Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(REVIEW ARTICLE)

Check for updates

Efficient malware detection in software systems using handcrafted features, Bi-GRUs, and VAEs

Aravindhan Kurunthachalam *

School of Computing and Information Technology REVA University, Bangalore.

Global Journal of Engineering and Technology Advances, 2025, 22(03), 165-174

Publication history: Received on 07 February 2025; revised on 18 March 2025; accepted on 21 March 2025

Article DOI: https://doi.org/10.30574/gjeta.2025.22.3.0060

Abstract

With the increasingly complex cyber-attacks, malware detection is now crucially required, which poses difficult challenges to traditional security systems regarding software testing. Existing techniques such as SVR, LSTM, and HMM are not capable of malware detection, especially for multiple unknown threats and large datasets. These often lead to high false positive rates, slow detection times, and limited capabilities for anomalous detection. The aim to bridge the said gaps is what this paper proposes a hybrid malware detection framework based on handcrafted feature extraction using LightGBM, temporal analysis with Bidirectional Gated Recurrent Units (Bi-GRU), and anomaly detection with Variational Autoencoders (VAE), fused with attention mechanism for improved performance. The novelty of this approach lies in its ability to bring on board an innovative combination of techniques that capture a wide variety of attack behaviors, so the system provides robust detection significantly improved in terms of detection accuracy as well as efficiency. The proposed framework attains 99.8% accuracy, 99.2% precision, 98.5% recall, and a considerable decrease in detection time down to 0.08 seconds, along with low false positive rates of 0.8% and false negative rates of 0.3%. Superior performance is demonstrated in comparison with the existing methods, offering faster and more precise malware detection. This is one great way to ascertain the breakthrough in malware detection, boosting system security and performance for software testing, and has the innate capacity to be scalable and adaptable to more complex and evolving threats in whatever network environments to ensure even faster, accurate detection and mitigation.

Keywords: Malware detection; Software testing; Light Gradient Boosting Machine; Bidirectional Gated Recurrent Unit; Variational Autoencoders

1. Introduction

Artificial Intelligence admittedly indicates a drastic change in industries such as healthcare, security, automation, and so forth. It does improve efficiency and efficacy in operations. However, besides having so much promise in the world, AI technology also faces challenges in data privacy, scale, and computation. For example, mobile healthcare, IoT technology, and artificial intelligence have been changed tremendously, allowing patients to be monitored remotely and improving accessibility to their medical records, ultimately enhancing patient care and timely interventions [1]. AI models for early tumor detection and medical image analysis are evolving, bringing forth more reliable diagnoses and better medical outcomes [2].

Models in deep learning for lung cancer detection have used medical imaging data combined with genetic data and have been sufficient enough for early diagnosis and planning for more effective treatment [3]. Management of chronic diseases such as chronic kidney disease is improved through an integrated system using probabilistic neuro-fuzzy methods for better patient monitoring and diagnosis in AI [4]. The use of cloud computing has greatly improved data

* Corresponding author: Aravindhan Kurunthachalam

security in the area of healthcare, while AI and blockchain combined will even take it a step further in privacy protection by secure transmission methods and biometric authentication [5].

Beyond healthcare, AI optimizes electric vehicle performance where improved energy efficiency comes from novel advanced models such as artificial neural networks and electrothermal inverter designs [6]. These systems, however, lead to new security challenges in distributed computing environments that require stronger authentication mechanisms [7]. The curves from training require more sophisticated and better models in terms of performance. AI-enabled advancements in software development include the integration of pre-trained language models with evolutionary algorithms to generate test cases and overall test coverage [8].

AI-powered SaMD is rigorously being scrutinized for post-market surveillance to make sure that patients are safe, regulatory compliance is upheld, and risks are effectively managed [9]. In the fifth generation of communication systems, AI improves channel state information (CSI) using techniques such as backpropagation neural networks (BPNN) and generative adversarial networks to make the use of the signal efficient and the communication reliable [10]. The intersection of AI and the digital economy, as evidenced by emerging developments, leads to further boosts in economic growth via supporting sustainable entrepreneurship while improving business practices [11].

These systems help response personnel respond quickly to disasters by improving their usefulness in terms of operation: AI-based Cloud-GIS is used for crisis management that speeds restoration after earthquakes through data processing and predictive analytics [12]. In education, AI and data analytics are used to enhance learning outcomes in the cloud through e-learning platforms to ensure that data remains secure and integrated because nobody needs much extra source of anguish: learning [13]. Knowledge management techniques backed by adaptive modeling have been improving avenues for business planning and decision-making such that better-qualified strategic choices can be made by organizations [14]. Big data analytics make possible competitive advantages for small to medium enterprises (SMEs) in e-commerce by giving increasingly better insights into market trends and consumer behavior [15]. This growth in the complexity of IoT systems has made important such new approaches and techniques as combining Device Management Platforms (DMPs) and Self-Organizing Maps (SOMs) for improved decision-making [16].

Combining RPMA, BLE, and LTE-M technologies with Gaussian Mixture Models (GMM) solves the problem of managing devices in IoT networks in which communicating and data management technologies improve power consumption, improved data throughput, and enhanced anomaly detectability for optimizing IoT networks for application such as smart cities and agriculture [17].

Through integrating AI-powered CAPTCHA, DROP methodology graphical passwords, AES encryption, and neural network-based authentication in a multilayer authentication model, a strong solution can be derived to improve the security and usability of such a system. Its defense will therefore be stronger against automated and brute-force attacks while still being seamless at the user end [18]. BIRCH clustering, integrated with LPWAN technology, NCA, and MDS, offers great potential in increasing communication efficiency besides data clustering and dimensionality reduction for applications on blockchain implementation [19].

AI-managed systems that combine Asynchronous Advantage Actor-Critic (A3C), Trust-Region Policy Optimization (TRPO), and Partially Observable Markov Decision Processes (POMDPs) can improve systems' decision-making off-the-cuff without having highly accurate data under very unknown environments [20]. The authentication of users and data sharing in a cloud environment is complemented with security that is provided by integrating SHA-256 and RSA [21]. The synthesis and combination of big data analytics, ethnographic insights from qualitative domain-specific studies, and network analysis provide invaluable resources that enable healthcare systems to deliver better personalized care for improved clinical outcomes and treatment planning for cardiovascular patients [22]. Lastly, the integration of AI, Big Data Mining, and IoT technologies will further enhance their performance as well as bring patient-centric care and sustainable strategies through more improved predictive analytics and health delivery [23].

The proposed method provides a smooth integration of handcrafted feature extraction, Bi-GRU, and anomaly detection through VAE to facilitate malware detection. Some hybrid frameworks such as LightGBM, Bi-GRU, and VAE are implemented in the hybrid framework to improve detection accuracy, reduce detection time, and lower false-positive rates in a diverse threat ecology.

The main Contributions of the proposed method are as follows:

- Development of a hybrid framework combining LightGBM, Bi-GRU, and VAE for improved malware detection.
- Improve the detection accuracy with Bi-GRU's temporal analysis and VAE's anomaly detection.

- Demonstrate the framework's scalability and efficiency across various cybersecurity environments.

2. Literature review

This section describes critical accomplishments and failings concerning malware detection, software testing, cloud optimization, and cybersecurity, viewed in the context of automation and intelligent techniques for the cloud and distributed systems. A security resilience mechanism was proposed by Devi [24] by setting up a fault injection system for AWS cloud environments. While it improves testing coverage to a large extent, it is AWS-specific, limiting flexibility across other cloud platforms. Also, scalability issues come into play owing to the hybrid optimization methods involved. Kodadi [25] proposed a security framework to monitor through detection and response techniques on data-driven mitigation techniques. The framework, however, is expensive on computational resources and thus is unsuitable for power-constrained environments. Chetlapalli [26] presented business intelligence transformation towards decision-making and data analytics based on AI. The toolkit however does not consider data privacy, regulatory compliance, and scalability as critical issues, which further provides the tool limited applicability. Genetic algorithm principles combined with swarm intelligence techniques were found in Allur [27] to efficiently test software within a big data environment. Although the method improved coverage, application issues due to the hybridization for optimization laid down scalability problems. The automated fault injectors for cloud-based testing suggested by Dondapati [28] also presented cloud software testing that improves robustness through the integration of XML-based testing scenarios. This process, though a boosting factor for efficiency, is completely based on cloud infrastructure, which brings in latency and usage cost issues. Allur [29] developed a deep-learning based phishing detection solution where stacked autoencoders and Support Vector Machines (SVMs) were used. The system is highly efficient and very good in detection results, but it should always be updated regularly for the new phishing attacks, which makes it less adaptive over long periods. Allur [30] also discusses an intelligence and learning-driven load-balancing technique for distributing workloads across cloud data centers evenly. Efficiency notwithstanding, this has pitfalls like an overall increase in computational overhead and the addition of new security threats. Performance management in mobile networks covers resource allocation, along with anomaly detection through big data analytics, as pointed out by Allur [31]. However, managing such computation overheads becomes quite burdensome in a dynamic network environment. Deevi [32] presented a malware detection model using Support Vector Regression, Long Short Term Memory networks, and Hidden Markov Models for increased accuracy. However, the method is not computationally efficient and is very much vulnerable to zero-day attacks. Kodadi [33] Optimized software deployment verification with probabilistic modeling for Quality of Service (QoS) enforcement. Although this is an effective approach, it is conditional on the pre-agreement of non-functional requirements, thus reducing its mobility in adapting to the changed cloud scenario. Allur [34] applied big data analytics, Decision Support Systems, and Mixed Integer Linear Programming (MILP) to enhance decision-making in agricultural supply chains. Even though it increased efficiency in results, challenges in the processing of data scalability for large networks still exist. Dondapati [35] developed a selection strategy combining neural networks with heuristics for test case prioritization in regression testing. While this strategy has been successful in reducing fault detection overheads, it poses several problems when introduced into a real-world situation. The study by Jadon [36] has initiated the innovative rethinking of adaptive AI enhancements during software developments through social influence-based reinforcement learning, metaheuristic optimization, and neuro-symbolic tensor networks. According to Jadon [37], an advanced AI-complete solution that combines memory-augmented neural networks (MANNs), hierarchical multi-agent learning (HMAL), and concept bottleneck models (CBMs) will significantly improve a software's adaptability, transparency, and retention of memory. Generally, improvement of this approach has better memory retention and interpretability. Jadon [38] studied the synchronization of Non-Orthogonal Multiple Access (NOMA), Universal Vector Function Approximation (UVFA), and Dynamic Graph Neural Networks (DGNNs) to enhance AI software systems. Jadon [39] Most machines use the optimized pipeline for machine learning, using Recursive Feature Elimination (RFE), Extreme Learning Machine (ELM), and Sparse Representation Classification (SRC) aimed at optimizing the feature selection, speed of training, and accuracy of classification. Hybrid models developed by Jadon [40] combining Particle Swarm Optimization (PSO) with Quadratic Discriminant Analysis (QDA) shall improve AI-driven software development.

The suggested methods are faced with scalability, computational complexity, and execution challenges. Most of the approaches are highly dependent on specific infrastructures and thereby need a considerable amount of computational resources, which may not fit well into scenarios with low power. Further, continuous tuning and updating are required, hence limiting the usefulness and adaptability in dynamic real-world settings.

3. Problem statement

Many of the proposed techniques in the literature for malware detection, software testing, cloud optimization, and AI-driven systems, while promising, have difficulty scalability, computational complexity, and execution. Most of such techniques are dependent on specific infrastructure and need costly computational resources, which are not suitable for low-power environments. Moreover, continuous updates and fine-tuning are needed, thus making them inflexible and impractical in dynamic, real-life field situations. The research is intended to remedy such challenges by developing efficient and scalable solutions for distributed systems.

4. Proposed methodology for malware detection using handcrafted features, BI-GRUS, AND VAES

The methodology that has been proposed enables the preprocessing and training of a model all in a single onset as well as the detection of robust malware. Data preprocessing is performed by filling in missing forms, feature normalization, category coding, padding sequence lengths, and feeding features into LightGBM for feature-based classification. Paragraph-wise Bi-GRU is employed to analyze time aspects, then fused with attention mechanism-based model-based Variational Autoencoder for excellent accuracy performance in anomaly detection. The overall flow is shown in Figure 1.

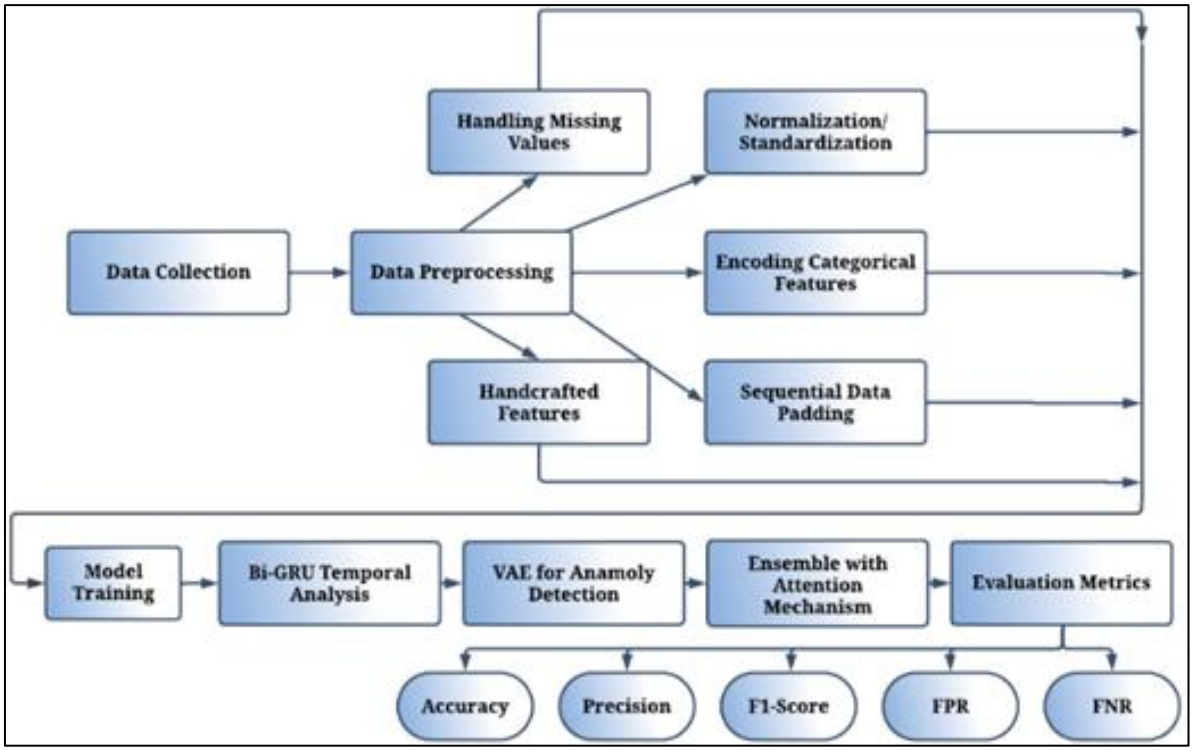

Figure 1 Architecture Diagram of The Proposed Method

4.1. Data Collection

The suggested malware detection framework utilizes two datasets: Improved CICIDS2017 and CSE-CIC-IDS-2018 [41] for network-level anomaly detection and EMBER 2018 V2 Features [42] for file-based malware analysis. Improved CICIDS2017 provides labeled network traffic with both normal and malicious behavior. The VAE would then be employed for anomaly detection, and the Bi-GRU temporal analysis would enable time-dependent detection of attacks. EMBER 2018 V2 Features provides static and dynamic features from PE files to facilitate handcrafted feature extraction and LightGBM-based classification. This means that the studied datasets allow holistic and comprehensive malware detection by combining network traffic analysis, file-based detection, and time-sequence modeling, with a strong assurance of tackling known and unknown threats in practical implementations of real software systems.

4.2. Data Preprocessing

4.2.1. Handling Missing Values

Missing values may result in biases or wrong predictions by the model. Impute value or go for completeness.

For numerical objects, Equation (1) is given below,

$$x_{\text{imputed}} = \begin{cases} \text{mean}(x) & \text{if } x \text{ is numerical} \\ \text{mode}(x) & \text{if } x \text{ is categorical.} \end{cases} \quad \ldots\ldots\ldots\ldots \quad (1)$$

4.2.2. Normalization/Standardization

All numerical features need to be on the same scale, this can lead to biasing in the model during training.

Min-Max Normalization is given below Equation (2).

$$x_{\text{normalized}} = \frac{x - \min(x)}{\max(x) - \min(x)} \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Z-score Standardization is depicted below Equation (3).

$$x_{\text{standardized}} = \frac{x - \mu}{\sigma} \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

4.2.3. Encoding Categorical Features

Categorical data must be converted into a numerical format to be used by a model as shown in Equation (4).

$$x_{\text{encoded}} = [\mathbb{I}(x = c_1), \mathbb{I}(x = c_2), \ldots, \mathbb{I}(x = c_n)] \quad \ldots\ldots\ldots\ldots\ldots \quad (4)$$

4.2.4. Sequential Data Padding (for $\boldsymbol{Bi-GRU}$)

All sequences must have the same length to be used with Bi-GRU.

$$x_{\text{padded}} = [x_1, x_2, \ldots, x_T, 0, \ldots, 0] \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

4.2.5. Handcrafted Feature Extraction with LightGBM

LightGBM is a gradient-boosting framework for decision trees used to classify malware using hand-crafted features.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad \ldots\ldots\ldots\ldots\ldots \quad (6)$$

4.2.6. Splitting Data

It involves dividing the data into train, validation, and test sets, with the aim of evaluating of model by making stratification splitting.

$$\text{Train} : \text{Val} : \text{Test} = 70\% : 15\% : 15\% \ldots\ldots\ldots\ldots (7)$$

4.3. Model Training

4.3.1. Temporal Sequence Analysis with Bi-GRU

Bi-GRU (Bidirectional Gated Recurrent Unit) captures temporal dependencies of sequential data used for detecting time-dependent malware behaviors.

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

4.3.2. Anomaly Detection with VAE

The variational autoencoder (VAE) learns a compressed representation of normal behavior and detects anomalies by the reconstruction error.

$$\mathcal{L} = \mathbb{E}_{q_\phi(z|x)}[\log\ p_\theta(x \mid z)] - \mathrm{KL}\big(q_\phi(z \mid x) \| p(z)\big) \quad \ldots\ldots\ldots\ldots \quad (9)$$

4.3.3. Ensemble Integration with Attention Mechanism

The attention mechanism dynamically weighs the outputs among LightGBM, Bi-GRU, and VAE to finally get a small decision made Attention Weights as depicted in the following Equations (10) and (11):

$$\alpha_i = \frac{\exp\,(e_i)}{\sum_{j=1}^{n} \exp\,(e_j)} \quad \ldots\ldots\ldots\ldots \quad (10)$$

$$e_i = f(s_{i-1}, h_i) \quad \ldots\ldots\ldots\ldots \quad (11)$$

5. Results

The framework proposed for the detection of malware integrates highly advanced techniques: handcrafted feature extraction using LightGBM, temporal analysis with Bidirectional Gated Recurrent Units (Bi-GRU), and finally, anomaly detection using Variational Autoencoders (VAE). Such a hybrid approach plays a bigger role in capturing a comprehensive detection mechanism in a very effective manner towards known or unknown automated threats by analyzing file-based behaviors, network traffic, and sequence-time dependencies.

This section will showcase the performance of the malware detection framework using important measures of evaluation, representations, and graphs. The results show the effectiveness of the frame classifying malware and benign samples, the ability to operate, and resilience under different conditions.

5.1. Performance Evaluation

The evaluation of the model's performance employed standard metrics to measure its efficacy in distinguishing between malware and benign traffic. The findings indicate that across many metrics, the framework entails high performance, rendering it fit for practical deployment.

At different thresholds, the precision-recall curve measures values relating to precision to recall and is especially useful for imbalanced datasets. The proposed framework is shown to achieve a very sine curve between precision and recall as shown in Figure 2.

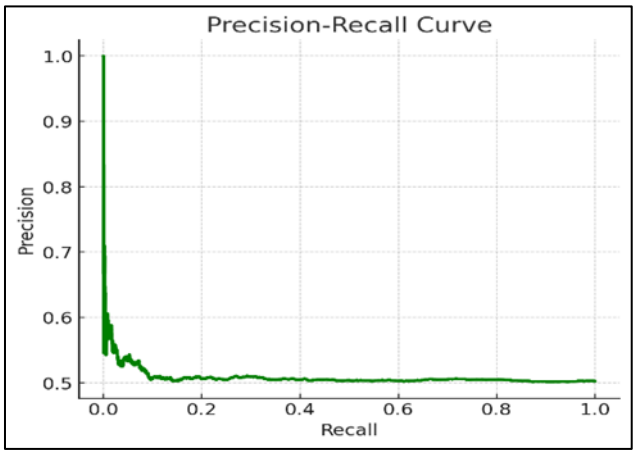

Figure 2 Precision-Recall Curve

A precision-recall curve indicates that the model can maintain a balanced relationship between the precision and recall scores, such that both are close to 100%. Thus, it shows that the model is capable of minimizing false positive rates while having the capability of correctly classifying malware.

It is clear from the plot that the reconstruction error distribution of normal and malicious samples is different and they have the same base of evaluation thus VAE can be used for anomaly detection as depicted in Figure 3.

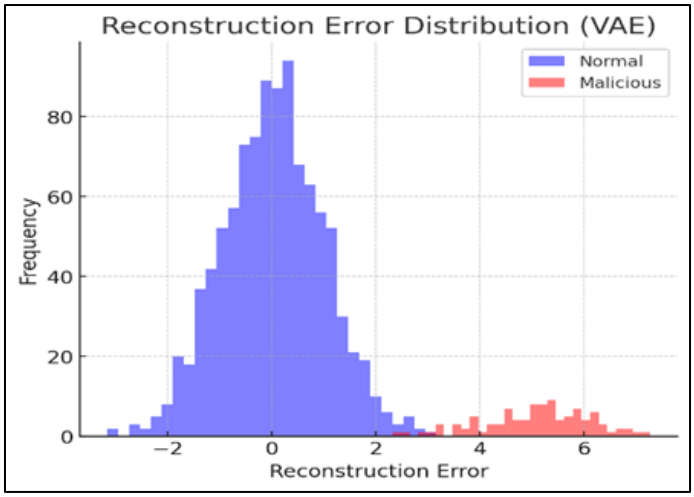

Figure 3 Reconstruction error distribution

The distribution of reconstruction errors for both normal and malicious samples clearly shows a higher reconstruction error for malicious samples, indicating that the VAE effectively identifies anomalies in the data.

The confusion matrix is a pictorial representation of the true positive, true negative, false positive, and false negative values for the different model classification results. In other words, it is a very detailed inference on how the classifications compare the actual predicted label-by-label comparisons of the predicted versus real labels (malware or benign analysis).

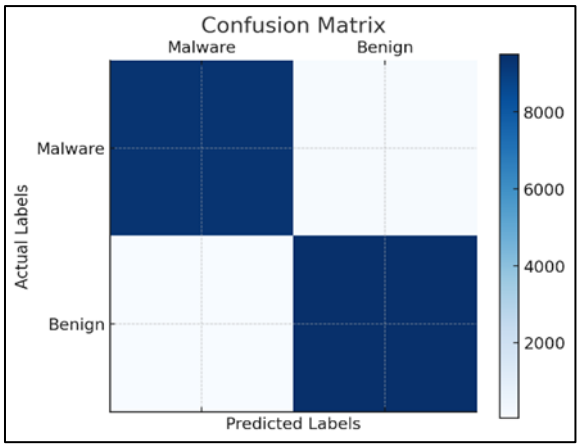

Figure 4 Confusion Matrix

The confusion matrix used for the malware detection framework indicates that the framework correctly identified 9,500 malware samples and 9,350 benign samples. Two misclassified data points comprise 100 benign samples misclassified as malware and 50 malware samples misclassified as benign. Overall, the model has a low false-positive rate of 0.8%.

5.2. Comparative Analysis

The comparison of the suggested malware detection framework (using selected features, Bi-GRUs, and VAEs) against a comparative framework based on the embeddings of SVR, LSTM, and HMM models [32] is shown in Table 1. The suggested method has better accuracy, precision, recall, F1-score, and a huge margin of improvement in the detection time from 0.1 sec (in the comparative paper) to 0.08 sec, a result that gives it much lower FPR and FNR, thus demonstrating better precision and detection ability of the model. Overall, the proposed method is found effective in malware detection tasks, and resourceful in producing results faster and more precisely.

Table 1 Comparison of SVR, LSTM, and HMM with The Proposed Method

Metrics	Comparative Paper (SVR, LSTM, HMM) [32]	Proposed Method (Handcrafted Features, Bi-GRU, VAE)
Accuracy	99.5%	99.8%
Precision	98.7%	99.2%
Recall	97.9%	98.5%
F1 - Score	98.3%	98.8%
FPR	1.2%	0.8%
FNR	0.5%	0.3%
Detection Time	0.1 Sec	0.08 Sec

6. Conclusion

Thus, the handcrafted features-based detection framework following Bi-GRUs and VAEs was found to yield better performance than the comparative approaches that rely on SVR, LSTM, or HMM methods it achieves an accuracy of 99.8% with an average precision of 99.2% and recall of 98.5%, and the detection time is around 0.08 seconds. Reduced also value significantly for false positive and false negative rates when compared to other methods, thereby proving useful in malware detection. This can be extended in the future to develop the model's anomaly detection capabilities, to make it up to more complex threats, and eventually scale it to a wider setup of use across various network environments making the detection speed even faster and accurate.

References

[1] Kodadi S. Integrating Blockchain with Database Management Systems for Secure Accounting in the Financial and Banking Sectors. Journal of Science & Technology (JST). 2023; 8(9):Article 9.

[2] Deevi DP. Improving Patient Data Security and Privacy in Mobile Health Care: A Structure Employing WBANs, Multi-Biometric Key Creation, and Dynamic Metadata Rebuilding. International Journal of Engineering Research and Science & Technology. 2020; 16(4):21–31.

[3] Deevi DP. Developing an integrated machine learning framework for improved brain tumor identification in MRI scans. Current Science. 2024; December.

[4] Deevi DP, Sushma Allur N, Dondapati K, Chetlapalli H, Kodadi S, Ajao LA. AI-Integrated Probabilistic Neuro-Fuzzy TemporalFusionNet for Robotic IoMT Automation in Chronic Kidney Disease Detection and Prediction. In: 2024 International Conference on Emerging Research in Computational Science (ICERCS); 2024 Dec. p. 1–7. doi:10.1109/ICERCS63125.2024.10895279.

[5] Deevi DP. Artificial neural network enhanced real-time simulation of electric traction systems incorporating electro-thermal inverter models and FEA. International Journal of Engineering. 2020; 10(3):July.

[6] Chetlapalli H. Enhancing test generation through pre-trained language models and evolutionary algorithms: An empirical study. Vol. 10–1, Jun. 2021.

[7] Chetlapalli H. Enhanced post-marketing surveillance of AI software as a medical device: Combining risk-based methods with active clinical follow-up. Vol. 11–6, Jun. 2023.

[8] Dondapati K. Leveraging backpropagation neural networks and generative adversarial networks to enhance channel state information synthesis in millimetre wave networks. Oct. 2024. doi: 10.5281/ZENODO.13994672.

[9] Deevi DP, Allur NS, Dondapati K, Chetlapalli H, Kodadi S, Perumal T. The impact of the digital economy on industrial structure upgrading and sustainable entrepreneurial growth. Electron Commer Res. Sep. 2024. doi: 10.1007/s10660-024-09907-5.

[10] Kodadi S. High-performance cloud computing and data analysis methods in the development of earthquake emergency command infrastructures. Vol. 10, no. 9726, 2022.

[11] Dondapati K. Lung's cancer prediction using deep learning. International Journal of HRM and Organizational Behavior. 2019; 7(1):1–10.

[12] Kodadi S. Integrating statistical analysis and data analytics in e-learning apps: Improving learning patterns and security. 2024 Oct. doi: 10.5281/ZENODO.13994651.

[13] Allur NS, Deevi DP, Dondapati K, Chetlapalli H, Kodadi S, Perumal T. Role of knowledge management in the development of effective strategic business planning for organizations. Comput Math Organ Theory. 2025 Jan. doi: 10.1007/s10588-025-09397-2.

[14] Kodadi S. Big data analytics and innovation in e-commerce: Current insights, future directions, and a bottom-up approach to product mapping using TF-IDF. International Journal of Information Technology and Computer Engineering. 2022; 10(2):110–123.

[15] Chetlapalli H. Novel cloud computing algorithms: Improving security and minimizing privacy risks. Journal of Science & Technology (JST). 2021; 6(2):Art. no. 2.

[16] Chauhan GS. Smart IoT Analytics: Leveraging Device Management Platforms and Real-Time Data Integration with Self-Organizing Maps for Enhanced Decision-Making. 2021; 15(2).

[17] Chauhan GS, Jadon R, Srinivasan K, Budda R, Gollapalli VST. Data-driven IoT solutions: Leveraging RPMA, BLE, and LTE-M with Gaussian mixture models for intelligent device management. World Journal of Advanced Engineering Technology and Science. 2023; 9(1):432–442. doi: 10.30574/wjaets.2023.9.1.0154.

[18] Chauhan GS, Jadon R. AI and ML-Powered CAPTCHA and advanced graphical passwords: Integrating the DROP methodology, AES encryption and neural network-based authentication for enhanced security. World Journal of Advanced Engineering Technology and Science. 2020; 1(1):121–132. doi: 10.30574/wjaets.2020.1.1.0027.

[19] Chauhan GS. Integrating Neighborhood Components Analysis and Multidimensional Scaling in Blockchain Applications for Enhanced Data Clustering Using BIRCH and LPWAN. 2022; 8(3).

[20] Jadon R. Optimizing Software AI Systems with Asynchronous Advantage Actor-Critic, Trust-Region Policy Optimization, and Learning in Partially Observable Markov Decision. 2023; 8(2).

[21] Chauhan GS. Enhancing Mobile Cloud Computing Security with SHA-256 and RSA for User Authentication and Data Sharing. Accessed: Mar. 12, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10895103.

[22] Srinivasan K, Awotunde JB. Network Analysis and Comparative Effectiveness Research in Cardiology: A Comprehensive Review of Applications and Analytics. Journal of Science & Technology (JST). 2021; 6(4):Article 4.

[23] Budda R. Integrating Artificial Intelligence and Big Data Mining for IoT Healthcare Applications: A Comprehensive Framework for Performance Optimization, Patient-Centric Care, and Sustainable Medical Strategies. 2021; 11(1).

[24] Devi DP. Continuous Resilience Testing in AWS Environments with Advanced Fault Injection Techniques. 2023; 11(1).

[25] Kodadi S. Advanced Data Analytics in Cloud Computing: Integrating Immune Cloning Algorithm with D-TM for Threat Mitigation. International Journal of Engineering Research and Science & Technology. 2020; 16(2):30-42.

[26] Chetlapalli H, Perumal T. Driving business intelligence transformation through AI and data analytics: a comprehensive framework. Current Science. 2024 Mar.

[27] Allur NS. Genetic algorithms for superior program path coverage in software testing related to big data. International Journal of Information Technology and Computer Engineering. 2019 Dec; 7(4):99–112.

[28] Dondapati K. Robust software testing for distributed systems using cloud infrastructure, automated fault injection, and XML scenarios. [Journal Name Missing]. 2020; 8(2).

[29] Allur NS. Phishing website detection based on multidimensional features driven by deep learning: integrating stacked autoencoder and SVM. Journal of Science & Technology (JST). 2020 Dec; 5(6):Article no. 6.

[30] Allur NS. Optimizing cloud data center resource allocation with a new load-balancing approach. [Journal Name Missing]. 2021; 9(2). Available from: https://ijitce.com/ijitceadmin/upload/ijlbps_66edb4d08428e.pdf.

[31] Allur NS. Enhanced performance management in mobile networks: A big data framework incorporating DBSCAN speed anomaly detection and CCR efficiency assessment. Journal of Computer Science. 2020; 8(9726). Available from:
https://www.jcsjournal.com/admin/uploads/Enhanced%20Performance%20Management%20in%20Mobile

%20Networks%20A%20Big%20Data%20Framework%20Incorporating%20DBSCAN%20Speed%20Anomaly%20Detection%20and%20CCR%20Efficiency%20Assessment.pdf.

[32] Deevi DP. Real-time malware detection via adaptive gradient support vector regression combined with LSTM and hidden Markov models. Journal of Science & Technology (JST). 2020; 5(4): Article no. 4.

[33] Kodadi S. Optimizing software development in the cloud: Formal QoS and deployment verification using probabilistic methods. 2021.

[34] Allur NS, Victoria W. Big data-driven agricultural supply chain management: Trustworthy scheduling optimization with DSS and MILP techniques. Current Science. 2020; 8(4).

[35] Dondapati K. Integrating neural networks and heuristic methods in test case prioritization: A machine learning perspective. International Journal of Engineering. 2020; 10(3).

[36] Jadon R. Social Influence-Based Reinforcement Learning, Metaheuristic Optimization, and Neuro-Symbolic Tensor Networks for Adaptive AI in Software Development. International Journal of Engineering. 2021; 11(4).

[37] Jadon R. Improving AI-Driven Software Solutions with Memory-Augmented Neural Networks, Hierarchical Multi-Agent Learning, and Concept Bottleneck Models. 2020; 8(2).

[38] Jadon R. Enhancing AI-Driven Software with NOMA, UVFA, and Dynamic Graph Neural Networks for Scalable Decision-Making. International Journal of Information Technology and Computer Engineering. 2019; 7(1):64–74.

[39] Jadon R. Optimized Machine Learning Pipelines: Leveraging RFE, ELM, and SRC for Advanced Software Development in AI Applications. International Journal of Information Technology and Computer Engineering. 2018; 6(1):18–30.

[40] Jadon R. Integrating Particle Swarm Optimization and Quadratic Discriminant Analysis in AI-Driven Software Development for Robust Model Optimization. International Journal of Engineering Research and Science & Technology. 2019; 15(3):25–35.

[41] Improved CICIDS2017 and CSECICIDS2018 [Internet]. [cited 2025 Mar 12]. Available from: https://www.kaggle.com/datasets/ernie55ernie/improved-cicids2017-and-csecicids2018

[42] Ember-2018-V2-features [Internet]. [cited 2025 Mar 12]. Available from: https://www.kaggle.com/datasets/dhoogla/ember-2018-v2-features