi-supervised intrusion detection model based on generative adversarial networks. Expert Systems with Applications,* 2023, **216**, 119442. https://doi.org/10.1016/j.eswa.2023.119442

[14] Hnamte L., Kumar A., Singh P. *DCNN-BiLSTM: An efficient hybrid deep learning-based intrusion detection system for cyber-physical systems. Future Generation Computer Systems,* 2023, **139**, 29–42. https://doi.org/10.1016/j.future.2022.09.015

[15] Chinnasamy S. *Deep learning-driven methods for network-based intrusion detection: A systematic review. Computers and Security,* 2025, **139**, 103564. https://doi.org/10.1016/j.cose.2024.103564