International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 12, No. 4, 2025, pp. 152-156.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Deep Learning-Based Anomaly Detection in Cyber-Physical Systems

Moti Ranjan Tandi [a], Archana Mishra [b]

[a] *Assistant Professor, Department of Computer Science, Kalinga University, Naya Raipur, Chhattisgarh, India. E-mail: ku.motiranjantandi@kalingauniversity.ac.in, Orcid: https://orcid.org/0009-0001-0346-1203*
[b] *Assistant Professor, Department of Computer Science, Kalinga University, Naya Raipur, Chhattisgarh, India. E-mail: ku.archanamishra@kalingauniversity.ac.in, Orcid: https://orcid.org/0009-0001-3989-616*

Abstract

Although anomaly identification is essential for identifying infrequent occurrences, these methods frequently fall short in differentiating anomaly from noise. While they are related, anomaly detection is different from noise removal. While anomaly detection is a component of data analysis, noises are undesired objects that impede the process. Although noise can be represented as a weak type of anomaly, this does not imply that all noise is abnormal. The elimination and identification of noise is crucial for anomaly detection, even though it is true that the distinction between the two is not exact. To tackle the challenge of anomaly detection, researchers have come up with various machine learning and data-mining algorithms. However, these methods come with their own set of issues, like forcing data points into specific clusters, class dominance, local convergence, and being overly sensitive to how clusters are chosen. Additionally, it's been observed that the same technique can produce different results when used on different datasets. Identifying abnormalities in cloud servers is vital for maintaining data confidentiality, especially with the rise of massive data storage and access in cloud computing. The security systems developed in the early stages of this study are more effective than previous access control methods. However, those earlier methods assumed that cloud servers could be trusted. In reality, the server in this system might be untrustworthy, as it often manipulates encrypted data and can be caught doing so.

Keywords: Anomaly Detection, Artificial Intelligence, Sensor, Deep Learning.

DOI:10.71086/IAJSE/V12I4/IAJSE1249
Received: 29 September 2025/Revised: 04 November 2025/Accepted: 05 December 2025/Published: 30 December 2025

Introduction

The goal of anomaly detection, a subset of data mining, is to identify unusual occurrences or patterns in data that substantially deviate from their expected behaviour. In various application sectors of literature, these uncommon patterns are referred to by various names, including anomalies, abnormalities, outliers, irregularities, deviations, exceptions, or oddities (Han et al., 2021). One of the most obvious yet well-known issues in data mining, which has several uses, is anomaly detection. Because abnormalities in a computer or network might disrupt the functioning of a data-overloaded modern society, it is relevant in the current Internet era. Since the Internet serves as a conduit for millions of networked computers and a source of access to all kinds of information, its security is a major problem. Any damage to a network like this could really affect its availability, confidentiality, or integrity. But not every anomaly is a bad thing or a sign that something's gone wrong. In fact, these unusual patterns can be quite significant, even though they can be tough to investigate (Fernández Maimó et al., 2019). For instance, unusual ECG artifacts or fatal heart rates may signal the risk of a heart attack, unusual credit card transactions may indicate credit card fraud, changes in consumer purchasing patterns may indicate the discovery of a recent trend, or unusual stock fluctuations may indicate stock market instability (Munir et al., 2019).

Methodology

Design Frame Work

This method's primary objective is to efficiently detect abnormalities in real-time based on usage behavior; that is, the system calls identifier sequences generated by cloud servers where big data representations of these behaviours are linked. In order to summarize the system, call identifier sequence of any length, effective data summarization is first built using several feature states. The streaming of raw system call sequences is then trained and tested using an anomaly detecting algorithm (Alloqmani et al., 2021).

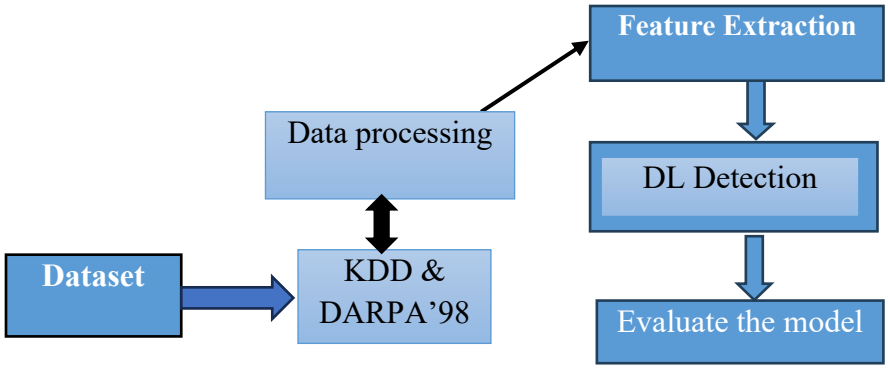


Fig. 1: System Flow Diagram

Log files play a crucial role in gathering observable system calls and tasks executed by cloud servers which is shown in Fig. 1. Whether the sequences are typical or unusual, the variations in this system call identifiers within a certain length tend to be random. By employing this method, we can monitor the sequences of system call identifiers for both normal and problematic services. To identify anomalies in cloud servers, we utilize Convolutional Neural Networks (CNNs), a deep learning technique that helps us learn and differentiate between normal and abnormal behavioural patterns (Duong et al., 2023).

Deep Learning Algorithm

Neural networks, which include many hidden layers (at least two), are central to the machine learning subfield of deep learning. One of the most challenging issues in many real-world applications is identifying the relevant high-level, abstract properties in raw data. In order to overcome this difficulty, deep learning uses representations that are definable in terms of simpler ones. An alternative perspective on deep learning is that it allows computers to learn intricate, multi-step routines. After carrying out a distinct set of instructions concurrently, each layer of representation functions similarly to the computer's memory state. This implies that more sequential tasks can be handled by deeper networks. Self-organizing maps (SOMs), autoencoders, and recurrent neural networks (RNNs) are some of the most popular unsupervised deep learning techniques in the literature. The autoencoder's encoder function converts the input data into a new representation, while the decoder function restores the original format of the new representation. Since anomalies are rarely found in the training data, the model essentially learns to reflect regular data behavior (Kwon et

al., 2019). This architecture is super user-friendly for spotting anomalies. You can use the reconstruction error to pinpoint anomalies during the testing phase. RNNs, or Recurrent Neural Networks, are another popular unsupervised deep learning method found in many of the studies we've looked at. They really shine when it comes to time series data, like performance metrics, because they're specifically designed to handle sequential information. However, RNNs do face challenges with issues like gradient vanishing or exploding. Long Short-Term Memory (LSTM) networks were presented as a solution to the gradient vanishing problem, resolving some of RNNs' drawbacks. In the context of Anomaly Detection Systems (ADS), most RNN-based models learn the sequential patterns in predict a probability distribution for the next set of values. An anomaly is flagged if the actual next value turns out to be an unlikely event based on the learned probability distribution. The Life-Log data anomaly detection system also utilizes RNN-LSTM units. To effectively detect anomalies in health log data collected from numerous devices, a model called Life-Log Anomaly Detection (LLAD) is developed in the cloud using an RNN-LSTM architecture (Gómez et al., 2023). The same data is examined using a multivariate RNN-LSTM technique to search for anomalies because RNN-LSTM performs exceptionally well with multivariate sequential data.

Results and Discussion

Verifying the accuracy and dependability of the suggested deep learning model is the main goal of this section. For this study, we selected the benchmark datasets from KDD CUP 1999 and DARPA 1998. Since both datasets exhibit normal and abnormal patterns of behavior relevant to network traffic, we use each dataset independently to train and evaluate the efficacy of the proposed anomaly detection method.

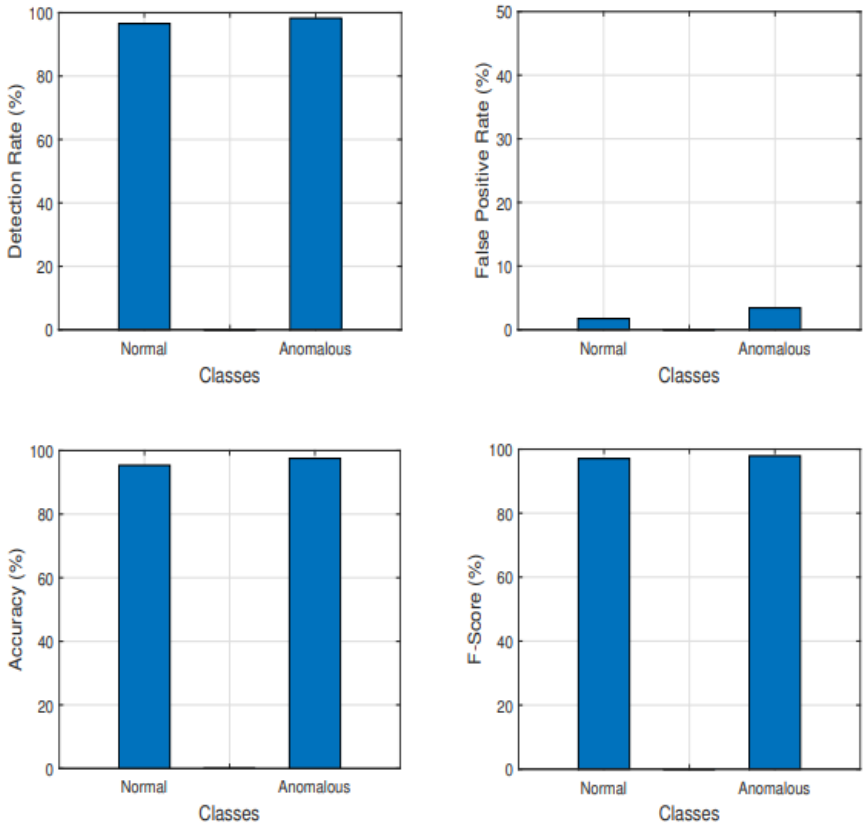


Fig. 2: Evaluation of Proposed Technique on DARPA'98 Dataset

Our trials employ 70% of the dataset for training, and the remaining part is reserved for testing which is shown in Fig. 2. It is challenging to discern between malicious and authentic packets in the DARPA dataset, which is made up of tcp dump data from packets leaving the network.

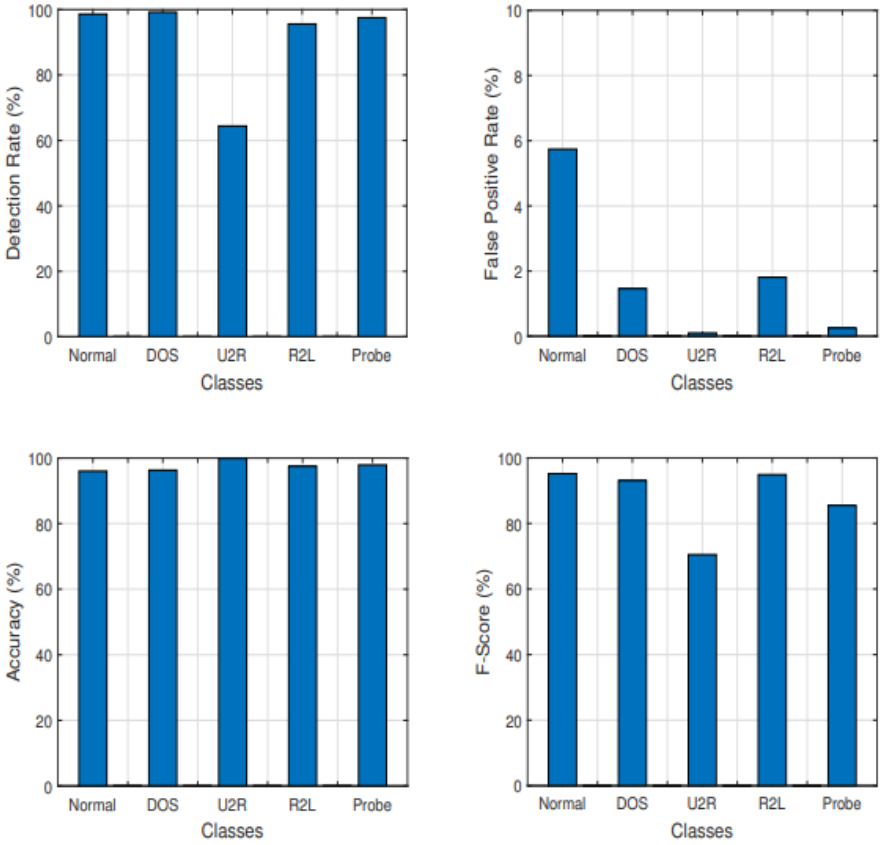


Fig. 3: Evaluation of Proposed Technique on KDD'99 Dataset

Consequently, each assault discovered in the DARPA'98 dataset is allocated to a single class, treating the anomaly detection problem as a binary-class classification problem. However, the variety of assault classes in the KDD'99 dataset makes anomaly identification a multi-class classification problem (Luo et al., 2021). For each dataset class, we calculate performance metrics such as accuracy, detection rate, false positive rate, and F-Score which is shown in Fig. 3.

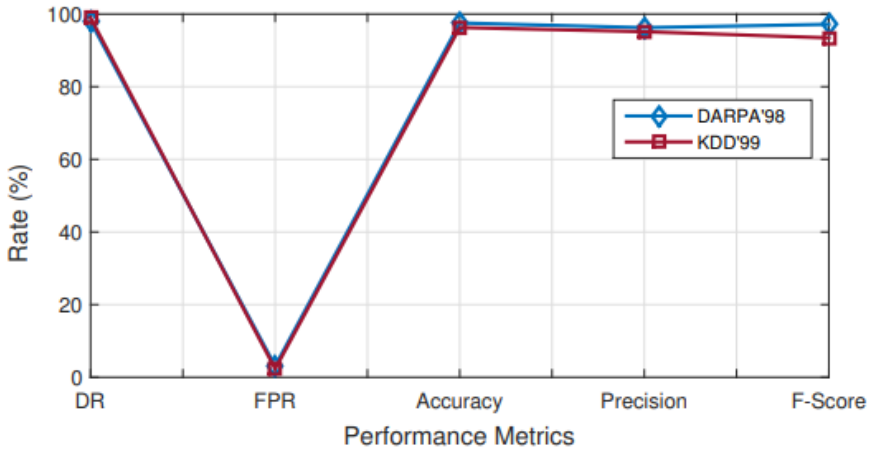


Fig. 4: Performance Evaluation of DL on DARPA'98 and KDD'99 Dataset

The overall performance of the proposed technique on the DARPA'98 and KDD'99 datasets, evaluated according to multiple criteria, is shown in Fig 4. The method achieves a high detection rate (DR) of more than 96% for both datasets while maintaining a low false positive rate (FPR).

Conclusions

The significant degree of data unpredictability frequently makes the task of detecting abnormalities more difficult. It is highly valuable IT security, finance, health care, safety critical systems, and vehicle tracking, among all other fields where sensors generate massive amounts of data in an exponential fashion. Anomaly detection approaches take into account that there are much more normal examples in the dataset than abnormal instances for all such applications. In the case of computer worm detection, for instance, the worms (anomalies) are far more frequent than typical occurrences, even though anomalies are not always uncommon. Therefore, anomaly detection can be thought of as using known information to decipher the unknown. In essence, it is a process of discovery that creates a baseline behavioral model for the data; any departures from this model are regarded as abnormal. It is evident from recent developments in this field that combining many anomaly detection strategies would yield superior outcomes compared to using them alone. But as information technology has advanced quickly, security risks in computer networks have increased.

References

Han, D., Wang, Z., Chen, W., Zhong, Y., Wang, S., Zhang, H., ... & Yin, X. (2021, November). Deepaid: Interpreting and improving deep learning-based anomaly detection in security applications. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security* (pp. 3197-3217). https://doi.org/10.1145/3460120.3484589

Fernández Maimó, L., Huertas Celdrán, A., Gil Pérez, M., García Clemente, F. J., & Martínez Pérez, G. (2019). Dynamic management of a deep learning-based anomaly detection system for 5G networks. *Journal of Ambient Intelligence and Humanized Computing*, *10*(8), 3083-3097. https://doi.org/10.1007/s12652-018-0813-4

Munir, M., Chattha, M. A., Dengel, A., & Ahmed, S. (2019, December). A comparative analysis of traditional and deep learning-based anomaly detection methods for streaming data. In *2019 18th IEEE international conference on machine learning and applications (ICMLA)* (pp. 561-566). IEEE.

Alloqmani, A., Abushark, Y. B., Khan, A. I., & Alsolami, F. (2021). Deep learning based anomaly detection in images: insights, challenges and recommendations. *International Journal of Advanced Computer Science and Applications*, *12*(4).

Duong, H. T., Le, V. T., & Hoang, V. T. (2023). Deep learning-based anomaly detection in video surveillance: A survey. *Sensors*, *23*(11), 5024. https://doi.org/10.3390/s23115024

Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., & Kim, K. J. (2019). A survey of deep learning-based network anomaly detection. *Cluster Computing*, *22*(Suppl 1), 949-961. https://doi.org/10.1007/s10586-017-1117-8

Gómez, Á. L. P., Maimó, L. F., Celdrán, A. H., & Clemente, F. J. G. (2023). SUSAN: A Deep Learning based anomaly detection framework for sustainable industry. *Sustainable Computing: Informatics and Systems*, *37*, 100842. https://doi.org/10.1016/j.suscom.2022.100842

Luo, Y., Xiao, Y., Cheng, L., Peng, G., & Yao, D. (2021). Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities. *ACM Computing Surveys (CSUR)*, *54*(5), 1-36. https://doi.org/10.1145/3453155