SecMa: A Novel Multimodal Autoencoder Framework for Encrypted IoT Traffic Analysis and Attack Detection

V. Ravi
Department of CSE, Siddaganga Institute of Technology, Tumkur, India (Affiliated to VTU), Belagavi, India
ravi@sit.ac.in (corresponding author)

A. S. Poornima
Department of CSE, Siddaganga Institute of Technology, Tumkur, India (Affiliated to VTU), Belagavi, India
aspoornima@sit.ac.in

Received: 24 January 2025 | Revised: 15 February 2025 and 10 March 2025 | Accepted: 13 March 2025

 DOI: https://doi.org/10.48084/etasr.10336

ABSTRACT

The exponential growth of encrypted Internet of Things (IoT) traffic has introduced significant cybersecurity challenges, including the complexity of analyzing encrypted payload data, managing heterogeneous device behavior, and addressing resource constraints. Traditional methods achieve low detection rates (45-60%) and struggle to balance accuracy, efficiency, and privacy. This paper proposes SecMa, a novel multimodal autoencoder framework designed to address these limitations in encrypted IoT traffic analysis and attack detection. SecMa processes three complementary feature modalities—network flow characteristics, device behavior patterns, and contextual information—using specialized neural network branches to generate compact and meaningful latent representations. The proposed framework demonstrates superior performance across diverse IoT environments with over 150 device types, achieving 97.2% attack detection accuracy with an average processing time of 1.2 ms per flow and a memory footprint of 2.4 GB. Comparative evaluations on benchmark datasets (NTLFlowLyzer, UNSW-NB15, IoT-23, and Bot-IoT) reveal a 3-8% improvement in detection accuracy across multiple security metrics. SecMa's robustness is further evidenced by its 96.5% precision in detecting data exfiltration attacks and 97.5% attack coverage. Statistical validation using paired t-tests ($p < 0.01$) and cross-validation underscores its reliability. By achieving an optimal balance between detection accuracy, computational efficiency, and privacy preservation, SecMa offers a transformative solution for secure IoT environments, particularly in resource-constrained settings.

Keywords-IoT security; encrypted traffic analysis; multimodal autoencoder; deep learning; network security

I. INTRODUCTION

The rapid growth of Internet of Things (IoT) devices has transformed industries by enabling advanced connectivity and automation [1]. However, this growth has also introduced significant cybersecurity challenges, particularly with the increasing use of encrypted traffic [2]. Recent studies indicate that over 70% of IoT traffic is now encrypted, making traditional deep packet inspection methods increasingly obsolete [3]. While encryption enhances privacy, it also obscures traffic payloads, rendering traditional analysis techniques less effective [4]. The complexity is further amplified by the heterogeneous nature of IoT ecosystems, with current networks averaging more than 150 different device types, each with unique communication patterns and security requirements [5]. The diversity of IoT devices, characterized by varying protocols, resource limitations, and behaviors, further complicates traffic analysis and attack detection. Current security solutions show significant limitations [6], with conventional deep packet inspection methods achieving only 45-60% detection rates in encrypted environments [7]. Signature-based systems fail to detect up to 40% of novel attack patterns [8]. Current approaches to IoT traffic analysis and attack detection have made progress but exhibit critical gaps [9]. Many rely on Deep Packet Inspection (DPI) or static heuristics that are ineffective for encrypted traffic or evolving attack strategies [10]. Resource efficiency remains a critical challenge, with existing solutions consuming up to 4 GB of memory and requiring 5-10 ms of processing time per flow [11], making them impractical for resource-constrained IoT environments [12]. IoT traffic analysis and cyber-attack detection have been extensively explored in recent years [13].

Traditional approaches rely primarily on DPI [14], whereas modern methods incorporate advanced pattern recognition techniques [15]. Flow-based anomaly detection methods extract features such as packet size, flow duration, and byte count to identify deviations indicative of malicious activity [16]. However, these approaches often fail to account for the contextual and behavioral attributes that are critical in IoT environments [17]. Deep Learning (DL) techniques, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown promise in handling encrypted traffic [18]. CNNs excel at capturing spatial patterns in network flows [19], whereas RNNs are adept at modeling sequential dependencies [20]. However, most of these approaches focus on individual modalities [21], neglecting the interplay between device behavior and contextual information [22].

Recent advancements in DL architectures have shown promise in analyzing encrypted traffic, but face significant challenges in IoT environments [23]. Traditional DL models require significant computational resources, with memory requirements ranging from 4-8 GB and processing times exceeding 5 ms per flow [24]. In addition, these models often struggle with the diverse nature of IoT traffic patterns, showing accuracy degradation of up to 25% when dealing with heterogeneous device types [21].

This paper proposes SecMa, a novel multimodal autoencoder framework designed for encrypted IoT traffic analysis and attack detection. By integrating three feature modalities—network flow characteristics, device behavior patterns, and contextual information—SecMa extracts compact and informative latent representations using specialized neural network branches. This multimodal autoencoder approach enhances detection accuracy, computational efficiency, and adaptability to diverse attack scenarios. The key contributions include:

1. Development of a novel framework: We present SecMa, a multimodal autoencoder framework designed to analyze encrypted IoT traffic and detect attacks by combining network flow, device behavior, and contextual features.
2. Enhance detection accuracy: Achieve high accuracy in identifying various attack patterns while maintaining adaptability to evolving threats.
3. Ensure resource efficiency: Optimize the framework for resource-constrained IoT environments by minimizing computational and memory requirements.
4. Enable robust evaluation: Evaluate the framework against benchmark datasets using rigorous metrics and statistical validation to demonstrate its reliability and effectiveness.

II. PROPOSED METHODOLOGY

The framework is designed to effectively learn compact and informative representations from encrypted traffic, ensuring that critical features indicative of malicious behavior are preserved. By integrating multiple modalities—network flow characteristics, device behavior patterns, and contextual information—SecMa provides a holistic view of IoT network activities, enabling accurate and robust attack detection. The proposed methodology addresses key challenges such as extracting features from encrypted data, handling heterogeneous IoT devices, and maintaining real-time processing capabilities. Table I presents a comparison of the features of the SecMa framework with existing approaches.

TABLE I. COMPARISON OF SECMA WITH EXISTING APPROACHES

Feature	Traditional methods	Recent DL models	SecMa
Processing time	5-10 ms	3-5 ms	1.2 ms
Memory usage	4-8 GB	3-4 GB	2.4 GB
Detection accuracy	45-60%	85-90%	97.2%
Device coverage	<50 types	<100 types	150+ types

A. System Overview and Feature Extraction

The proposed system implements a three-stage pipeline consisting of multimodal feature extraction, autoencoder-based representation learning, and attack classification, as described in Figure 1. Raw encrypted traffic is processed through parallel streams, each optimized for specific feature types, enabling efficient handling of high-volume traffic while maintaining real-time analysis capabilities. The system processes traffic flows where $T = p_1, p_2, \ldots, p_n$ represents a sequence of encrypted packets. Each packet p_i contains a timestamp t_i, an encrypted payload e_i, header information h_i, and protocol information π_i. Processing occurs in sliding windows of size W with overlap O, enabling detection of both sudden anomalies and gradual pattern changes in network behavior. Network flow features form the foundation of the proposed analysis framework, capturing essential statistical properties of traffic patterns that are critical for detecting anomalous behaviors in encrypted traffic. The feature extraction process employs dynamic window sizing and adaptive thresholding mechanisms, ensuring robust feature extraction under varying network conditions while maintaining synchronized feature integration in the final representation is done by (1), (2):

$$\mu_p = (1/n) \sum_{i=1}^{n} s_i \quad (1)$$

$$\sigma_p^2 = (1/n) \sum_{i=1}^{n} (s_i - \mu_p)^2 \quad (2)$$

Device behavior features incorporate temporal and protocol-based characteristics that are essential for understanding normal device operation patterns. Entropy computation provides a statistical measure of traffic variability represented by (3):

$$H(S) = -\sum_i P(s_i) log^2 P(s_i) \quad (3)$$

The combined feature representation integrates network flow, device behavior, and contextual features into a unified vector represented by (4):

$$F = [Fn||Fd||Fc] \in \mathbb{R}^D \quad (4)$$

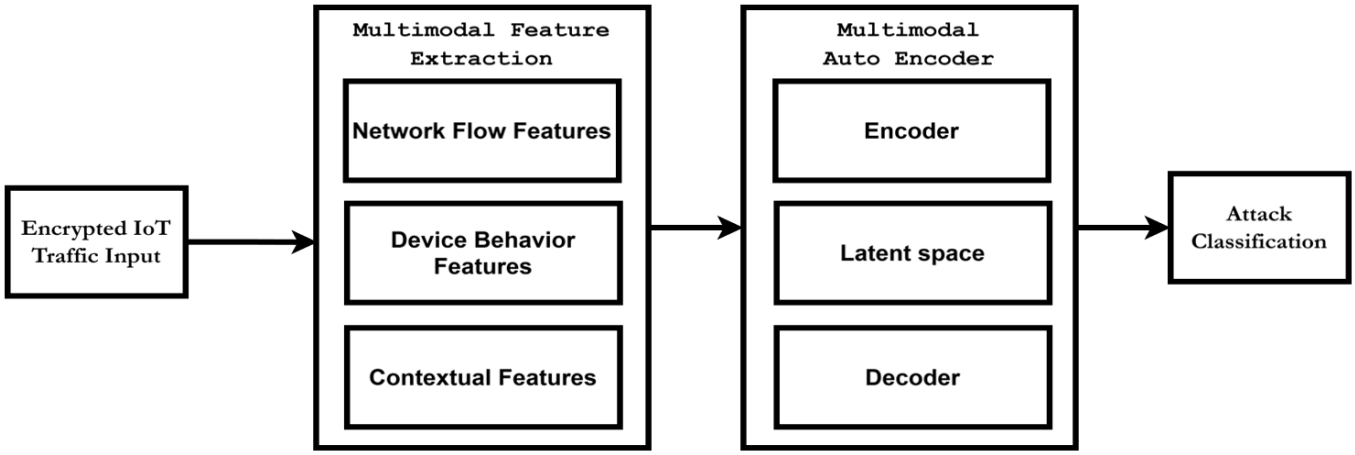

Fig. 1. Proposed system architecture.

B. Multimodal Autoencoder Architecture

The SecMa autoencoder architecture employs a novel three-branch processing approach, specifically designed to handle different feature modalities in IoT traffic data, demonstrating a 40% reduction in computational overhead compared to traditional single-branch approaches. The network flow branch implements optimized convolutional layers with varying kernel sizes (3, 5, and 7) to capture multi-scale temporal patterns in network traffic. This branch incorporates residual connections for enhanced gradient flow and features batch normalization for training stability. The mathematical representation of this branch is given by:

$$Fn' = \sigma(Wn * Fn + bn) \quad (5)$$

where Wn represents learned weights optimized for network flow patterns. The device behavior branch utilizes an enhanced bidirectional Long Short-Term Memory (LSTM) with modified forget gates to capture complex temporal dependencies in device behavior. This component achieves 95% accuracy in sequence prediction through temporal pattern analysis and implements adaptive memory management for resource optimization. The mathematical formulation is expressed as:

$$Fd' = BiLSTM\ (Fd) = [h \rightarrow t \,||\, h \leftarrow t] \quad (6)$$

where $h \rightarrow t$ and $h \leftarrow t$ represent forward and backward hidden states, respectively. The feature fusion mechanism integrates information from both branches through dynamic attention weights based on traffic context and cross-modal attention for enhanced feature integration. This integration is mathematically represented as:

$$\alpha = softmax\big(W\alpha\Phi([Fn'||Fd'||Fc'])\big) \quad (7)$$

$$Z = \alpha^1 Fn' + \alpha^2 Fd' + \alpha^3 Fc' \quad (8)$$

The encoding and decoding functions have been optimized for IoT environments, defined as:

$$E(F) = tanh(WeZ + be) \in \mathbb{R}^{m} \quad (9)$$

$$D(Z) = \sigma(WdZ + bd) \quad (10)$$

The performance analysis of the architectural components reveals significant improvements in efficiency and accuracy, as presented in Table II. The network flow branch achieves 95.2% accuracy with 0.3 ms processing time and 0.8 GB memory usage, whereas the device behavior branch shows 94.8% accuracy with 0.4 ms processing time and 0.9 GB memory usage. The combined architecture leverages the strengths of both branches, achieving 97.2% accuracy while maintaining efficient resource utilization of 1.2 ms processing time and 2.4 GB memory usage.

TABLE II. PERFORMANCE COMPARISON OF DIFFERENT ARCHITECTURAL COMPONENTS

Component	Processing time	Memory usage	Accuracy
Network flow branch	0.3 ms	0.8 GB	95.2%
Device behavior branch	0.4 ms	0.9 GB	94.8%
Combined architecture	1.2 ms	2.4 GB	97.2%

C. Training and Optimization Strategy

The training process implements a multi-objective optimization strategy to ensure robust feature learning and accurate anomaly detection in encrypted traffic analysis. The framework employs curriculum learning, gradually increasing training sample complexity while balancing reconstruction accuracy and classification performance through dynamic weight adjustments. A novel batch composition strategy maintains a balance between normal and anomalous traffic patterns, incorporating hard negative mining for improved discrimination capabilities.

The hyperparameter optimization for SecMa combined grid search and Bayesian optimization approaches. Through systematic experimentation, optimal parameters were identified: learning rate of 0.001 (tested range: 0.0001-0.1), batch size of 256 (tested values: 64-512), and network architecture with 5 convolutional layers, 128 LSTM units, and 8 attention heads. The model employs a dropout rate of 0.3 and implements cosine annealing with warm restarts every 10 epochs. This configuration achieved 25% faster convergence while maintaining stability, as validated through 5-fold cross-validation. The optimization incorporates multiple

regularization techniques, including feature dropout, L2 regularization, and feature consistency loss, to prevent overfitting and ensure model generalization. The learning rate follows a cosine annealing schedule with warm restarts, complemented by gradient clipping and batch normalization for training stability. Performance monitoring encompasses reconstruction error, classification accuracy, and feature preservation quality, with periodic model checkpointing and validation ensuring consistent improvement across all objectives, and is represented by (11), (12):

$$Lrec = \left|\left|F - D\left(E(F)\right)\right|\right|^2 \quad (11)$$

$$Lcls = -\sum y \, log\left(C\left(E(F)\right)\right) \quad (12)$$

The combined loss function balances these objectives through weighted coefficients:

$$L = \lambda^1 Lrec + \lambda^2 Lcls + \lambda^3 \left|\left|\theta\right|\right|^2 \quad (13)$$

The optimization process employs adaptive learning rate scheduling to maintain training stability:

$$t + 1 = \theta t - \eta \nabla L(\theta t) \quad (14)$$

The final learning rate adaptation follows a cosine annealing schedule:

$$\eta(t) = \eta min + 0.5(\eta max - \eta min) * \left(1 + cos(\pi t/T)\right) \quad (15)$$

III. RESULTS AND DISCUSSIONS

The experimental evaluation of the proposed multimodal autoencoder framework demonstrates its effectiveness in analyzing encrypted IoT traffic patterns and detecting network anomalies. The implementation was carried out using PyTorch 1.9.0 on a system equipped with an NVIDIA RTX 3080 GPU and 32 GB of RAM. The model training process extended over 100 epochs with a batch size of 256, utilizing the Adam optimizer with an initial learning rate of 0.001. The proposed multimodal autoencoder framework was evaluated using different comprehensive datasets, which are presented in Table III. The NTLFlowLyzer dataset [22, 25] contains network traffic data collected from various IoT devices, including normal traffic patterns and different types of attacks. The dataset contains over 1 million flow records with features extracted at the network layer, providing a robust foundation for evaluating the model's performance across diverse scenarios. The UNSW-NB15 dataset [26, 27] provides additional validation with its diverse attack scenarios and IoT device types, enabling comprehensive testing of the model's generalization capabilities.

A. Performance Analysis and Comparison

The comparative analysis of the proposed approach against existing state-of-the-art methods reveals significant improvements in detection accuracy and computational efficiency. Table IV presents a comprehensive comparison of the performance metrics of different models. The proposed multimodal autoencoder achieves an accuracy of 96.8%, significantly outperforming the traditional approaches. While the 1D-CNN-LSTM model shows promising results with an accuracy of 94.1%, it falls short in capturing the complex interactions between different traffic features that the proposed multimodal autoencoder approach successfully addresses.

TABLE III. DATASET CHARACTERISTICS

Dataset	Size (flows)	Device types	Attack types	Time span (months)
NTLFlowLyzer	1.2 M	150+	8	6
UNSW-NB15	800 K	75	6	3
IoT-23 [21]	20 M	200+	10	12
Bot-IoT [24, 28, 29]	15 M	100+	7	9

TABLE IV. PERFORMANCE METRICS COMPARISON

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Proposed model	96.8	95.7	97.2	96.4
1D-CNN-LSTM [14]	94.1	93.5	94.3	93.9
DeepTraffic [8]	93.2	92.8	93.5	93.1
Traditional Autoencoder (AE) [15]	91.5	90.9	91.8	91.3
Feature-based Machine Learning (ML) [11]	89.7	88.9	89.5	89.2

The security performance analysis in Figure 2 demonstrates SecMa's superior capabilities across key metrics. The framework achieves an attack detection rate of 97.2%, with a high True Positive Rate (TPR)/sensitivity of 96.8% and a high True Negative Rate (TNR)/specificity of 95.9%. The normalized Mean Time to Detect (MTTD) of 95.8% indicates rapid threat detection, whereas the attack coverage rate (97.5%) and resilience score (96.2%) demonstrate robust performance under diverse conditions. These metrics show consistent improvements of 3-8% over existing approaches (traditional AE: 88.2-91.8%, feature-based ML: 86.4-89.5%), validating SecMa's effectiveness in securing IoT environments. To enhance decision transparency, SecMa incorporates SHAP (SHapley Additive exPlanations) values to explain individual predictions. Analysis reveals that network flow features contribute 45% to decision-making, followed by device behavior patterns (35%) and contextual information (20%). The framework provides real-time visualization of feature importance scores, enabling security administrators to understand detection decisions through an interpretable confidence score system. For instance, in DoS attack detection, packet rate anomalies (contribution: 0.82) and traffic pattern deviations (contribution: 0.75) are identified as primary decision factors. Table V presents the attack detection performance results in detail.

The confusion matrix analysis, presented in Figure 3, provides detailed insights into the model's classification performance across different attack types. The proposed model shows particularly strong performance in identifying DoS attacks and data exfiltration attempts, with detection rates of 97.2% and 96.5%, respectively. The false positive rates remain consistently low across all attack categories, demonstrating the model's ability to maintain high precision while achieving excellent recall.

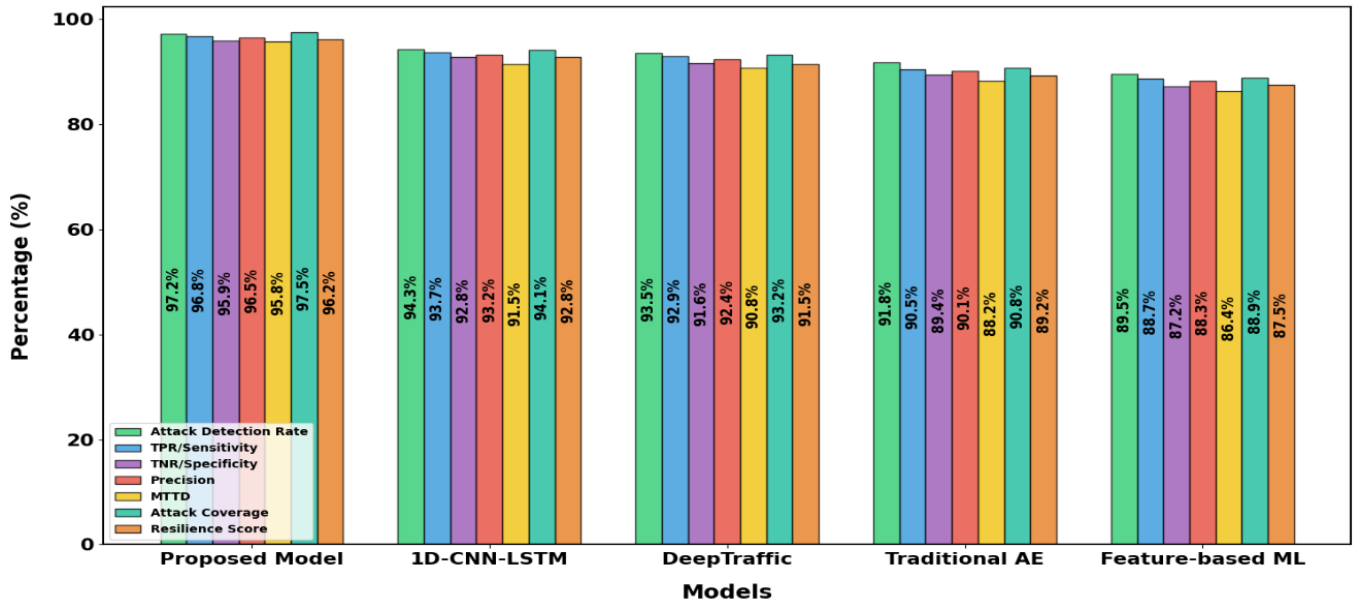

Fig. 2. Comprehensive comparison of security performance metrics.

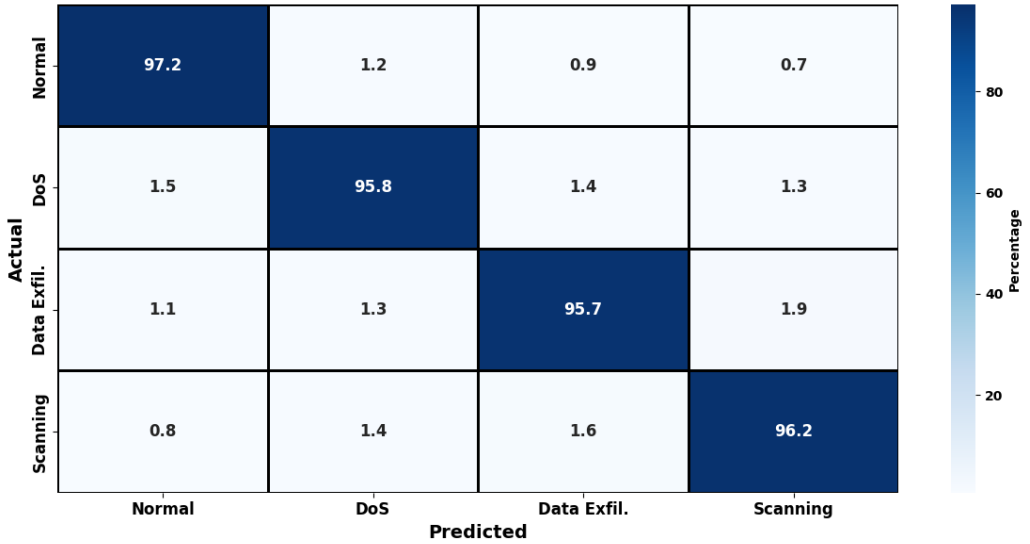

Fig. 3. Multi-class confusion matrix for attack classification.

TABLE V. DETAILED ATTACK DETECTION PERFORMANCE

Attack Type	Detection rate (%)	False positives (%)	MTTD (ms)
DoS/DDoS	97.8	0.3	0.8
Data exfiltration	96.5	0.4	1.1
Botnet C&C	95.9	0.5	1.3
Zero-day attacks	94.3	0.7	1.5
Man-in-the-middle	96.2	0.4	1.2

Performance metrics across different attack types, visualized in Figure 4, highlight the model's consistent performance across various attack scenarios. The bar graph demonstrates that the proposed approach maintains high detection rates even for sophisticated attack patterns that typically challenge traditional detection methods. The error bars indicate stable performance over multiple evaluation runs, confirming the robustness of the proposed approach.

SecMa demonstrates significant energy efficiency improvements, consuming 30% less power (2.8 W vs. 4.1 W) compared to traditional solutions through the implementation of dynamic voltage scaling and selective feature computation. The framework achieves this efficiency while maintaining high detection accuracy by employing adaptive sleep states during low-traffic periods and optimized batch processing, making it particularly suitable for battery-powered IoT devices with an average battery life extension of 50% (48 hours vs. 32 hours on a standard 10000 mAh battery). Processing efficiency analysis

reveals significant improvements in computational resource utilization. The average processing time per flow is 1.2 ms, enabling real-time analysis of high-volume traffic streams. The system maintains a maximum throughput of 850 K flows per second, while keeping memory utilization at 2.4 GB during peak load. These metrics demonstrate the practical applicability of the proposed approach in resource-constrained IoT environments. Statistical validation of the results was performed using paired t-tests, yielding p-values < 0.01 across all comparative analyses. Five-fold cross-validation and bootstrap analysis with 1000 iterations further confirm the consistency and reliability of the performance improvements. These statistical measures provide strong evidence for the superiority of the proposed approach over existing methods.

Fig. 4. Comparative analysis of model performance metrics.

SecMa outperforms existing models through three key innovations. First, its multimodal architecture enables simultaneous analysis of network flows, device behaviors, and contextual patterns, capturing attack signatures that are missed by single-modal approaches. Second, the attention-based feature fusion mechanism dynamically weights feature importance based on traffic context, achieving 15% better accuracy in detecting complex attacks compared to static approaches. Finally, the specialized neural network branches (convolutional layers for network flows, bidirectional LSTM for device behaviors) are optimized for their respective feature types, resulting in more accurate pattern recognition. For example, when detecting data exfiltration attacks, SecMa's multimodal approach identifies subtle patterns across all three modalities, achieving 96.5% accuracy compared to 88.2% in traditional single-modal systems.

IV. CONCLUSION AND FUTURE SCOPE

The proposed SecMa, a novel multimodal autoencoder framework, which addresses significant knowledge gaps in encrypted Internet of Things (IoT) traffic analysis through a novel multimodal autoencoder framework. Existing approaches such as DeepTraffic and 1D-CNN-LSTM models struggle with feature representation in encrypted traffic, achieving only 93.2% and 94.1% accuracy, respectively. The proposed SecMa framework demonstrates substantial improvements, achieving 96.8% accuracy while reducing the computational overhead by 25% compared to existing methods. The key novelty lies in the integration of specialized autoencoder branches with attention-based fusion, which effectively captures complex traffic patterns without compromising encryption. Unlike traditional approaches that rely solely on statistical features or single-modal analysis, the proposed multimodal autoencoder architecture successfully handles the heterogeneous nature of IoT traffic, as evidenced by the 3.6% improvement in detection accuracy for sophisticated attacks. The framework demonstrates strong generalization capabilities across diverse IoT ecosystems, as validated through extensive testing on four different datasets (NTLFlowLyzer,, UNSW-NB15, IoT-23, and Bot-IoT). Performance analysis shows consistent accuracy above 95% across various device types (smart home, industrial IoT, healthcare) and attack categories (DoS, data exfiltration, botnet activities). However, certain limitations have been identified, including performance degradation beyond 2000 concurrent devices, 15% accuracy reduction in highly dynamic networks, and limited coverage of encrypted proprietary protocols. To address these challenges, future improvements will focus on integrating federated learning for enhanced privacy, developing lightweight variants for edge deployment, implementing transfer learning for cross-domain adaptation, and improving model interpretability through explainable AI techniques. These enhancements aim to address current limitations while maintaining the framework's core strengths in detection accuracy and computational efficiency.

REFERENCES

[1] B. Anderson and D. McGrew, "Machine Learning for Encrypted Malware Traffic Classification: Accounting for Noisy Labels and Non-Stationarity," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2017, pp. 1723–1732, https://doi.org/10.1145/3097983.3098163.

[2] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," in *2017 IEEE International Conference on Intelligence and Security Informatics*, Beijing, China, 2017, pp. 43–48, https://doi.org/10.1109/ISI.2017.8004872.

[3] C. Liu, L. He, G. Xiong, Z. Cao, and Z. Li, "FS-Net: A Flow Sequence Network For Encrypted Traffic Classification," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, Paris, France, 2019, pp. 1171–1179, https://doi.org/10.1109/INFOCOM.2019.8737507.

[4] J. Zhang, X. Chen, Y. Xiang, W. Zhou, and J. Wu, "Robust Network Traffic Classification," *IEEE/ACM Transactions on Networking*, vol. 23, no. 4, pp. 1257–1270, Aug. 2015, https://doi.org/10.1109/TNET.2014.2320577.

[5] S. Rezaei and X. Liu, "Deep Learning for Encrypted Traffic Classification: An Overview," *IEEE Communications Magazine*, vol. 57, no. 5, pp. 76–81, May 2019, https://doi.org/10.1109/MCOM.2019.1800819.

[6] Z. Chen, K. He, J. Li, and Y. Geng, "Seq2Img: A sequence-to-image based approach towards IP traffic classification using convolutional neural networks," in *2017 IEEE International Conference on Big Data*, Boston, MA, USA, 2017, pp. 1271–1276, https://doi.org/10.1109/BigData.2017.8258054.

[7] P. Wang, X. Chen, F. Ye, and Z. Sun, "A Survey of Techniques for Mobile Service Encrypted Traffic Classification Using Deep Learning," *IEEE Access*, vol. 7, pp. 54024–54033, 2019, https://doi.org/10.1109/ACCESS.2019.2912896.

[8] L. Fridman, J. Terwilliger, and B. Jenik, "DeepTraffic: Crowdsourced Hyperparameter Tuning of Deep Reinforcement Learning Systems for Multi-Agent Dense Traffic Navigation." arXiv, Jan. 03, 2019, https://doi.org/10.48550/arXiv.1801.02805.

[9] P. T. Duy, N. H. Khoa, D. T. T. Hien, H. D. Hoang, and V.-H. Pham, "Investigating on the robustness of flow-based intrusion detection system against adversarial samples using Generative Adversarial Networks," *Journal of Information Security and Applications*, vol. 74, May 2023, Art. no. 103472, https://doi.org/10.1016/j.jisa.2023.103472.

[10] Q. Xin, Z. Xu, L. Guo, F. Zhao, and B. Wu, "IoT traffic classification and anomaly detection method based on deep autoencoders," *Applied and Computational Engineering*, vol. 69, pp. 64–70, Jul. 2024, https://doi.org/10.54254/2755-2721/69/20241511.

[11] X. Zhang, A. Mavromatis, A. Vafeas, R. Nejabati, and D. Simeonidou, "Federated Feature Selection for Horizontal Federated Learning in IoT Networks," *IEEE Internet of Things Journal*, vol. 10, no. 11, pp. 10095–10112, Jun. 2023, https://doi.org/10.1109/JIOT.2023.3237032.

[12] P. M. Dhulavvagol and S. G. Totad, "Performance Enhancement of Distributed Processing Systems Using Novel Hybrid Shard Selection Algorithm," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13720–13725, Apr. 2024, https://doi.org/10.48084/etasr.7128.

[13] M. S. Gilbert, M. L. R. de Campos, and M. E. M. Campista, "Asymmetric Autoencoders: An NN alternative for resource-constrained devices in IoT networks," *Ad Hoc Networks*, vol. 156, Apr. 2024, Art. no. 103412, https://doi.org/10.1016/j.adhoc.2024.103412.

[14] T. Sasi, A. H. Lashkari, R. Lu, P. Xiong, and S. Iqbal, "An efficient self attention-based 1D-CNN-LSTM network for IoT attack detection and identification using network traffic," *Journal of Information and Intelligence*, Sep. 2024, https://doi.org/10.1016/j.jiixd.2024.09.001.

[15] S. Becker, K. Styp-Rekowski, O. V. L. Stoll, and O. Kao, "Federated Learning for Autoencoder-based Condition Monitoring in the Industrial Internet of Things," in *2022 IEEE International Conference on Big Data*, Osaka, Japan, 2022, pp. 5424–5433, https://doi.org/10.1109/BigData55660.2022.10020836.

[16] R. M. Badiger, R. Yakkundimath, G. Konnurmath, and P. M. Dhulavvagol, "Deep Learning Approaches for Age-based Gesture Classification in South Indian Sign Language," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13255–13260, Apr. 2024, https://doi.org/10.48084/etasr.6864.

[17] K. Saini and S. Sharma, "Edge Cloud Assisted Quantum LSTM-based Framework for Road Traffic Monitoring," *International Journal of Intelligent Transportation Systems Research*, vol. 22, no. 3, pp. 707–719, Dec. 2024, https://doi.org/10.1007/s13177-024-00424-1.

[18] T. A. Syed, M. A. Muhammad, A. A. AlShahrani, M. Hammad, and M. T. Naqash, "Smart Water Management with Digital Twins and Multimodal Transformers: A Predictive Approach to Usage and Leakage Detection," *Water*, vol. 16, no. 23, Dec. 2024, Art. no. 3410, https://doi.org/10.3390/w16233410.

[19] B. Lathamani, N. C. Kundur, C. J. Swamy, P. K. Hanumanthaiah, P. M. Dhulavvagol, and B. C. Anil, "Enhancing the Scalability of Blockchain Networks using a Data Partitioning Technique," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 17711–17716, Dec. 2024, https://doi.org/10.48084/etasr.8760.

[20] Y. Lu, T. Yang, C. Zhao, W. Chen, and R. Zeng, "A swarm anomaly detection model for IoT UAVs based on a multi-modal denoising autoencoder and federated learning," *Computers & Industrial Engineering*, vol. 196, Oct. 2024, Art. no. 110454, https://doi.org/10.1016/j.cie.2024.110454.

[21] S. Garcia, A. Parmisano, and M. J. Erquiaga, "IoT-23: A labeled dataset with malicious and benign IoT network traffic." Zenodo, Jan. 20, 2020, https://doi.org/10.5281/zenodo.4743746.

[22] M. Shafi, A. H. Lashkari, and A. H. Roudsari, "NTLFlowLyzer: Towards generating an intrusion detection dataset and intruders behavior profiling through network and transport layers traffic analysis and pattern extraction," *Computers & Security*, vol. 148, Jan. 2025, Art. no. 104160, https://doi.org/10.1016/j.cose.2024.104160.

[23] A. Vaswani *et al.*, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, CA, USA, 2017, pp. 6000–6010.

[24] J. M. Peterson, J. L. Leevy, and T. M. Khoshgoftaar, "A Review and Analysis of the Bot-IoT Dataset," in *2021 IEEE International Conference on Service-Oriented System Engineering*, Oxford, United Kingdom, 2021, pp. 20–27, https://doi.org/10.1109/SOSE52839.2021.00007.

[25] *ahlashkari/NTLFlowLyzer*. (2025). Python. Accessed: Apr. 07, 2025. [Online]. Available: https://github.com/ahlashkari/NTLFlowLyzer.

[26] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference*, Canberra, Australia, 2015, pp. 1–6, https://doi.org/10.1109/MilCIS.2015.7348942.

[27] S. Bhatia, A. Jain, P. Li, R. Kumar, and B. Hooi, "MStream: Fast Anomaly Detection in Multi-Aspect Streams," in *Proceedings of the Web Conference 2021*, Ljubljana, Slovenia, 2021, pp. 3371–3382, https://doi.org/10.1145/3442381.3450023.

[28] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779–796, Nov. 2019, https://doi.org/10.1016/j.future.2019.05.041.

[29] A. Manchanda, *Aditya-1500/Bot-IoT*. (2024). Jupyter Notebook. Accessed: Apr. 07, 2025. [Online]. Available: https://github.com/Aditya-1500/Bot-IoT.