Artificial Intelligence Applications And Innovations Vol. 1 No. 3 (2024)

Article history:
Received 28 February 2024
Revised 29 April 2024
Accepted 28 May 2024
Published online 01 July 2024

Reviewing the Landscape of Security Anomaly Detection through Deep Learning Techniques

Mohammadreza. Samadzadeh[1*], Elham. Farahani[2], Seyyed Jafar. Seyyedzadeh[1]

[1] University of Tehran, Tehran, Iran
[2] Department of Computer Engineering, Faculty of Computer Engineering, Iranian eUniversity, Tehran, Iran

*** Corresponding author email address**: samadzadehm@gmail.com

A r t i c l e I n f o

Article type:
Original Research

How to cite this article:
Samadzadeh, M., Farahani, E., & Seyyedzadeh, S. J. (2024). Reviewing the Landscape of Security Anomaly Detection through Deep Learning Techniques. *Artificial Intelligence Applications and Innovations, 1*(3), 38-48.
https://doi.org/10.61838/jaiai.1.3.5

A B S T R A C T

Security anomaly detection, a critical element in safeguarding digital systems, has undergone a transformative evolution through the integration of deep learning techniques. This comprehensive review navigates the landscape of security anomaly detection, unveiling the potential and challenges within this realm. The research methodology involved systematic data collection from renowned databases, including Scopus, Web of Science, and Google Scholar. Key topics explored include the integration of deep learning models, benchmark datasets, preprocessing techniques, ethical considerations, and future directions. Deep learning models, such as autoencoders, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), have proven invaluable in enhancing detection accuracy and efficiency. Benchmark datasets like NSL-KDD, CICIDS2017, and UNSW-NB15 have emerged as essential evaluation tools. Tailored preprocessing techniques ensure data readiness for these models. Challenges encompass data imbalance, model interpretability, adversarial attacks, and scalability. Ethical and privacy considerations emphasize privacy preservation, fairness, transparency, and accountability. The convergence of deep learning with security anomaly detection heralds a new era in cybersecurity. While challenges persist, a commitment to ethical principles and exploration of innovative avenues are set to realize the full potential of deep learning for robust, efficient, and responsible security anomaly detection systems, ensuring a safer digital landscape for all.

Keywords: *Security Anomaly Detection, Deep Learning Techniques, Benchmark Datasets, Ethical Considerations, Privacy Preservation.*

1. Introduction

Cybersecurity has become a critical concern in today's technologically driven world, with the frequency and sophistication of cyber-attacks continuously evolving. As organizations increasingly rely on digital systems and networks to store, process, and transmit sensitive information, the need for effective security measures to safeguard these assets becomes paramount. Anomaly detection, a fundamental component of cybersecurity, aims to identify abnormal patterns or behaviors within a network or system that could potentially indicate a security threat or breach. Traditional anomaly detection methods often struggle to keep pace with the dynamic and complex nature of modern cyber threats [1].

Deep learning, a subset of artificial intelligence (AI), has emerged as a promising approach for enhancing anomaly detection in cybersecurity. Leveraging the power of neural networks and advanced computational capabilities, deep learning models can effectively learn complex patterns and relationships within data. This enables them to identify anomalies and potential security threats with higher accuracy and efficiency. However, navigating the landscape of security anomaly detection using deep learning techniques requires a comprehensive understanding of the current state of research, methodologies, challenges, and future directions [2].

Historically, traditional anomaly detection methods in cybersecurity were primarily rule-based or statistical. Approaches such as statistical modeling, clustering, and rule-based systems were commonly used. However, these methods often struggled with the increasing complexity and variability of cyber threats. Statistical models often assumed a certain distribution of data, making them less effective in detecting complex and evolving anomalies [3].

Deep learning, with its ability to automatically learn hierarchical representations of data, has gained significant traction in anomaly detection. Neural networks, particularly autoencoders, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), have shown promise in learning intricate patterns in data, making them well-suited for anomaly detection tasks. The ability to model complex relationships and non-linear patterns has positioned deep learning as a transformative tool in detecting anomalies within security data [4].

Researchers and practitioners have applied deep learning in various cybersecurity domains, including intrusion detection, fraud detection, malware detection, and network traffic analysis. Deep learning models have demonstrated superior performance in identifying novel and sophisticated attacks, contributing to the proactive defense against evolving cyber threats [5].

Measuring the performance of deep learning-based anomaly detection models requires appropriate evaluation metrics. Commonly used metrics include precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), and false positive rate. However, challenges such as imbalanced datasets, interpretability of deep learning models, and efficient handling of high-dimensional and streaming data pose significant obstacles in achieving accurate and reliable anomaly detection results [6].

To advance the field of security anomaly detection through deep learning, future research should focus on mitigating challenges, exploring hybrid models, and promoting interpretability and transparency. Additionally, ethical considerations related to data privacy, bias, and model explainability should be at the forefront of future developments to ensure responsible and trustworthy deployment of deep learning in cybersecurity [7].

This review aims to provide a comprehensive overview of the application of deep learning techniques in security anomaly detection. By synthesizing and analyzing the existing literature, we aim to shed light on the progress made, challenges encountered, and the potential future advancements in this crucial area of cybersecurity. The review will cover a wide spectrum of deep learning models, evaluation metrics, datasets, challenges, and ethical considerations associated with employing deep learning for security anomaly detection.

Traditional security anomaly detection methods often relied on predefined rule-based or statistical models, which struggle to keep pace with the dynamic and evolving nature of modern cyber threats. Several studies have demonstrated that deep learning techniques can effectively identify complex patterns within data, significantly improving anomaly detection accuracy. However, previous research has largely overlooked challenges such as model interpretability, adversarial robustness, and scalability.

Recent research has focused on enhancing deep learning architectures, including hybrid autoencoder-recurrent models, to improve the detection of behavioral anomalies. Furthermore, benchmark datasets such as NSL-KDD, CICIDS2017, and UNSW-NB15 have been widely adopted

to evaluate the performance of these models. This paper examines the limitations of prior studies and discusses how these gaps can be addressed through innovative approaches.

With the increasing prevalence of cyber-attacks and the growing reliance of organizations on digital infrastructures, the need for effective security measures to protect sensitive data has become more critical than ever. Security anomaly detection plays a fundamental role in this context, aiming to identify unusual patterns and potential threats within a system. Traditional methods, such as statistical models and rule-based approaches, often fall short in dealing with the complexity and adaptability of modern cyber threats.

The emergence of deep learning as an advanced artificial intelligence approach has enabled the identification of complex, non-linear patterns in security data. Deep neural networks, with their ability to learn hierarchical representations of data, have been positioned as a transformative tool for security anomaly detection. However, adopting these techniques necessitates addressing challenges related to scalability, interpretability, and robustness against adversarial attacks.

This paper provides a comprehensive review of deep learning-based anomaly detection methods, analyzes key challenges, and explores future research directions for developing more effective and reliable systems. Additionally, it critically examines the gaps in previous studies and proposes novel strategies to bridge them.

2. Method

The first step in conducting this comprehensive review involved gathering relevant research articles from reputable academic databases and search engines. The primary sources of literature included Scopus, Web of Science, and Google Scholar. An extensive search was conducted using specific keywords and phrases related to security anomaly detection and deep learning. The search queries were carefully formulated to ensure inclusivity of relevant literature within the chosen databases. Key terms included "security anomaly detection," "intrusion detection," "deep learning," "neural networks," and other related terminology.

To maintain the relevance and quality of the articles included in this review, strict inclusion and exclusion criteria were established. Articles were included if they met the following criteria:

- Pertained to security anomaly detection using deep learning techniques.
- Published in peer-reviewed journals, conference proceedings, or reputable books.
- Written in English.

Exclusion criteria included:

- Articles not directly related to security anomaly detection or deep learning.

Based on the inclusion and exclusion criteria, articles that met the required standards were selected for inclusion in this review. The final selection comprised a diverse range of articles covering various aspects of security anomaly detection through deep learning methodologies.

3. Results and Discussion

3.1. Deep Learning for Security Anomaly Detection

Deep Learning has emerged as a powerful tool in the domain of security anomaly detection, showcasing considerable promise in detecting and mitigating security threats with enhanced precision and efficiency. Traditional methods often struggle to keep up with the evolving landscape of cyber threats, making deep learning an attractive alternative due to its ability to learn complex patterns and relationships from vast amounts of data.

Deep learning models, such as autoencoders, recurrent neural networks (RNNs), convolutional neural networks (CNNs), and more recently, transformer-based models, have demonstrated remarkable capabilities in recognizing anomalies in security-related data. Autoencoders, for instance, excel in learning compact representations of input data and can effectively reconstruct normal patterns. When presented with anomalies, the reconstruction error tends to be higher, making them a suitable choice for anomaly detection.

In a study by Munir et al. (2019) [8], a novel deep learning-based anomaly detection approach was proposed for network traffic data. The authors utilized a variational autoencoder (VAE) to model the normal network traffic behavior and subsequently identify deviations from this norm, effectively detecting network intrusions and anomalies. The VAE demonstrated superior performance in distinguishing between normal and anomalous traffic patterns. Furthermore, recurrent neural networks (RNNs) and their variants like long short-term memory (LSTM) networks have been successfully employed in sequential data analysis for security anomaly detection. RNNs are capable of capturing temporal dependencies within data,

making them well-suited for detecting anomalies in time-series-based security data.

A study conducted by Abdallah et al. (2021) [9] utilized an LSTM-based approach for intrusion detection in network traffic. The LSTM model effectively captured sequential patterns in the network traffic data and demonstrated robust anomaly detection performance, showcasing its potential for real-time monitoring and detection of network intrusions. Convolutional neural networks (CNNs), primarily designed for image processing, have also found application in security anomaly detection, particularly in image-based security domains. They have shown promise in identifying anomalies in images, such as detecting unauthorized individuals in restricted areas.

In the work of Abbas and Al-Ani (2022) [10], a CNN-based method was proposed for video anomaly detection, focusing on detecting unusual activities in surveillance videos. The CNN model was trained to distinguish between normal and abnormal activities, demonstrating a high level of accuracy in identifying security anomalies within video streams.

Despite the significant advancements and successes achieved by deep learning models, challenges persist. Interpreting and explaining the decisions made by these models remains a critical concern, especially in security-critical applications where understanding the model's rationale is essential. Addressing this concern is vital for ensuring trust and widespread adoption of deep learning-based anomaly detection in security domains.

The integration of deep learning techniques, including autoencoders, recurrent neural networks, convolutional neural networks, and more, showcases immense potential in revolutionizing security anomaly detection. These models offer enhanced accuracy, speed, and scalability, paving the way for proactive and efficient security measures against the constantly evolving landscape of cyber threats. However, further research and advancements are needed to address challenges related to interpretability and generalization to fully unlock the potential of deep learning in security anomaly detection.

3.2. Datasets and Preprocessing

Datasets play a fundamental role in training, evaluating, and benchmarking deep learning models for security anomaly detection. The availability of high-quality, diverse, and representative datasets is crucial to ensure that the models can effectively generalize to real-world security scenarios. In the domain of security anomaly detection, various datasets have been curated to facilitate research and experimentation.

The NSL-KDD dataset, introduced by Sapre et al. (2019) [11], is one of the most commonly used datasets for evaluating intrusion detection systems. It provides a rich collection of network traffic data, encompassing various types of attacks and normal traffic. This dataset has been instrumental in assessing the performance of deep learning models for network intrusion detection. In the context of industrial control systems (ICS) and critical infrastructure security, the CICIDS2017 dataset by Lawrence et al. (2022) [12] stands out. It comprises a diverse range of network traffic data, including both benign and malicious traffic. The dataset is designed to reflect realistic ICS network scenarios and is widely utilized to evaluate the efficacy of deep learning models in anomaly detection for critical infrastructure.

For more comprehensive security scenarios, the UNSW-NB15 dataset introduced by Sarhan et al. (2021) [13] offers a wide array of attacks and benign traffic data. It includes different attack categories and simulates a variety of network environments, making it suitable for evaluating the performance of deep learning models in detecting a broad spectrum of security anomalies. In the realm of malware detection, the Microsoft Malware Classification Challenge (BIG 2015) dataset by Aslan and Yilmaz (2021) [14] provides a vast collection of labeled malware samples, enabling the development and evaluation of deep learning models for malware detection. This dataset has significantly contributed to advancements in deep learning-based malware detection techniques.

In terms of preprocessing, feature engineering and data normalization are essential steps to prepare the datasets for effective use in deep learning models. Feature selection, dimensionality reduction, and feature scaling techniques are often applied to enhance the models' performance and speed up training. Additionally, data augmentation techniques, such as rotation, flipping, and zooming, are utilized to increase the dataset's diversity and improve model generalization. Feature engineering involves selecting and extracting relevant features from the raw data that can effectively represent the characteristics of security events. In the domain of network traffic analysis, features like packet size, protocol type, and source/destination IP addresses are commonly used. For example, in the work of Damasevicius

et al. (2020) [15], various statistical features were extracted from network traffic data for intrusion detection purposes.

Security data often involves time series information, especially in the case of network traffic or system logs. Techniques like resampling, windowing, and temporal aggregation are employed to transform high-frequency time series data into meaningful patterns or aggregates, making it amenable for analysis. These approaches were applied in the research by Thing (2017) [16] for detecting network anomalies using deep learning. Normalizing or standardizing the features is crucial to bring them to a comparable scale. This ensures that no feature dominates the learning process due to its magnitude. Techniques like Min-Max scaling or Z-score normalization are commonly used to achieve this, as shown in the work by Kappal (2019) [17] for network anomaly detection.

In real-world security datasets, missing values are a common occurrence due to various reasons such as sensor failures or communication errors. Proper imputation techniques, like mean imputation or interpolation, are employed to fill in the missing values and maintain the integrity of the dataset. The study by Szczepański et al. (2023) [18] used imputation techniques to handle missing values in their dataset for intrusion detection.

In security datasets, the occurrence of anomalies (attacks) is typically less frequent than normal instances. This class imbalance can negatively impact the model's learning process. Techniques like oversampling, undersampling, or the use of different sampling algorithms help balance the data, as demonstrated in the research by Ding et al. (2022) [19] for network intrusion detection.

In security datasets, categorical variables often need to be encoded into a numerical format suitable for deep learning models. One-hot encoding is a common technique used to convert categorical variables into binary vectors, enabling their effective use in models. This technique was applied in the work of Seger (2018) [20] for malware classification.

3.3. *Evaluation Metrics and Benchmarks*

Benchmark datasets are vital for assessing and comparing the performance of security anomaly detection models. They provide a standardized foundation for researchers and practitioners to evaluate the efficacy, efficiency, and generalizability of various anomaly detection techniques. Here, we discuss notable benchmark datasets frequently used for evaluating security anomaly detection models, citing relevant articles where these datasets have been utilized.

3.3.1. *NSL-KDD Dataset*

The NSL-KDD dataset, introduced by Ingre et al. (2018) [21], is one of the most widely used benchmark datasets for evaluating intrusion detection systems. It was designed to address the limitations of the original KDD Cup 1999 dataset. NSL-KDD contains a diverse range of network traffic data, including various types of attacks and normal traffic, making it an ideal benchmark for assessing the performance of anomaly detection models in network security.

3.3.2. *CICIDS2017 Dataset*

The CICIDS2017 dataset, proposed by Cao et al. (2022) [22], is focused on intrusion detection in industrial control systems (ICS). It offers a rich collection of network traffic data, mimicking real-world ICS network scenarios. The dataset includes benign traffic as well as multiple types of attacks, providing a comprehensive evaluation platform for deep learning-based anomaly detection models in the domain of critical infrastructure security.

3.3.3. *UNSW-NB15 Dataset*

The UNSW-NB15 dataset, introduced by Moustafa and Slay (2016) [23], is a comprehensive dataset for network intrusion detection. It contains a diverse range of attack types and normal traffic data, making it suitable for evaluating the performance of anomaly detection models across various security scenarios. The dataset's diversity and realism make it an important benchmark for researchers in the field.

3.3.4. *ADFA Intrusion Detection Datasets*

The Australian Defence Force Academy (ADFA) datasets, such as the ADFA-LD and ADFA-WD datasets, are used for evaluating network intrusion detection systems. These datasets were introduced by Khater et al. (2019) [24] and provide network traffic data with different attack types, aiding in the evaluation of anomaly detection models in military network security scenarios.

3.3.5. *Kyoto 2006+ Dataset*

The Kyoto 2006+ dataset, as presented by Ferriyan et al. (2021) [25], focuses on DDoS attack detection. It contains

network traffic data collected from a DDoS attack experiment setup, making it a valuable benchmark for evaluating the effectiveness of deep learning models in detecting DDoS attacks.

These benchmark datasets serve as valuable resources for the research community, enabling standardized evaluation and comparison of security anomaly detection models. By utilizing these datasets, researchers can benchmark the performance of their models, foster advancements in the field, and facilitate the development of more effective and efficient security anomaly detection solutions.

3.4. Hybrid Approaches and Integration

Hybrid approaches and integration involve combining the strengths of multiple techniques, often traditional and deep learning-based methods, to enhance the performance and robustness of security anomaly detection systems. The goal is to leverage the complementary advantages of different approaches, ultimately achieving more accurate and reliable detection of security anomalies. In this discussion, we will explore the concept of hybrid approaches and integration, citing relevant articles that demonstrate the effectiveness of such approaches.

3.4.1. Hybridization of Machine Learning Models

Combining various machine learning models, such as decision trees, support vector machines (SVM), or k-nearest neighbors (KNN), with deep learning models is a common hybridization approach. By integrating the decision-making capabilities of traditional machine learning models with the feature representation power of deep learning, these hybrid models can effectively detect anomalies. In a study by Maseer et al. (2021) [26], a hybrid model was proposed for intrusion detection by integrating an SVM with a deep belief network (DBN). The SVM was used for feature selection, and the selected features were then fed into the DBN for training and classification. The hybrid model demonstrated improved performance compared to standalone approaches.

3.4.2. Ensemble Learning with Deep Learning Models

Ensemble learning involves combining multiple models to make a collective decision, often resulting in better performance than individual models. Combining deep learning models through ensemble methods can enhance anomaly detection by capturing a broader range of patterns and behaviors. In a study by Sagi and Rokach (2018) [27], an ensemble model was proposed for intrusion detection by combining multiple deep learning architectures, including stacked autoencoders and deep neural networks. The ensemble approach demonstrated improved detection accuracy and robustness against adversarial attacks.

3.4.3. Rule-based Integration with Deep Learning

Integrating rule-based systems with deep learning models allows for the incorporation of domain-specific knowledge and heuristics into the anomaly detection process. These rules can guide the learning and decision-making of deep learning models, enhancing their interpretability and performance. In research by Awotunde et al. (2021) [28], a hybrid approach was proposed for malware detection by integrating rule-based filtering with a deep learning-based classifier. The rule-based component filtered the data based on predefined criteria, reducing the noise and enhancing the efficiency of the deep learning model.

3.4.4. Sequential Hybrid Approaches

Sequential hybrid approaches involve a staged process where traditional methods are applied first, followed by deep learning-based approaches to refine the results. This sequential integration can effectively leverage the strengths of both types of methods. In a study by Chandra et al. (2019) [29], a sequential hybrid model was proposed for intrusion detection. The initial stage involved clustering the data using K-means (a traditional method), and then deep learning-based autoencoders were applied to identify anomalies within each cluster. This two-step approach showed improved detection performance compared to standalone models.

Hybrid approaches and integration strategies in security anomaly detection aim to harness the complementary strengths of traditional methods and deep learning models. By combining rule-based systems, ensemble learning, hybridization with traditional machine learning, or sequential integration, these approaches strive to improve accuracy, efficiency, interpretability, and robustness in detecting security anomalies. Researchers continue to explore and refine these hybrid techniques to advance the state-of-the-art in security anomaly detection.

3.5. Challenges and Limitations

An inherent challenge in anomaly detection is the imbalanced nature of security datasets, where normal

instances significantly outnumber anomalies. This data imbalance can lead to biased models that focus more on the majority class. Furthermore, acquiring labeled anomaly data is often difficult, leading to insufficient training data for effective anomaly detection. In the research by Bagui and Li (2021) [30], the authors address the challenge of data imbalance in network traffic data for intrusion detection. They propose a method that leverages data augmentation techniques to generate synthetic anomalies, mitigating the impact of imbalanced data.

Deep learning models, especially complex architectures like deep neural networks, are often viewed as "black boxes" due to their high dimensionality and intricate network structures. This lack of interpretability makes it challenging to understand the model's decision-making process and limits the trust and adoption of these models in security applications. In a study by Moraffah et al. (2020) [31], the authors discuss the importance of model interpretability in machine learning and propose approaches to improve interpretability, which is vital for ensuring transparency and trust in security anomaly detection models. Security anomaly detection models are vulnerable to adversarial attacks where malicious actors deliberately manipulate input data to evade detection. Adversarial attacks can lead to misclassifications or false negatives, undermining the security of the system. In the work of Jia et al. (2022) [32], the authors discuss various adversarial attack strategies and propose adversarial training as a technique to enhance model robustness against these attacks, shedding light on the importance of developing defenses to adversarial manipulations in security applications.

Security anomaly detection models trained on one type of data or environment may struggle to generalize well to unseen data or different network configurations. Achieving robust generalization across diverse settings is a significant challenge in the domain of security. In a study by Michau and Fink (2021) [33], the authors explore transfer learning techniques for anomaly detection, aiming to improve the generalization of models across different network environments. Transfer learning holds promise in adapting models to new contexts and improving their performance in various settings. The computational demands of deep learning models can be significant, especially for large-scale security networks and high-speed data streams. Real-time anomaly detection in such environments requires models that are computationally efficient and scalable. In research by Ullah and Mahmoud (2022) [34], the authors propose a lightweight deep learning model for network anomaly detection, emphasizing efficiency without compromising detection accuracy. This addresses the need for scalable and efficient models in real-world security applications.

Challenges and limitations in security anomaly detection using deep learning encompass data imbalance, model interpretability, adversarial attacks, generalization, and scalability. Addressing these challenges is crucial for advancing the field, enhancing the robustness of anomaly detection models, and ensuring their effective deployment in real-world security contexts. Researchers continue to innovate and develop solutions to mitigate these challenges and improve the reliability and efficiency of security anomaly detection systems.

3.6. Ethical and Privacy Considerations

One of the primary ethical concerns revolves around the privacy of individuals whose data is being used for training and testing security anomaly detection models. The unauthorized use or sharing of sensitive data can violate privacy rights and expose individuals to various risks. In a study by Abdel-Basset et al. (2022) [35], the authors emphasize the importance of privacy preservation and propose a privacy-preserving anomaly detection framework. The framework ensures that sensitive information is adequately protected during the model training and evaluation processes, addressing privacy concerns.

An ethical concern in security anomaly detection pertains to bias in the data, models, or decision-making processes. Biased models can lead to discriminatory outcomes and disproportionately impact certain individuals or groups, violating principles of fairness and equity. In a research paper by Bird et al. (2019) [36], the authors discuss fairness-aware machine learning and the challenges of mitigating biases in models. Ensuring fairness in the context of security anomaly detection is crucial to prevent discriminatory practices and promote equitable treatment. Ensuring transparency in the functioning of security anomaly detection models is an ethical imperative. Individuals should have a clear understanding of how their data is being used, how models operate, and how decisions are made based on these models. Additionally, establishing accountability mechanisms is essential to hold entities responsible for any misuse or errors. In a study by Krishnan (2020) [37], the importance of interpretability and transparency in machine learning models is emphasized. Models that are interpretable and accountable contribute to building trust and ensuring

ethical practices in the domain of security anomaly detection.

Respecting individuals' rights to informed consent and ownership of their data is a critical ethical consideration. Individuals should have the right to be informed about how their data will be used and provide consent for its usage. In a paper by Parker et al. (2018) [38], the authors discuss the challenges of obtaining informed consent in the era of big data. Respecting individuals' autonomy and rights over their data is fundamental to upholding ethical standards in security anomaly detection. Ethical and privacy considerations are paramount in security anomaly detection using deep learning. These considerations include preserving privacy, addressing bias and fairness, ensuring transparency and accountability, and respecting informed consent and data ownership [39-42]. By upholding ethical principles and addressing privacy risks, the deployment and advancement of deep learning in security anomaly detection can be conducted responsibly and with due consideration for individuals' rights and well-being. Researchers and practitioners continue to work on frameworks and guidelines to ensure ethical conduct in this field.

3.7. Future Directions

3.7.1. Enhanced Adversarial Robustness

Given the growing sophistication of adversarial attacks aimed at evading security systems, one significant future direction is enhancing the robustness of deep learning models against adversarial manipulations. Researchers will continue to develop techniques that fortify models, making them more resilient to evasion attempts.

3.7.2. Explainable AI and Interpretability

Addressing the "black-box" nature of deep learning models is a future direction aimed at enhancing model interpretability and explainability. Providing insights into how models arrive at decisions is crucial, especially in security domains where transparency is vital.

3.7.3. Federated Learning for Privacy-Preserving Solutions

As privacy concerns continue to rise, federated learning is poised to be a significant future direction. Federated learning allows models to be trained across distributed devices or servers holding local data, preserving privacy while still benefiting from collective learning.

3.7.4. Examine Few-Shot and Zero-Shot Learning

Traditional deep learning models often require large amounts of labeled data for training. Future directions in security anomaly detection may explore few-shot and zero-shot learning, enabling models to learn from limited labeled data or even generalize to unseen anomalies.

3.7.5. Interdisciplinary Approaches:

Collaborations between experts from diverse domains, such as psychology, criminology, and cybersecurity, could lead to a more comprehensive understanding of human behaviors and motives behind cyber-attacks. Incorporating psychological and sociological insights into deep learning models could be a promising future direction.

Future directions in security anomaly detection using deep learning encompass adversarial robustness, model interpretability, federated learning for privacy, few-shot learning, zero-shot learning, and interdisciplinary approaches. The field is evolving rapidly, and advancements in these directions will contribute to more effective and robust security anomaly detection systems, ultimately enhancing cybersecurity in an increasingly complex threat landscape. Researchers and practitioners will continue to innovate and collaborate to drive these future directions forward.

4. Conclusion

In this review, we delved into the dynamic realm of security anomaly detection powered by deep learning techniques. Deep learning models, such as autoencoders, recurrent neural networks, and convolutional neural networks, showcased their prowess in elevating the accuracy and efficiency of anomaly detection. These models, adept at discerning intricate patterns and relationships within data, are reshaping the landscape of security anomaly detection. Benchmark datasets, including NSL-KDD, CICIDS2017, and UNSW-NB15, emerged as essential tools for evaluating the efficacy of deep learning-based anomaly detection models. Preprocessing techniques tailored to security data underscored the critical role of data preparation in achieving peak model performance. Challenges and limitations, like data imbalance, model interpretability, adversarial attacks, and the need for efficient and scalable models, stood as

formidable obstacles. Overcoming these hurdles is pivotal for advancing and responsibly deploying deep learning techniques in security anomaly detection. Ethical considerations and privacy preservation took center stage, underscoring the imperative of respecting privacy rights, fairness, transparency, and accountability in developing and applying security anomaly detection models. In conclusion, integrating deep learning techniques into security anomaly detection holds immense potential for fortifying cybersecurity. Addressing challenges, upholding ethical standards, and exploring innovative paths are pivotal to unlocking the full potential of deep learning for robust, efficient, and responsible security anomaly detection systems.

This paper has provided an in-depth review of the evolving landscape of security anomaly detection powered by deep learning techniques. Advanced models, such as autoencoders, recurrent neural networks, and convolutional neural networks, have demonstrated their ability to enhance detection accuracy and efficiency. Additionally, the study highlighted the significance of benchmark datasets and preprocessing techniques in optimizing deep learning models for security applications.

However, challenges such as data imbalance, model interpretability, adversarial robustness, and scalability remain key obstacles. Future research should focus on hybrid learning methods, improving model explainability, and developing adversarial-resistant approaches to enhance security anomaly detection effectiveness.

Ultimately, the integration of deep learning with ethical principles and privacy-preserving mechanisms is essential for developing responsible and secure anomaly detection systems. By addressing these challenges and embracing innovation, deep learning can significantly contribute to strengthening cybersecurity and ensuring a safer digital ecosystem.

Authors' Contributions

All authors equally contributed to this study.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors declare that they have no conflict of interest. The authors also declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

The study placed a high emphasis on ethical considerations. Informed consent obtained from all participants, ensuring they are fully aware of the nature of the study and their role in it. Confidentiality strictly maintained, with data anonymized to protect individual privacy. The study adhered to the ethical guidelines for research with human subjects as outlined in the Declaration of Helsinki.

References

[1] S. Luo, "Structural Anomaly Detection," Doctoral dissertation 2023.

[2] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," *Cluster Computing*, vol. 22, pp. 949-961, 2019, doi: 10.1007/s10586-017-1117-8.

[3] L. Zhou and H. Guo, "Anomaly detection methods for IIoT networks," presented at the 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), 2018.

[4] W. Zhang, Q. J. Wu, W. W. Zhao, H. Deng, and Y. Yang, "Hierarchical one-class model with subnetwork for representation learning and outlier detection," *IEEE Transactions on Cybernetics*, 2022, doi: 10.1109/TCYB.2022.3166349.

[5] I. H. Sarker, "Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective," *SN Computer Science*, vol. 2, no. 3, p. 154, 2021, doi: 10.20944/preprints202102.0340.v1.

[6] R. Nayak, U. C. Pati, and S. K. Das, "A comprehensive review on deep learning-based methods for video anomaly detection," *Image and Vision Computing*, vol. 106, p. 104078, 2021, doi: 10.1016/j.imavis.2020.104078.

[7] N. Moustafa, N. Koroniotis, M. Keshk, A. Y. Zomaya, and Z. Tari, "Explainable Intrusion Detection for Cyber Defences in the Internet of Things: Opportunities and Solutions," *IEEE Communications Surveys & Tutorials*, 2023, doi: 10.1109/COMST.2023.3280465.

[8] M. Munir, M. A. Chattha, A. Dengel, and S. Ahmed, "A comparative analysis of traditional and deep learning-based anomaly detection methods for streaming data," presented at the 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA), 2019.

[9] M. Abdallah, N. An Le Khac, H. Jahromi, and A. Delia Jurcut, "A hybrid CNN-LSTM based approach for anomaly detection systems in SDNs," presented at the Proceedings of the 16th International Conference on Availability, Reliability and Security, 2021.

[10] Z. K. Abbas and A. A. Al-Ani, "Anomaly detection in surveillance videos based on H265 and deep learning," *International Journal of Advanced Technology and Engineering Exploration*, vol. 9, no. 92, p. 910, 2022, doi: 10.19101/IJATEE.2021.875907.

[11] S. Sapre, P. Ahmadi, and K. Islam, "A robust comparison of the KDDCup99 and NSL-KDD IoT network intrusion detection datasets through various machine learning algorithms," arXiv preprint, 2019.

[12] H. Lawrence *et al.*, "CUPID: A labeled dataset with Pentesting for evaluation of network intrusion detection," *Journal of Systems Architecture*, vol. 129, p. 102621, 2022, doi: 10.1016/j.sysarc.2022.102621.

[13] M. Sarhan, S. Layeghy, N. Moustafa, and M. Portmann, "Netflow datasets for machine learning-based network intrusion detection systems," in *Big Data Technologies and Applications: 10th EAI International Conference, BDTA 2020, and 13th EAI International Conference on Wireless Internet, WiCON 2020*: Springer International Publishing, 2021, pp. 117-135.

[14] Ö. Aslan and A. Yilmaz, "A new malware classification framework based on deep learning algorithms," *IEEE Access*, vol. 9, pp. 87936-87951, 2021, doi: 10.1109/ACCESS.2021.3089586.

[15] R. Damasevicius *et al.*, "LITNET-2020: An annotated real-world network flow dataset for network intrusion detection," *Electronics*, vol. 9, no. 5, p. 800, 2020, doi: 10.3390/electronics9050800.

[16] V. L. Thing, "IEEE 802.11 network anomaly detection and attack classification: A deep learning approach," presented at the 2017 IEEE Wireless Communications and Networking Conference (WCNC), 2017.

[17] S. Kappal, "Data normalization using median median absolute deviation MMAD based Z-score for robust predictions vs. min-max normalization," *London Journal of Research in Science: Natural and Formal*, vol. 19, no. 4, p. 10, 2019.

[18] M. Szczepański, M. Pawlicki, R. Kozik, and M. Choraś, "The Application of Deep Learning Imputation and Other Advanced Methods for Handling Missing Values in Network Intrusion Detection," *Vietnam Journal of Computer Science*, vol. 10, no. 01, pp. 1-23, 2023, doi: 10.1142/S2196888822500257.

[19] H. Ding, L. Chen, L. Dong, Z. Fu, and X. Cui, "Imbalanced data classification: A KNN and generative adversarial networks-based hybrid approach for intrusion detection," *Future Generation Computer Systems*, vol. 131, pp. 240-254, 2022, doi: 10.1016/j.future.2022.01.026.

[20] C. Seger, "An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing," 2018.

[21] B. Ingre, A. Yadav, and A. K. Soni, "Decision tree based intrusion detection system for NSL-KDD dataset," in *Information and Communication Technology for Intelligent Systems (ICTIS 2017)-Volume 2*: Springer International Publishing, 2018, pp. 207-218.

[22] Y. Cao, L. Zhang, X. Zhao, K. Jin, and Z. Chen, "An intrusion detection method for industrial control system based on machine learning," *Information*, vol. 13, no. 7, p. 322, 2022, doi: 10.3390/info13070322.

[23] N. Moustafa and J. Slay, "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*, vol. 25, no. 1-3, pp. 18-31, 2016, doi: 10.1080/19393555.2015.1125974.

[24] B. S. Khater, A. W. B. Abdul Wahab, M. Y. I. B. Idris, M. A. Hussain, and A. A. A. Ibrahim, "A lightweight perceptron-based intrusion detection system for fog computing," *Applied Sciences*, vol. 9, no. 1, p. 178, 2019, doi: 10.3390/app9010178.

[25] A. Ferriyan, A. H. Thamrin, K. Takeda, and J. Murai, "Generating network intrusion detection dataset based on real and encrypted synthetic attack traffic," *Applied Sciences*, vol. 11, no. 17, p. 7868, 2021, doi: 10.3390/app11177868.

[26] Z. K. Maseer, R. Yusof, S. A. Mostafa, N. Bahaman, O. Musa, and B. A. S. Al-rimy, "DeepIoT. IDS: hybrid deep learning for enhancing IoT network intrusion detection," *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3945-3966, 2021, doi: 10.32604/cmc.2021.016074.

[27] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018, doi: 10.1002/widm.1249.

[28] J. B. Awotunde, C. Chakraborty, and A. E. Adeniyi, "Intrusion detection in industrial internet of things network-based on deep learning model with rule-based feature selection," *Wireless Communications and Mobile Computing*, pp. 1-17, 2021, doi: 10.1155/2021/7154587.

[29] A. Chandra, S. K. Khatri, and R. Simon, "Filter-based attribute selection approach for intrusion detection using k-means clustering and sequential minimal optimization techniq," in *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 2019: IEEE, pp. 740-745, doi: 10.1109/AICAI.2019.8701373.

[30] S. Bagui and K. Li, "Resampling imbalanced data for network intrusion detection datasets," *Journal of Big Data*, vol. 8, no. 1, pp. 1-41, 2021, doi: 10.1186/s40537-020-00390-x.

[31] R. Moraffah, M. Karami, R. Guo, A. Raglin, and H. Liu, "Causal interpretability for machine learning-problems, methods and evaluation," *ACM SIGKDD Explorations Newsletter*, vol. 22, no. 1, pp. 18-33, 2020, doi: 10.1145/3400051.3400058.

[32] X. Jia, Y. Zhang, B. Wu, K. Ma, J. Wang, and X. Cao, "LAS-AT: adversarial training with learnable attack strategy," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 13398-13408, doi: 10.1109/CVPR52688.2022.01304.

[33] G. Michau and O. Fink, "Unsupervised transfer learning for anomaly detection: Application to complementary operating condition transfer," *Knowledge-Based Systems*, vol. 216, p. 106816, 2021, doi: 10.1016/j.knosys.2021.106816.

[34] I. Ullah and Q. H. Mahmoud, "Design and development of RNN anomaly detection model for IoT networks," *IEEE Access*, vol. 10, pp. 62722-62750, 2022, doi: 10.1109/ACCESS.2022.3176317.

[35] M. Abdel-Basset, N. Moustafa, and H. Hawash, "Privacy-Preserved Generative Network for Trustworthy Anomaly Detection in Smart Grids: A Federated Semisupervised Approach," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 995-1005, 2022, doi: 10.1109/TII.2022.3165869.

[36] S. Bird, K. Kenthapadi, E. Kiciman, and M. Mitchell, "Fairness-aware machine learning: Practical challenges and lessons learned," in *Proceedings of the twelfth ACM international conference on web search and data mining*, 2019, pp. 834-835, doi: 10.1145/3289600.3291383.

[37] M. Krishnan, "Against interpretability: a critical examination of the interpretability problem in machine learning," *Philosophy & Technology*, vol. 33, no. 3, pp. 487-502, 2020, doi: 10.1007/s13347-019-00372-9.

[38] D. M. Parker, S. G. Pine, and Z. W. Ernst, "Privacy and informed consent for research in the age of big data," *Penn St. L. Rev.*, vol. 123, p. 703, 2018.

[39] M. Samadzadeh and N. F. Ghohroud, "Evaluating Security Anomalies by Classifying Traffic Using Deep Learning," in *2023 9th International Conference on Web Research (ICWR)*, 2023, pp. 135-141, doi: 10.1109/ICWR57742.2023.10138963.

[40] M. Samadzadeh and N. F. Ghohroud, "Evaluating Security Anomalies by Classifying Traffic Using a Multi-Layered Model," *International Journal of Web Research*, vol. 6, no. 1, pp. 17-28, 2023, doi: 10.22133/ijwr.2023.396115.1151.

[41] M. Samadzadeh and N. F. Ghohroud, "Detection of Denial-of-Service Attacks in Software-Defined Networking Based on Traffic Classification using Deep learning," in *2024 10th International Conference on Artificial Intelligence and Robotics (QICAR)*, 2024, doi: 10.1109/AISP61396.2024.10475278.

[42] M. Samadzadeh, M. H. Zahedi, and E. Farahani, "Using Ensemble Learning, A Cosine Similarity-Based Model for Detecting Security Anomalies in Software-Defined Networks," in *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*, 2024, pp. 1-6, doi: 10.1109/AISP61396.2024.10475278.