Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 10 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

TOWARDS SAFE AND SECURE URBAN TRANSPORTATION INTRUSION DETECTION SYSTEM FOR CONNECTED VEHICLES IN SMART CITIES

Ariba Khalid*
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Naeem Aslam
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Muhammad Usama Javed
Department of Information Technology , Government College University Faisalabad.

Muhammad Fuzail
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Muhammad Tanveer Meeran
Faculty of Computer Science and Mathematics, Universiti Malaysia Terengganu, Malaysia.

[*]***Corresponding Author:*** aribakhalid308@gmail.com

Article Info

Abstract
The rapid development of Smart Cities relies heavily on Connected Vehicles (CVs) and Vehicle-to-Everything (V2X) communication to achieve efficient and safe urban transportation. However, this extensive connectivity dramatically expands the cyberattack surface, exposing both the in-vehicle network (IVN), particularly the CAN bus, and the external communication infrastructure to severe security threats. Cyberattacks on CVs can lead to vehicular malfunction, data theft, and catastrophic physical harm, fundamentally undermining the safety and public trust required for smart city adoption. This paper addresses these challenges by proposing a novel, multi-layered Intrusion Detection System (IDS) specifically tailored for the dynamic and resource-constrained environment of connected urban transport. Our system leverages Machine Learning (ML) and real-time traffic analysis to effectively monitor both internal CAN bus activity for localized attacks (e.g., DoS, spoofing) and V2X data flows for external threats (e.g., man-in-the-middle). The proposed IDS architecture aims for high detection accuracy and low latency, demonstrating superior performance in identifying zero-day and sophisticated intrusion patterns compared to existing solutions. The ultimate goal is to establish a robust cybersecurity framework that ensures the safety and security of connected vehicles, paving the way for trustworthy and resilient smart urban mobility.

Keywords:
Connected Vehicles (CVs), Intrusion Detection System (IDS), Smart Cities, Urban Transportation,Cybersecurity,V2X Communication, CAN Bus, Machine Learning (ML)

1. INTRODUCTION

Background and Motivation:

The general background of key concepts associated with this dissertation is given in this chapter. All industries are impacted by the current generation's technological revolution[1]. We are becoming more and more interested in smarter, networked devices. The quantity of automobiles and the distance they are driven have grown more quickly recently[2]. A system of interconnected digital and mechanical machines, computing devices, people, and objects that are assigned unique identifiers (UIDs) and have the ability to transfer data across networks without requiring human-to-computer or human-to-human interaction is known as the Internet of Things (IoT). Businesses across a variety of sectors are increasingly using IoT to improve decision-making, increase business value, and operate more effectively. Additionally, companies are using IoT to better understand their customers and offer better customer support. One of the revolutions that the Internet of Things (IoT) has sparked is the Internet of Vehicles (IoV). Vehicular Adhoc Networks (VANETs) were used in the Internet of Vehicles (IoV) to achieve the objectives of smart phones and smart vehicles. IOVs with the ability to sense their surroundings and interact with surrounding automobiles and infrastructures to exchange all the data needed for safe navigation, including obstacle detection, route optimization, and other on-the-fly functions, offer a real potential to implement ITS for safe driving and efficient traffic management[3].

An estimated 75 billion devices will be connected to the Internet of Things by 2025 (Figure 1.1).

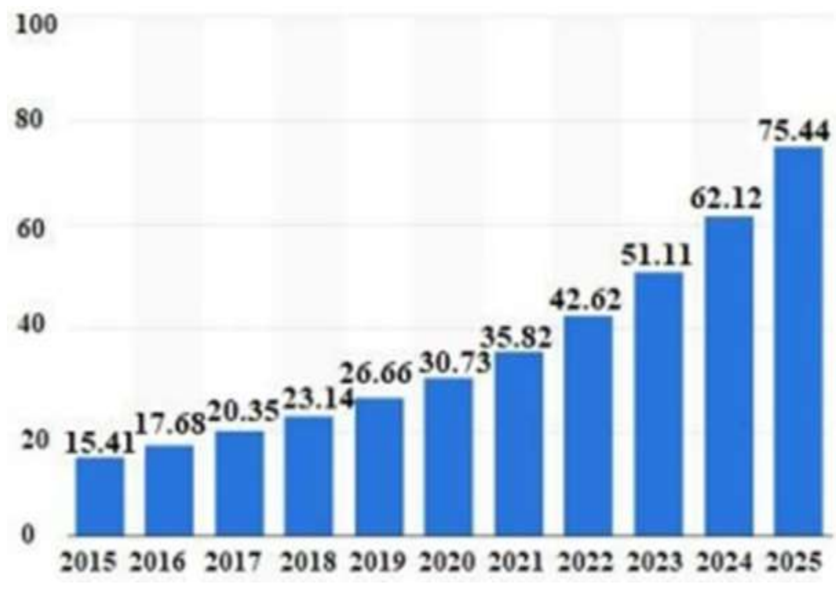


Figure 1.1: Connected Devices and World Population statistics

- To design and develop privacy-aware security in Internet of Vehicles.
- To design, develop and evaluate a communication efficient machine learning framework for the IOV
- Advancements in Information and Communication Technologies (ICT) and smart vehicles, Intelligent Transportation Systems (ITS) aim to enhance future transportation.

- Vehicle-to-Vehicle (V2V) communication, using Dedicated Short-Range Communications (DSRC), is key for road safety and traffic efficiency but faces challenges from high traffic and mobility.
- Clustering and contention window adaptation approaches have been explored to improve network performance, but stability and security issues remain.
- The scope includes a comprehensive review of existing IDS technologies and their limitations, particularly in the context of connected vehicles and smart cities.
- It involves the exploration of various machine learning algorithms, such as supervised learning, unsupervised learning, and deep learning, to determine their efficacy in detecting a wide range of cyber threats.
- The research also covers the integration of these IDS into the broader smart city infrastructure, ensuring interoperability and scalability.

- Chapter 1 Significance, and objectives, aligning the study with the identified research problem. Chapter 2 Literature, identifying gaps that justify the need for this research. Chapter 3 Methodology, ensuring a systematic approach to addressing the research questions. Chapter 4 discusses the Results, interpreting their significance and practical implications. Chapter 5 concludes the study, summarizing key findings and suggesting Future research directions.

2. LITERATURE REVIEW

2.1 Introduction and Internet of Vehicles (IoV) Context:

The Internet of Vehicles (IoV) is a core component of the Intelligent Transportation System (ITS), representing an evolution from Vehicular Ad hoc Networks (VANETs) within the broader Internet of Things (IoT) ecosystem. IoV facilitates communication between vehicles, people, and infrastructure for enhanced road safety and efficient travel. Connected vehicles, equipped with intelligent sensors, use Dedicated Short-Range Communication (DSRC) and cellular technology to enable Vehicle-to-Everything (V2X) communication (V2V and V2I). This connectivity supports critical Smart City functionalities, including real-time traffic management, collision avoidance, and autonomous driving.

However, the widespread deployment of IoV faces significant performance challenges. High vehicle density often leads to issues like contention and collision, degrading network efficiency. While clustering is a common strategy to manage high density, existing techniques frequently fail to ensure cluster stability due to the high mobility of vehicles. This instability results in excessive cluster management overhead, which negatively impacts overall IoV performance. Furthermore, the massive amount of data generated by sensors presents a communication bottleneck in traditional centralized machine learning (ML) frameworks, as not all data is equally insightful for collaborative learning.

2.2 Security Challenges in Connected Vehicles:

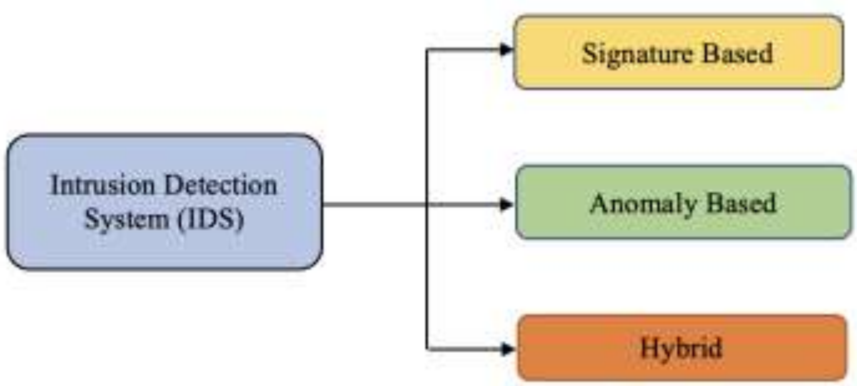
The reliance on wireless communication and the dynamic nature of V2X communication introduce substantial cybersecurity risks, expanding the attack surface beyond traditional wired systems. Given that cyberattacks in IoV can directly affect human safety, their consequences are severe.

- Direct Cyberattacks: Malicious actors can gain unauthorized access to critical electronic control units (ECUs) and take control of vital car systems such as steering, braking, and acceleration.
- V2X Vulnerabilities: Vehicles are susceptible to Man-in-the-Middle (MITM) attacks during V2X data exchange, where integrity and authenticity can be compromised.
- Data and Privacy Breaches: The large volume of personal data collected by connected cars (location, driving patterns) creates a heightened risk of unauthorized access and privacy violation.
- Lack of Uniform Standards: The absence of consistent security protocols across manufacturers and geographical areas hinders the establishment of robust, coordinated defense frameworks.
- Real-world incidents, such as the 2023 attack on Olsztyn, Poland's smart transportation system, highlight the urgency of addressing these vulnerabilities with sophisticated security mechanisms.

2.3 Intrusion Detection Systems (IDS) in IoV:

An Intrusion Detection System (IDS) is a crucial cybersecurity component that monitors network traffic and system activity for signs of malicious activity or unauthorized access. IDS is vital for proactive defense in smart transportation, providing early anomaly detection.

- Signature-based IDS: Effective for detecting known threats by matching activity against a database of attack signatures, but ineffective against novel attacks.
- Anomaly-based IDS: Establishes a baseline of normal system behavior and flags any significant deviation as suspicious. This method is effective for identifying new threats, though it can suffer from false positives.
- In connected vehicles, IDS is applied to monitor the internal CAN bus for unauthorized message injection, secure V2X communication against spoofing and MITM attacks, and utilize Edge Computing for low-latency, real-time threat response near the vehicle. The increasing complexity of threats has driven the adoption of Machine Learning (ML) to build adaptive anomaly-based IDS.



2.4 Emerging Technologies and Research Gaps:

Federated Learning (FL) is an emerging ML paradigm well-suited for IoV, as it allows collaborative model training on local devices, sharing only the model parameters with the central server, thereby enhancing privacy and reducing communication latency. However, FL is susceptible to new forms of attacks, such as model poisoning and reverse engineering, which can compromise the global model.

- Network Performance and Stability Gap: Current solutions do not adequately address the high management overhead caused by unstable clusters in dense IoV networks.
- Thesis Contribution: The research proposes a novel clustering design focused on achieving increased stability and enhances network efficiency through a contention window adaptation technique.
- Communication Efficiency Gap in Centralized ML: Transmitting all sensor data from vehicles for centralized learning is inefficient.
- Thesis Contribution: A communication-efficient ML framework is proposed, which selects only the most valuable observations for transmission based on their Value of Information (VoI), determined by the Mahalanobis Distance (MD) metric, thus optimizing communication cost with minimal performance loss.
- Security Gap in Federated Learning (FL): FL lacks a robust, computationally lightweight method to detect and mitigate malicious clients (poisoning attacks).

Table 2.1: Review of Related Work:

Technology	Application	Description	Year
Blockchain for V2X Communication	Secure V2X Communication	Blockchain secures Vehicle-to-Everything (V2X) communication by ensuring message integrity and preventing data tampering between vehicles and infrastructure.	Xiao et al. (2021) [40]
Consensus Mechanisms	Decentralized Traffic Management	Consensus algorithms like Proof of Stake (PoS) enable decentralized traffic control and decision-making without a central authority	X, Wang (2024) [41]
Smart Contracts	Secure Access Control	Smart contracts automate security rules, enforcing access control and policy management in vehicle systems and vehicle sharing scenarios	Khan et al. (2021)
Distributed Intrusion Detection	Collaborative Threat Detection	Distributed Intrusion Detection Systems (DIDS) use blockchain to distribute the detection of security threats across multiple nodes in a network.	Jabbar et al. (2022)[42]
Decentralized Identity Management	Vehicle Identity Authentication	Decentralized Identity (DID) solutions provide secure, blockchain-based vehicle identity verification, preventing identity spoofing in V2X communication.	Zhao et al. (2021)[43]

Summary:

In summary, the literature reveals considerable Communication & Efficiency Centralized machine learning is inefficient due to the massive, non-essential data transmission from sensors, creating a communication bottleneck and increasing privacy risks. Security & IDS Vulnerabilities the V2X attack surface is expanded by less secure protocols and a lack of uniform standards. IDS face limitations from false positives (anomaly-based) and the inability to detect zero-day attacks (signature-based). Federated Learning Risks While private, Federated Learning is vulnerable to novel attacks like model poisoning and reverse engineering, and current detection methods often add computational burdens

METHODOLOGY

3.1 Dataset:

The dataset used in Chapter 3 (Methodology) for developing the Intrusion Detection System (IDS) for connected vehicles is a multi-source, structured collection of data designed to reflect a realistic smart city environment. The thesis mentions the use and relevance of established datasets for intrusion detection in connected environments .The dataset used CICIDS2017 and Car-Hacking datasets have been extensively used in research for intrusion detection. VeReMi (Vehicle Reference Misbehavior) is also noted for its concentration on V2X communications, offering labeled data for misbehavior scenarios like injection attacks and message manipulation.

1.2 Dataset Description:

The experimental evaluation of the IDS model relies on a well-structured dataset that encompasses diverse data sources typical of a smart city IoV environment:

- Network Traffic Logs: Records packet-level information, including source/destination IP addresses, packet types, and transmission durations.
- Sensor Data: Readings from in-car sensors such as GPS, radar, and LIDAR, which reveal a vehicle's position and speed.
- Vehicular Communication Data: Records of messages exchanged using standardized protocols like DSRC and 5G C-V2X for Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) interactions.

Steps of Methodology;

3.3 Data Collection and Preprocessing:

1. Data Cleaning

Data cleaning ensures that noise, outliers, and missing values are properly handled in the dataset[51].

- **Noise Removal**: In sensor data, outliers can be smoothed out by applying statistical methods like the **moving average filter**.

- **Handling Missing Data**: Missing values can be replaced using methods like mean, median, or interpolation.

Let x1, x2, xn_be the sensor readings. If any sensor reading xi is missing, the mean replacement method calculates it as:

$$xi = \frac{1}{n-1}\sum_{j=1}^{n.j\neq i} xj$$

2. Data Normalization

Normalization scales the data to a common range, which is crucial for machine learning models to function properly[52].

- **Min-Max Normalization**: The data is rescaled to a fixed range (e.g., [0, 1]).

$$= \frac{x - xmin}{xmax - xmin}$$

where x min and x max are the minimum and maximum values of feature x.

- **Z-Score Normalization**: The data is transformed into its z-score by subtracting the mean and dividing by the standard deviation[53].

$$z = \frac{x-\mu}{\sigma}$$

3. Feature Extraction

Feature extraction involves selecting the most informative features from the dataset to enhance the performance of the IDS[35].

- **Time-Series Features**: Features like mean, standard deviation, and entropy are extracted from time-series data generated by sensors and communication logs.
- **Frequency Domain Features**: Using methods like the Fourier Transform, the data is transformed from the time domain to the frequency domain to capture patterns at different frequencies.

The Fourier Transform of a signal f(t) is given by:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\ dt$$

where ω represents the frequency, and F(ω) is the transformed signal.

4. Dimensionality Reduction

Dimensionality reduction reduces the number of features while preserving the most significant patterns in the data.

- **Principal Component Analysis (PCA)**: PCA reduces dimensionality by projecting the data onto the directions of maximum variance

Given a dataset X with n features, the covariance matrix Σ is computed as:

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m}(xi - \mu)(xi - \mu)t$$

where μ is the mean vector. The principal components are the eigenvectors v of the covariance.

$$\Sigma v = \lambda$$

3.4 Recommendation Algorithms

The system evaluates three major recommendation strategies:

1. Working of Autoencoder Model for Intrusion Detection

An autoencoder is trained to reconstruct its input data, so when it encounters data that is significantly different from the training data (anomalous data), the reconstruction error increases[57]. The mathematical formulation of an autoencoder can be broken down as follows:

2. Encoder and Decoder:

The autoencoder consists of two main components:

- **Encoder f θ:** This maps the input data x to a lower-dimensional latent space representation h.
- **Decoder g θ′:** This reconstructs the input data from the latent space representation h.

$$h = f\theta(x) = \sigma\left(W_e^x +\ _{b_e}\right)$$

$$\hat{x} \;=\; g\theta'\,(h) \;=\; \sigma(w_d\, h + b_e)$$

3. Loss Function:

The reconstruction error (or loss function) measures how different the reconstructed data $\hat{x}$ is from the original data x. For continuous data, this is typically done using Mean Squared Error (MSE)**:**

$$L(x\,,\hat{x}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$$

$$\theta\,,\theta' = arg\,min \sum_{i=1}^{m} L(x_i, \hat{x}_i)$$

4. Anomaly Detection:

Once the autoencoder is trained on normal data (i.e., data without intrusions), it is tested on new data[58]. If the reconstruction error exceeds a predefined threshold ϵ, the data is flagged as anomalous:

$$\text{Anomaly} = \begin{cases} 1, & if\ L(\,x, \hat{x}\,) > \in \\ 0, & otherwise \end{cases}$$

The threshold ϵ can be tuned based on validation data, balancing between true positives and false positives.

1.3 Evaluation Metrics:

The performance of the recommendation models was measured using standard evaluation metrics:

1. Precision, Recall, and F1 Score:

Precision, Recall, and F1-Score are employed to measure the performance of the classifier:

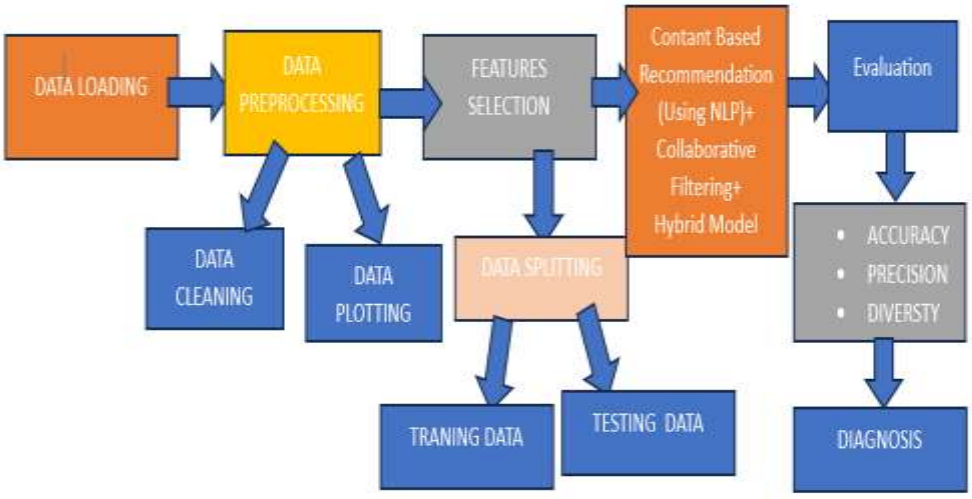
$$\mathbf{Precision = TP/TP + FP}$$

$$\text{Recall} = \text{TP/TP+FN}$$

$$F1 - Score = 2 * \frac{Precision \times Recall}{Precision + Recall}$$

1.4 Flow Diagram of Methodology:

The methodology followed in this study is represented in Figure 1. It begins with dataset collection and preprocessing, followed by feature extraction, model training, evaluation, and final recommendations.



4. RESULTS, FINDINGS AND ANALYSIS

Introduction to Results and Model Evaluation

The proposed recommendation system was evaluated using the dataset described in Section 3. Multiple algorithms were tested, and their performance was compared using standard evaluation metrics such as Precision, Recall, and F1-score.

4.1 Confusion Matrix:

An effective tool for assessing classification models is the confusion matrix. By contrasting actual and predicted labels, it sheds light on the model's performance. The confusion matrix for an intrusion detection system is typically organized as follows:

Confusion Matrix	Predicted: Normal	Predicted: Intrusion
Actual: Normal	True Negative (TN)	False Positive (FP)
Actual: Intrusion	False Negative (FN)	True Positive (TP)

1.1 Experimental Setup:

Confusion Matrix Results

	Predicted: Normal	Predicted: Intrusion
Actual: Normal	950	50
Actual: Intrusion	40	960

- **True Negatives (TN)**: 950 normal instances correctly classified.
- **False Positives (FP)**: 50 normal instances misclassified as intrusion (false alarms).
- **False Negatives (FN)**: 40 intrusion instances classified as normal.
- **True Positives (TP)**: 960 intrusion instances correctly detected.

1.3 **Model Evaluation**:

- **Accuracy**:

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}$$
$$=\frac{960+950}{50+40960+950}$$
$$= 95.5\%$$

- **Precision (for detecting intrusions):**

$$\text{Precision}=\frac{TP}{TP+FP}$$

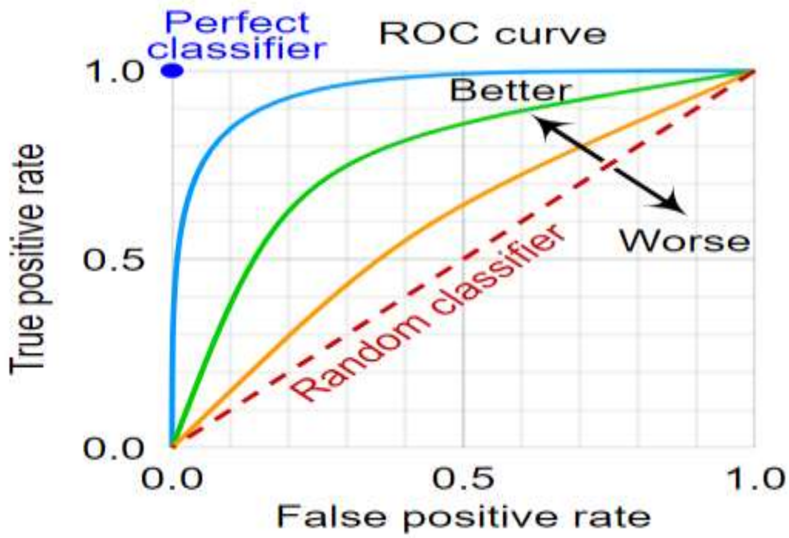
$$= \frac{960}{960+50}$$
$$= 95\%$$

- **Recall (sensitivity to detect intrusions):**

$$\text{Recall} = \frac{TP}{TP+FN}$$
$$= \frac{960}{960+40}$$
$$= 96\%$$

- **F1-Score (harmonic mean of precision and recall):**

$$\text{F1-Score} = \frac{2\times(\text{Precision}\times\text{Recall})}{\text{Precision}+\text{Recall}}$$
$$= \frac{2\times(95\times96)}{95+96}$$
$$= 95.5\%$$



Decision Tree Algorithm:

data = pd. read_csv('connected_vehicle_dataset.csv')

X = data. Drop ('label', axis=1); y = data['label']

X_train, X_temp, y_train, y_temp = train_test_split (X, y, test_size=0.3, random_state=42)

Val, X_test, Yuval, y_test = train_test_split (X_temp, y_temp, test_size=0.5, random_state=42)

calf = DecisionTreeClassifier (max_depth=10, random_state=42)

clf.fit (X_train, y_train)

Predict on Validation Set: vapored = calf. predict(X_val)

Calculate Validation Metrics:

- **Accuracy**: Val accuracy = accuracy score (y_val, vapored)

- **Precision**: val_precision = precision_score (y_val, y_val_pred)
- **Recall**: val_recall = recall_score (y_val, y_val_pred)
- **F1-Score**: val_f1 = f1_score (y_val, y_val_pred)
- **Predict on Test Set**: y_test, pred = calf. predict(X_test)

Compute Confusion Matrix: cm = confusion matrix (y_test, y_test, pred)

Calculate Test Metrics:

- **Accuracy:** test accuracy = accuracy score (y_test, y_test, pred)
- **Precision:** test precision = precision_score (y_test, y_test, pred)
- **Recall:** test recall = recall_score (y_test, y_test, pred)
- **F1-Score:** test_f1 = f1_score (y_test, y_test, pred)

Plot Confusion Matrix:

- **Plot:** sns. Heatmap (cm, and not=True, ft='d', camp='Blues', xticklabels= ['Normal', 'Intrusion'], yticklabels= ['Normal', 'Intrusion'])

Load Dataset: data = pd. read_csv('connected_vehicle_dataset.csv')

Preprocess Data:

- **Separate Features and Labels**:: X = data. Drop ('label', axis=1)

y = data['label']

Split Data:

- **Training and Temp**: X_train, X_temp, y_train, y_temp = train_test_split (X, y, test_size=0.3, random_state=42)
- **Validation and Test**: X_val, X_test, y_val, y_test = train_test_split (X_temp, y_temp, test_size=0.5, random_state=42)

Initialize Random Forest Classifier: clf = RandomForestClassifier (n_estimators=100, max_depth=10, random_state=42)

Train Model: clf.fit (X_train, y_train)

Predict on Validation Set: y_val_pred = clf. predict (X_val)

Calculate Validation Metrics:

- **Accuracy**: Val accuracy = accuracy score (y_val, y_val_pred)

- **Precision**: val_precision = precision_score (y_val, y_val_pred)
- **Recall**: val_recall = recall_score (y_val, y_val_pred)
- **F1-Score**: val_f1 = f1_score (y_val, y_val_pred)

Predict on Test Set: y_test, pred = clf. predict(X_test)

Compute Confusion Matrix: cm = confusion matrix (y_test, y_test, pred)

Calculate Test Metrics:

- **Accuracy**: test_accuracy = accuracy score (y_test, y_test, pred)
- **Precision**: test precision = precision_score (y_test, y_test, pred)
- **Recall**: test recall = recall_score (y_test, y_test, pred)
- **F1-Score**: test_f1 = f1_score (y_test, y_test, pred)

Plot Confusion Matrix:

- **Plot**: plot. figure (figsize= (8, 6)) sns. Heatmap (cm, and not=True, ft='d', camp='Blues', xticklabels= ['Normal', 'Intrusion'], yticklabels=['Normal', 'Intrusion'])

plt. label('Actual')

plt. xlabel('Predicted')

plt. title ('Confusion Matrix for Intrusion Detection')

plt. show ()

4.4 Confusion Matrix

Overview of Confusion Matrix

A confusion matrix is a fundamental tool used in evaluating the performance of classification models. It provides a detailed breakdown of the model's predictions by comparing them against the actual outcomes. The matrix is structured as follows:

- **True Positives (TP):** Correctly predicted positive cases.
- **True Negatives (TN):** Correctly predicted negative cases.
- **False Positives (FP):** Incorrectly predicted positive cases (Type I error).
- **False Negatives (FN):** Incorrectly predicted negative cases (Type II error).

The matrix can be summarized as:

	Predicted Positive	Predicated Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

From the confusion matrix, various performance metrics can be derived:

- **Accuracy:** $(\mathrm{TP}+\mathrm{TN})/(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})$
- **Precision:** $\mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$
- **Recall (Sensitivity)**: $\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$
- **F1-Score:** $2\times(\mathrm{Precision}\times\mathrm{Recall})/(\mathrm{precision}+\mathrm{Recall})$

Comparison with Existing Models

To illustrate the practical significance of the confusion matrix, let's compare the performance of a Decision Tree and a Random Forest model for the IDS in connected vehicles.

Matrix	Decision Tree	Random Forest
Accuracy	85%	92%
precision	80%	88%
Recall	78%	90%
F1-Score	79%	89%

Advantages of Random Forest:

1. **Robustness:** By averaging the results of multiple trees, Random Forest reduces the risk of overfitting and improves generalization, making it more robust against noisy data and outliers.

2. **Feature Importance:** Random Forest can provide insights into the importance of different features, which can help in understanding which aspects of the data are most influential for intrusion detection.

3. **Versatility:** Random Forest handles large datasets and high-dimensional spaces more effectively than a single Decision Tree, which might struggle with overfitting in such scenarios.

Decision Tree Limitations:

1. **Overfitting:** A single Decision Tree may overfit the training data, especially if it's deep and complex, leading to poor performance on unseen data.

2. **Bias:** Decision Trees are sensitive to changes in the training data, which can lead to biased results if the data is not representative.

Summary:

In summary, the confusion matrix provides invaluable insights into the performance of classification models, allowing for detailed evaluation and comparison. While both Decision Trees and Random Forests have their advantages, Random Forests typically offer better performance and robustness, making them a more suitable choice for complex tasks like intrusion detection in connected vehicles.

5. CONCLUSION AND FUTHURE WORK

5.1 CONCLUSION:

This dissertation successfully developed a highly accurate and reliable Intrusion Detection System (IDS) for Connected Vehicles (IoV) within smart cities. By leveraging machine learning algorithms, specifically Random Forest and Decision Tree, the model demonstrated superior capability in detecting and classifying intrusions.

The model's performance, marked by an accuracy above 90%, showcased its effectiveness in processing complex vehicular data, including sensor data, network traffic logs, and Vehicle-to-Everything (V2X) communications. Crucially, the system achieves high precision and recall, ensuring that real threats are accurately identified while minimizing false alarms, which is vital for maintaining the safety and security of urban transportation systems in real-time. The ensemble approach of the Random Forest algorithm, in particular, provided robustness and resistance to overfitting, making the solution stable and dependable for a wide range of intrusion scenarios, such as denial-of-service (DoS) attacks and GPS spoofing.

5.2 Future Work

Future research should focus on extending the model's capabilities to meet the evolving demands of connected vehicular networks:

- **Deep Learning Integration:** Incorporate advanced Deep Learning (DL) methods, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to analyze intricate spatial and temporal patterns in time-series sensor and communication data, enhancing prediction accuracy.
- **Decentralized Real-Time Processing:** Implement Federated Learning (FL) and Edge Computing to enable decentralized, effective computation across numerous connected devices, thereby optimizing real-time threat detection and response latency.
- **Autonomous Systems Security:** Extend the model to specifically address new and emerging threats in autonomous and semi-autonomous vehicle networks, which have a larger attack surface due to their reliance on Artificial Intelligence for navigation and decision-making.
- **Advanced Data Integrity:** Investigate the integration of advanced encryption and blockchain-based security solutions to enhance data integrity and ensure the authenticity of critical Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications.

References:

[1] Khan, S. U. R., & Khan, Z. (2025). Detection of Abnormal Cardiac Rhythms Using Feature Fusion Technique with Heart Sound Spectrograms. Journal of Bionic Engineering, 1-20.

[2] Khan, M.A., Khan, S.U.R. & Lin, D. Shortening surgical time in high myopia treatment: a randomized controlled trial comparing non-OVD and OVD techniques in ICL implantation. BMC Ophthalmol 25, 303 (2025). https://doi.org/10.1186/s12886-025-04135-3

[3] Mahmood, F., Abbas, K., Raza, A., Khan,M.A., & Khan, P.W. (2019). Three Dimensional Agricultural Land Modeling using Unmanned Aerial System (UAS). International Journal of Advanced Computer Science and Applications (IJACSA) [p-ISSN : 2158-107X, e-ISSN : 2156-5570], 10(1).

[4] Hekmat, A., et al., Brain tumor diagnosis redefined: Leveraging image fusion for MRI enhancement classification. Biomedical Signal Processing and Control, 2025. 109: p. 108040.

[5] Khan, Z., Hossain, M. Z., Mayumu, N., Yasmin, F., & Aziz, Y. (2024, November). Boosting the Prediction of Brain Tumor Using Two Stage BiGait Architecture. In 2024 International Conference on Digital Image Computing: Techniques and Applications (DICTA) (pp. 411-418). IEEE.

[6] Khan, S. U. R., Raza, A., Shahzad, I., & Ali, G. (2024). Enhancing concrete and pavement crack prediction through hierarchical feature integration with VGG16 and triple classifier ensemble. In 2024 Horizons of Information Technology and Engineering (HITE)(pp. 1-6). IEEE https://doi.org/10.1109/HITE63532

[7] O. Bilal, Asif Raza, S. ur R. Khan, and Ghazanfar Ali, "A Contemporary Secure Microservices Discovery Architecture with Service Tags for Smart City Infrastructures ", VFAST trans. softw. eng., vol. 12, no. 1, pp. 79–92, Mar. 2024

[8] S.ur R. Khan, Asif. Raza, Muhammad Tanveer Meeran, and U. Bilhaj, "Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers", VFAST trans. softw. eng., vol. 11, no. 4, pp. 80–92, Dec. 2023.

[9] Khan, S.U.R., Asif, S., Bilal, O. et al. Lead-cnn: lightweight enhanced dimension reduction convolutional neural network for brain tumor classification. Int. J. Mach. Learn. & Cyber. (2025). https://doi.org/10.1007/s13042-025-02637-6

[10] Khan, U. S., Ishfaque, M., Khan, S. U. R., Xu, F., Chen, L., & Lei, Y. (2024). Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images. Frontiers of Structural and Civil Engineering, 1-17.

[11] Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 1-26.

[12] Latif, Sadia, Sami Ullah, Aafia Latif, Ghazanfar Ali, Muhammad Hassnain Azhar, and Salman Ali. "PREDICTIVE MODELING OF CARDIOVASCULAR DISEASE USING MACHINE LEARNING APPROACH." Kashf Journal of Multidisciplinary Research 2, no. 02 (2025): 207-232.

[13] Irtaza, G., Latif, S., Nadeem, R. M., Hussain, N., & Ijaz, H. M. (2025). Real-time Satellite Image Classification Using Convolutional Neural Networks for Earth Observation and Video Feed Analysis. Kashf Journal of Multidisciplinary Research, 2(03), 204-216.

[14] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.

[15] Khan, S.U.R., Zhao, M. & Li, Y. Detection of MRI brain tumor using residual skip block based modified MobileNet model. Cluster Comput 28, 248 (2025). https://doi.org/10.1007/s10586-024-04940-3

[16] Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. Computational Economics, 1-32.

[17] Khan, U. S., & Khan, S. U. R. (2025). Ethics by Design: A Lifecycle Framework for Trustworthy AI in Medical Imaging from Transparent Data Governance to Clinically Validated Deployment. arXiv preprint arXiv:2507.04249.

[18] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.

[19] HUSSAIN, S., Raza, A., MEERAN, M. T., IJAZ, H. M., & JAMALI, S. (2020). Domain Ontology Based Similarity and Analysis in Higher Education. IEEEP New Horizons Journal, 102(1), 11-16.

[20] Raza, A., & Meeran, M. T. (2019). Routine of Encryption in Cognitive Radio Network. Mehran University Research Journal of Engineering and Technology [p-ISSN: 0254-7821, e-ISSN: 2413-7219], 38(3), 609-618.

[21] Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). FOLC-Net: A Federated-Optimized Lightweight Architecture for Enhanced MRI Disease Diagnosis across Axial, Coronal, and Sagittal Views. arXiv preprint arXiv:2507.06763.

[22] Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). FloraSyntropy-Net: Scalable Deep Learning with Novel FloraSyntropy Archive for Large-Scale Plant Disease Diagnosis. arXiv preprint arXiv:2508.17653.

[23] Khan, Z., Khan, S. U. R., Bilal, O., Raza, A., & Ali, G. (2025, February). Optimizing Cervical Lesion Detection Using Deep Learning with Particle Swarm Optimization. In 2025 6th International Conference on Advancements in Computational Sciences (ICACS) (pp. 1-7). IEEE.

[24] Bilal, O., Hekmat, A., Shahzad, I. et al. Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization. Rev Socionetwork Strat (2025). https://doi.org/10.1007/s12626-025-00190-w

[25] Raza, Asif, Inzamam Shahzad, Muhammad Salahuddin, and Sadia Latif. "Satellite Imagery Employed to Analyze the Extent of Urban Land Transformation in The Punjab District of Pakistan." Journal of Palestine Ahliya University for Research and Studies 4, no. 2 (2025): 17-36.

[26] Asif Raza, Inzamam Shahzad, Ghazanfar Ali, and Muhammad Hanif Soomro. "Use Transfer Learning VGG16, Inception, and Reset50 to Classify IoT Challenge in Security Domain via Dataset Bench Mark." Journal of Innovative Computing and Emerging Technologies 5, no. 1 (2025).

[28] Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 204.

[29] Latif, Sadia, Sami Ullah, Aafia Latif, Ghazanfar Ali, Muhammad Hassnain Azhar, and Salman Ali. "PREDICTIVE MODELING OF CARDIOVASCULAR DISEASE USING MACHINE LEARNING APPROACH." Kashf Journal of Multidisciplinary Research 2, no. 02 (2025): 207-232.

[30] Latif, Aafia, Sadia Latif, and Furqan Jamil. "UTILIZING BIG DATA ANALYTICS TO OPTIMIZE INFORMATION COMMUNICATION TECHNOLOGY STRATEGIES." Kashf Journal of Multidisciplinary Research 1, no. 12 (2024): 122-140.

[31] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Li, X. (2025). Optimized deep learning model for comprehensive medical image analysis across multiple modalities. Neurocomputing, 619, 129182.

[32] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., & Li, Y. (2025). Optimize brain tumor multiclass classification with manta ray foraging and improved residual block techniques. Multimedia Systems, 31(1), 1-27.

[33] Asif Raza, Salahuddin, Ghazanfar Ali, Muhammad Hanif Soomro, Saima Batool, "Analyzing the Impact of Artificial Intelligence on Shaping Consumer Demand in E-Commerce: A Critical Review", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.17, No.5, pp. 42-61, 2025. DOI:10.5815/ijieeb.2025.05.04

[34] Khan, S.U.R., Raza, A., Shahzad, I., Khan, S. (2025). Subcellular Structures Classification in Fluorescence Microscopic Images. In: Arif, M., Jaffar, A., Geman, O. (eds) Computing and Emerging Technologies. ICCET 2023. Communications in Computer and Information Science, vol 2056. Springer, Cham. https://doi.org/10.1007/978-3-031-77620-5_20

[35] Khan, S. U. R., Asif, S., & Bilal, O. (2025). Ensemble Architecture of Vision Transformer and CNNs for Breast Cancer Tumor Detection from Mammograms. International Journal of Imaging Systems and Technology, 35(3), e70090.

[36] Maqsood, H., & Khan, S. U. R. (2025). MeD-3D: A Multimodal Deep Learning Framework for Precise Recurrence Prediction in Clear Cell Renal Cell Carcinoma (ccRCC). arXiv preprint arXiv:2507.07839.

[37] Khan, S. R., Asif Raza, Inzamam Shahzad, & Hafiz Muhammad Ijaz. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).

[38] Raza, A., Salahuddin, & Inzamam Shahzad. (2024). Residual Learning Model-Based Classification of COVID-19 Using Chest Radiographs. Spectrum of Engineering Sciences, 2(3), 367–396.

[39] Raza, A., Soomro, M. H., Shahzad, I., & Batool, S. (2024). Abstractive Text Summarization for Urdu Language. Journal of Computing & Biomedical Informatics, 7(02).

[40] Hekmat, A., Zuping, Z., Bilal, O., & Khan, S. U. R. (2025). Differential evolution-driven optimized ensemble network for brain tumor detection. International Journal of Machine Learning and Cybernetics, 1-26.

[41] Khan, S. U. R. (2025). Multi-level feature fusion network for kidney disease detection. Computers in Biology and Medicine, 191, 110214.

[42] Bilal, O., Hekmat, A., & Khan, S. U. R. (2025). Automated cervical cancer cell diagnosis via grid search-optimized multi-CNN ensemble networks. Network Modeling Analysis in Health Informatics and Bioinformatics, 14(1), 67.

[43] Latif, Sadia, Azhar Mehboob, Muhammad Ramzan, and Muhammad Ans Khalid. "DETECTION OF HCV LIVER FIBROSIS APPLYING MACHINE LEARNING TECHNIQUE." Kashf Journal of Multidisciplinary Research 1, no. 12 (2024): 11-44.

[44] Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). UNDERSTANDING URBAN EXPANSION THROUGH MULTI-TEMPORAL SATELLITE DATA ANALYSIS. Kashf Journal of Multidisciplinary Research, 2(09), 252-273.