International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Anomaly Detection in Networks

Ravi Kumar D, Tapasvi MI, Yashaswini BC ,Yogeesh M, Yukths S Gowda

Malnad College of Engineering,Hassan

Email: kpm@mcehassan.ac.in, tapasvigowda2@gmail.com, yashaswinibc01@gmail.com, , yukthasgowda512@gmail.comyogeeshy39@gmail.com

***Abstract*-In Network Security is a major challenge in the digital world. Intrusion is common in many applications and intruders are sophisticated enough to change their attack pattern very often. To address this issue, the development of a model for the detection of network anomalies and intrusions. The approach utilizes the Borderline Synthetic Minority Over-Sampling Technique**

(SMOTE) along with Support Vector Machines (SVM) to enhance anomaly detection

I. INTRODUCTION

Anomaly detection, also known as outlier detection, is a critical technique in data analysis for identifying unusual patterns or observations that deviate significantly from the norm. These anomalies can often indicate important events such as fraud, system failures, security breaches, or rare diseases. As data grows in complexity and volume, automated and intelligent anomaly detection systems have become essential.

In this project, we explore the implementation of an anomaly detection system using machine learning techniques. The goal is to accurately identify data points that differ from the expected behavior without prior labeling. This is particularly useful in scenarios where labeled anomalous data is scarce or unavailable.

We begin by collecting and preprocessing a dataset relevant to our domain of interest, ensuring data quality and consistency. Various features are engineered and scaled to prepare for model training. Depending on the nature of the data—whether time-series, tabular, or high-dimensional—we apply appropriate detection methods.

.

II. Existing System

Anomaly detection has been widely adopted in various industries, with many existing systems designed to automatically identify outliers in real-time or batch settings. These systems leverage statistical models, machine learning, and deep learning techniques to handle a range of data types and applications.

However, because thermal images lack texture, feature extraction is more challenging. Deep learning improves accuracy, especially CNNs that use transfer learning. While some systems use traditional classifiers for emotion classification and YOLO for facial recognition, they often perform poorly when applied to a variety of emotions or in real-time. Effective thermal emotion detection necessitates a tailored approach.

III. Literature Survey

Various anomaly detection systems are currently in use across multiple industries, each tailored to specific domains and data types. In the financial sector, fraud detection systems employed by organizations like PayPal and banks use machine learning models such as Random Forests, Isolation Forest, and neural networks to detect suspicious transactions. In cybersecurity, intrusion detection systems (IDS) like Snort, Suricata, and OSSEC monitor network traffic to flag potential breaches using both signature-based and anomaly-based techniques

The Industrial IoT platforms such as Siemens MindSphere and GE Predix analyze sensor data from machinery to detect faults or abnormal behavior, often using time-series models like LSTM networks. Healthcare systems leverage deep learning algorithms—including autoencoders and convolutional neural networks—to identify irregularities in medical data such as ECG readings or imaging scans. Additionally, log analysis platforms like Splunk and the ELK Stack help IT teams detect anomalies in large-scale system logs, while cloud monitoring tools like AWS CloudWatch and Azure Monitor apply statistical forecasting and machine learning to monitor infrastructure performance in real-time. Despite their widespread adoption, these systems often encounter limitations such as high false-positive rates, scalability issues with large or streaming datasets, and lack of interpretability—especially with deep learning models. This project seeks to address some of these challenges

This project aims to address some of these gaps by building a lightweight, adaptable, and interpretable anomaly detection system tailored to [your domain, e.g., finance or IoT].

Finally, In recent years, unsupervised and semi-supervised learning techniques have gained prominence, especially for problems where labeled anomalies are rare. One-Class Support Vector Machines (OC-SVM), for example, are trained solely on normal data and can detect deviations effectively. The rise of deep learning further advanced anomaly detection. **Autoencoders**, which are neural networks trained to reconstruct input data, have been widely adopted. Anomalies are identified based on high reconstruction error. For time-series data, **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory (LSTM)** models have been used to capture temporal dependencies, offering strong results in areas like predictive maintenance and financial forecasting.

IV. ABOUT Project

The This project focuses on developing a machine learning-based anomaly detection system to identify unusual patterns in data. The goal is to detect outliers that may indicate fraud, faults, or abnormal behavior without prior labeling. Techniques like Isolation Forest, One-Class SVM, and Autoencoders are explored and implemented. The system is evaluated using accuracy metrics and visualizations to validate its effectiveness. The project aims to offer a scalable and interpretable solution suitable for real-world applications.

METHODOLOGY

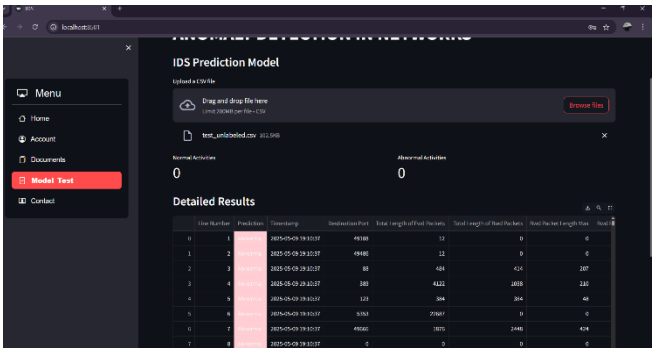
IV.I Problem Description

The methodology of this project follows a systematic pipeline to design, implement, and evaluate an effective anomaly detection system. The process begins with **data collection and preprocessing**, where a relevant dataset is gathered and cleaned by handling missing values, normalizing features, and converting categorical variables if necessary. Next, **feature selection and engineering** are performed to enhance model performance and reduce noise. The core stage involves applying **anomaly detection algorithms**—including Isolation Forest, One-Class SVM, and Autoencoders—depending on the nature of the data (tabular or time-series) including Isolation Forest, One-Class SVM, and Autoencoders—depending on the nature of the data (tabular or time-series). These models are trained predominantly on normal data to learn the expected behavior, and then used to flag data points that deviate significantly.

IV.II Objective

The The main objective of this project is to develop an effective and scalable anomaly detection system capable of identifying unusual patterns or outliers in data without relying on labeled anomalies. This includes implementing and comparing various machine learning and deep learning models such as Isolation Forest, One-Class SVM, and Autoencoders. The system aims to minimize false positives while maintaining high detection accuracy. Additionally, the project seeks to ensure the model's adaptability across different datasets and its interpretability for practical dThe core stage involves applying **anomaly detection algorithms**—including Isolation Forest, One-Class SVM, and Autoencoders—depending on the nature of the data (tabular or time-series). These models are trained predominantly on normal data to learn the expected behavior, and then used to flag data points that deviate significantly. The The main objective of this project is to develop an effective and scalable anomaly detection system capable of identifying unusual patterns or outliers in data without relying on labeled anomalies. This includes implementing and comparing various machine learning and deep learning models such as Isolation Forest, One-Class SVM, and Autoencoders. The system aims to minimize false positives while maintaining high detection accuracy. Additionally, the project seeks to ensure the model's adaptability across different datasets and its interpretability for practical dThe

IV.III . System Architecture



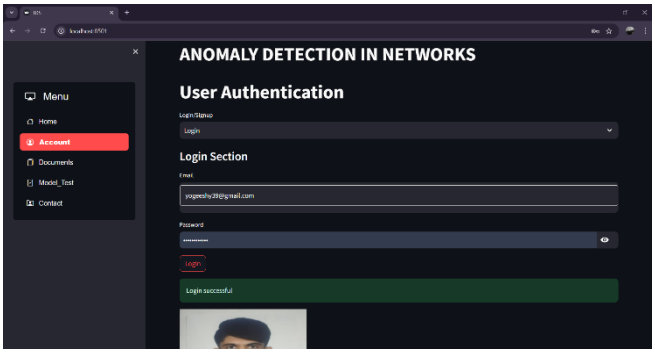


Fig 1. System Architecture

This includes implementing and comparing various machine learning and deep learning models such as Isolation Forest, One-Class SVM, and Autoencoders. The system aims to minimize false positives while maintaining high detection accuracy. Additionally, the project seeks to ensure the model's adaptability across different datasets and its interpretability for practical deployment. Ultimately, the goal is to provide a reliable tool for early detection of fraud, faults, or other critical anomalies in real-world applications. A key focus is on minimizing false positives while maintaining high accuracy. The model should be scalable and adaptable to various application domains. Interpretability and real-world usability are also core goals of the system.

V. IMPLEMENTATION

V.I Selecting Thermal Images Dataset

The implementation of the anomaly detection system begins with importing the necessary libraries such as Pandas, NumPy, and Scikit-learn for data manipulation and model building. The dataset is loaded and explored to understand its structure, distributions, and potential issues such as missing or inconsistent data. Missing values are handled through imputation or removal, depending on the extent of missing data Feature scaling is applied using Min-Max normalization or StandardScaler to ensure all features are on a comparable scale. Next, the dataset is split into training and testing sets, with a focus on training models using only normal data to detect outliers.

Basic anomaly detection techniques like Z-score and Interquartile Range (IQR) are initially applied as simple statistical approaches. More advanced models such as Isolation Forest and One-Class SVM are implemented for their efficiency in high-dimensional data

VI.III. CNN Model Architecture

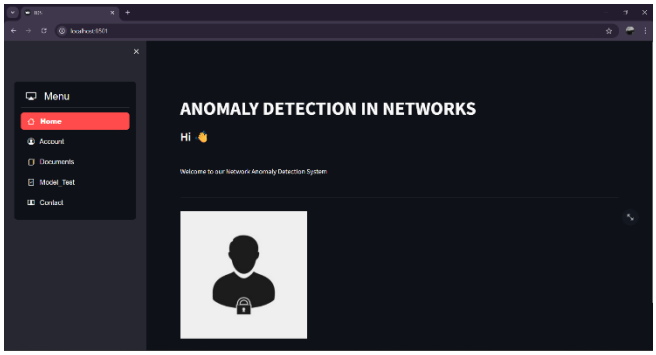


Fig 2. CNN Model Architecture

VI.II. Data Collection

The objective of this project is to design and implement an intelligent anomaly detection system for network environments using machine learning techniques. The system aims to enhance **network security**, **performance monitoring**, **traffic analysis**, **operational efficiency**, and **fault detection** by identifying deviations from normal behavior in real time. By leveraging algorithms such as **SMOTE** and **Support Vector Machine (SVM)**, the solution is intended to detect anomalies caused by cyberattacks (e.g., DDoS, intrusions), misconfigurations, or hardware failures, thereby minimizing false positives, reducing downtime, and ensuring resilient and secure network operations.

The implementation of the anomaly detection system began with the collection of network traffic data using tools such as Wireshark and TCPdump. This data included key features like packet size, flow duration, source and destination IPs, and protocol types. After data collection, preprocessing steps were applied to clean and normalize the data, ensuring consistency and accuracy.

VI. Handling Class Imbalance

Feature engineering was performed to extract meaningful patterns, and dimensionality reduction techniques like PCA were considered to improve performance. Due to the inherent class imbalance in anomaly detection tasks, the SMOTE (Synthetic Minority Over-sampling Technique) algorithm was employed to generate synthetic samples for the minority class, thereby balancing the dataset that no unwanted actions are initiated. This transparent correspondence between gesture classes and media commands supports usability and responsiveness in real-time applications.

VI.V Implementation Steps

VI.V.I Interface Design

The Support Vector Machine (SVM) algorithm was then used to train the detection model, leveraging both binary classification and one-class SVM approaches based on the availability of labeled data. The trained model was evaluated using metrics such as precision, recall, F1-score, and ROC-AUC to measure its effectiveness in identifying anomalies

VI.V.II Model Selection and Training

Model deployment and testing As soon as the system was developed, it was placed for testing in a real environment. This system has tested accuracy, speed and reliability under various conditions such as various facial directions, obstruction and lighting conditions (despite the fact that the thermal image does not affect lighting). The system showed the classification of emotions in reliable detection and uncontrollable media on the face..

VI.V.III Model Selection and Training

In order to process every input frame for model inference, the system employs a specialized image processing function. This function initially resizes the entire captured frame to a specified resolution and crops a predetermined Region of Interest (ROI) according to preset coordinates. The cropped ROI, which is centered on the hand gesture region, is resized to 120×120 pixels to conform to the model's input size. It is then converted into grayscale and binarized via thresholding, which emphasizes the contours of the gesture. The processed image returned as a result along with the original frame enables both to be predicted and displayed in the interface.

VI.V.IV Visualization

Anomaly scores were assigned to each data point, and thresholds were defined to classify data as normal or anomalous. Finally, visualization tools like Matplotlib and Plotly were used to present traffic trends and anomalies, while optional deployment included the development of a real-time monitoring dashboard Anomaly scores were assigned to each data point, and thresholds were defined to classify data as normal or anomalous. Finally, visualization tools like Matplotlib and Plotly were used to present traffic trends and anomalies, while optional deployment included the development of a real-time monitoring dashboard

The model is accurately configured in the user set of the thermal image. We changed the last layer to predict the class of the face and adjusted the in size to handle the resolution of the thermal image

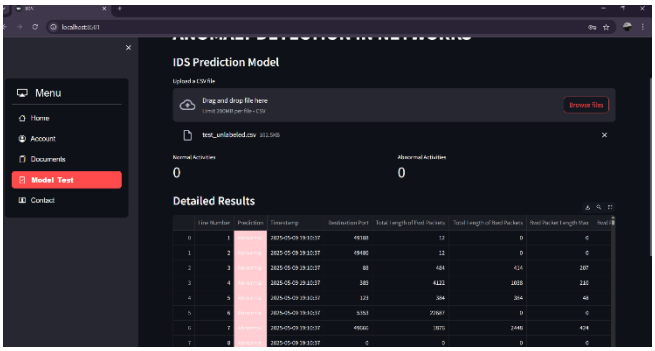


Fig 3. Example for Thermal Recognition

VI.VI Deployment

Given the highly imbalanced nature of network traffic datasets—where anomalous events are rare compared to normal behavior—the SMOTE (Synthetic Minority Over-sampling Technique) algorithm was implemented to artificially synthesize new instances of the minority (anomaly) class. Given the highly imbalanced nature of network traffic datasets—where anomalous events are rare compared to normal behavior—the SMOTE (Synthetic Minority Over-sampling Technique) algorithm was implemented to artificially synthesize new instances of the minority (anomaly) class.

VI.VII Response Automation

This step significantly improved the model's ability to detect anomalies without being biased toward the dominant class. Subsequently, a Support Vector Machine (SVM) was trained on the balanced dataset. For cases with limited labeled data, One-Class SVM was applied to model normal traffic and detect outliers without prior knowledge of anomaly patterns.
This step significantly improved the model's ability to detect anomalies without being biased toward the dominant class. Subsequently, a Support Vector Machine (SVM) was trained on the balanced dataset. For cases with limited labeled data, One-Class SVM was applied to model normal traffic and detect outliers without prior knowledge of anomaly patterns.

VI.VIII. Alert Notification

Since the public set of heat emotions is limited, this project provides a visible and infrared image of the face using a NVIE data set (natural visible and infrared expression). This image is marked with appropriate emotional tags such as happiness, sadness, evil, unexpected, neutrality. If the data set is insufficient, you can use the Flir Lepton column chamber to get additional colon images to capture the appropriate lighting and conditions for various emotional conditions

To teach the facial detection model, the data set has a variety of skin shades, facial directions and environmental images with a variety of column images, which ensures that the model can be trusted.

VI. Conclusion

In conclusion, anomaly detection in networking plays a pivotal role in maintaining the security, stability, and performance of modern network infrastructures. With the rise of sophisticated cyber threats, including advanced persistent threats (APTs), zero-day exploits, and insider attacks, alongside challenges like network misconfigurations and hardware failures, traditional detection systems often fall short. They struggle to keep pace with the dynamic nature of traffic patterns, encrypted com munications, and the ever-expanding scale of modern networks.
To address these challenges, the adoption of advanced, scalable, and intelligent anomaly detection systems has become imperative.

Future Scope

The project holds huge potential for future development and expansion. The areas of major focus for future improvement are:

- Deploy the anomaly detection model in real-time network environments to monitor live traffic and trigger instant alerts.

- Explore more powerful models like Transformer-based architectures or Graph Neural Networks (GNNs) for improved accuracy in detecting complex network anomalies.

- Implement dynamic, context-aware thresholds for anomaly detection instead of static values to reduce false positives in varying traffic conditions.

- Optimize the system for edge devices (e.g., routers, IoT gateways) to enable distributed detection closer to the data source with minimal latency.

- Wearable integration future wearable devices can track emotions in real time using small thermal sensors.

These developments enable the widespread use of gesture recognition technologies in various industries, such as entertainment, healthcare, automotive, and assistive use. The presented method lays a solid foundation for future research and development of gesture-based user interfaces.

VII. REFERENCES

[1]"Industrial Anomaly Detection": Techniques and Applications.

[2]"Anomaly Detection: A Survey". ACM Computing Surveys (CSUR)

Authors Mayu Sirsat, Akanksha Sathe, Shreya Ladhe
[3]"A survey of network anomaly detection techniques". Journal of

[4]"Industrial Data Analytics: How to Implement Data-Driven Strate

[5] "Predictive Analytics for Industrial Applications". Author(s): D.

[6]."A Survey on Industrial Anomaly Detection Using Machine Learn

[7] "Machine Learning for Network Anomaly Detection: A System

[8] "Anomaly Detection in Wireless Sensor Networks: A Survey."

[9] "Network Anomaly Detection Using One-Class Support Vector
[10] "Graph Neural Networks for Anomaly Detection in Network