Shahid Beheshti University

Robust Traffic Sign Recognition Through Collaborative Perception in VANETs Under Adverse Conditions

Bachelor's Thesis

Handed in by:	Mohammad Karbalaee Shabani
Matriculation number:	99222085
Supervisor:	Prof. Ziba Eslami
Faculty:	Mathematical Sciences
Department:	Computer Science
Editing time:	Sep 2024– Jan 2025

Abstract

Contents

Introduction

Traffic sign recognition plays a pivotal role in autonomous driving systems by enabling vehicles to interpret and respond to road signs in real-time. Accurate recognition is essential for ensuring the safety and efficiency of these systems. However, real-world scenarios introduce significant challenges. Environmental conditions such as poor lighting, rain, or fog, physical obstructions like overgrown trees or dirt-covered signs, and damaged or unclear signage can all hinder reliable recognition. These challenges emphasize the need for innovative solutions that address the limitations of traditional standalone recognition models [1].

Integrating traffic sign recognition with **V2X!** (**V2X!**) communication presents a promising avenue for overcoming these challenges. V2X communication fosters a connected environment where vehicles and infrastructure exchange data in real-time, enabling collaborative decision-making [2]. By leveraging V2X, vehicles can validate their recognition results through shared observations, reducing the risks associated with isolated errors and enhancing overall system reliability. This synergy holds the potential to improve traffic sign recognition by combining the strengths of machine learning techniques and connected vehicular ecosystems.

1 Problem Statement

Despite advancements in traffic sign recognition technology, existing systems often struggle in real-world conditions due to environmental factors, damaged signage, and obstructions. These limitations pose a risk to road safety, as errors in recognizing critical traffic signs can lead to incorrect or delayed responses.[3] Standalone recognition systems further exacerbate the problem by lacking a mechanism to cross-verify observations, leaving room for inaccuracies that may compromise autonomous driving systems' reliability.

While V2X communication offers a potential solution by facilitating real-time data sharing among vehicles, its implementation presents several challenges. These include ensuring data security, minimizing latency, and developing efficient mechanisms for aggregating shared data to derive consensus. Furthermore, research in this area remains limited, particularly in the context of applying V2X communication to enhance traffic sign recognition. Addressing these gaps is crucial for the safe and effective deployment of autonomous vehicles in complex, real-world scenarios.

2 Objectives

This thesis seeks to improve the reliability and accuracy of traffic sign recognition systems through the integration of V2X communication. The specific objectives of the research are as follows:

- **Develop a robust traffic sign recognition model:** Create a system capable of operating effectively under real-world conditions, accounting for challenges like environmental variability and occlusions.
- **Design a reliable consensus mechanism:** Develop an efficient method to aggregate recognition data from multiple vehicles, improving decision-making accuracy.
- **Simulate real-world scenarios:** Evaluate the performance of the integrated system in terms of recognition reliability, security, and efficiency within simulated environments.

3 Scope of the Study

This study contributes to the advancement of intelligent transportation systems by addressing critical challenges in traffic sign recognition and vehicular communication. Its findings are expected to enhance the safety, reliability, and efficiency of autonomous driving systems. By bridging the gap between recognition accuracy and collaborative data sharing through V2X, this research underscores the importance of secure and reliable vehicular communication in building public trust in autonomous technologies.

1 Background

This section provides an overview of the foundational concepts and technologies relevant to this research. It discusses V2X communication, its enabling technologies, applications, and associated challenges, as well as TSR as a critical component of intelligent transportation systems. Additionally, it introduces the dataset used for this study, highlighting its significance in training and evaluating machine learning models for traffic sign recognition.

1.1 Vehicle-to-Everything (V2X) Communication

Vehicle-to-Everything (V2X) communication is a groundbreaking technology that enables vehicles to exchange data with their surroundings, including other vehicles, infrastructure, pedestrians, and cloud-based systems. This interconnected framework is a cornerstone of modern **ITS!** (**ITS!**), designed to enhance road safety, improve traffic flow, and facilitate autonomous driving.

1.1.1 Types of V2X Communication

V2X encompasses several key components. **V2V!** (**V2V!**) communication allows direct data exchange between vehicles, enabling applications such as collision avoidance and coordinated lane changes. **V2I!** (**V2I!**) extends this interaction to roadside elements like traffic lights and road sensors, which provide vehicles with vital updates about traffic conditions or hazards. Additionally, **V2P!** (**V2P!**) communication ensures vehicles are aware of nearby pedestrians, even in scenarios with poor visibility. Finally, **V2C!** (**V2C!**) links vehicles to cloud servers for updates on navigation, weather, or software improvements [4].

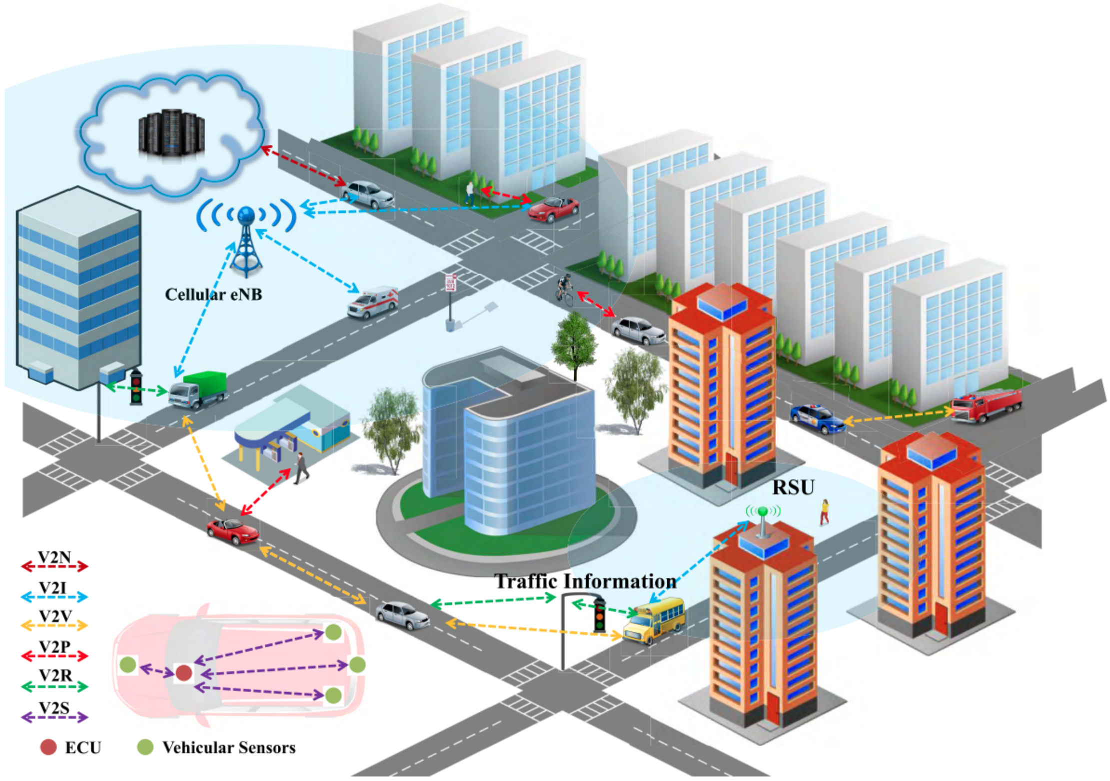

Figure 1: An overview of V2X scenario

Figure 1 illustrates a V2X communication network in a smart city environment, showcasing the interactions between vehicles, infrastructure, pedestrians, and networks. Various types of V2X communication are represented: **V2N!** (**V2N!**) connects vehicles to cloud-based systems via the Cellular **eNB!** (**eNB!**), which serves as the backbone of the cellular communication infrastructure. The Cellular eNB provides real-time updates and broad connectivity by leveraging 4G LTE and 5G technologies, enabling vehicles to access services such as navigation, traffic information, and emergency alerts [5].

Vehicle-to-Infrastructure (V2I) is enabled through **RSUs!** (**RSUs!**), which are positioned near roadways and intersections. RSUs act as intermediaries between vehicles and the infrastructure, collecting and disseminating localized traffic information such as signal timings, road hazards, or construction updates. These units enhance traffic management and safety by maintaining a continuous flow of communication with nearby vehicles and infrastructure elements like traffic lights and road signs [6].

1.1.2 Technologies Enabling V2X Communication

The technology behind V2X is built on two major standards. **DSRC!** (**DSRC!**), a Wi-Fi-based protocol, is optimized for low-latency, reliable communication, making it suitable for safety-critical applications like emergency braking. **C-V2X!** (**C-V2X!**), on the other hand, leverages 4G LTE and 5G networks to support broader connectivity, enabling advanced functionalities such as real-time updates and large-scale data sharing [7].

1.1.3 Applications of V2X Communication

Applications of V2X are vast and transformative. In addition to enhancing safety through collision prevention, V2X optimizes traffic management by reducing congestion and enabling efficient vehicle platooning. For autonomous vehicles, V2X complements onboard sensors like cameras and LiDAR, providing an additional layer of environmental awareness [8, 9].

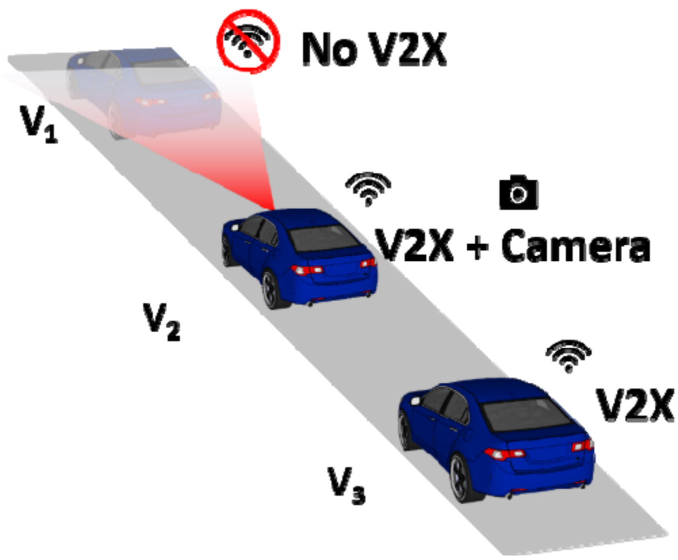

Figure 2: Illustration of the collaborative perception concept

1.1.4 Security Challenges in V2X Communication

Despite its potential, V2X faces several challenges. Security and privacy concerns arise from the constant exchange of real-time data, while ensuring seamless interoperability across manufacturers remains a significant hurdle [10].

A critical subset of V2X communication is the **VANET!** (**VANET!**), which enables vehicles to form dynamic, self-organized communication networks without relying on fixed infrastructure. In VANETs, vehicles act as both transmitters and receivers, exchanging information with other vehicles (V2V) and infrastructure (V2I). The decentralized nature of VANET allows for real-time communication, making it essential for time-sensitive applications. However, this decentralized architecture also introduces unique security and privacy challenges.

The challenges of VANET stem primarily from the misuse of information provided by vehicles. Clearly, the use of incorrect messages can lead to accidents and the adoption of erroneous strategies by traffic control centers. Therefore, the completeness and authenticity of messages must be verified before use [10].

Additionally, one of the essential security requirements of VANET is conditional privacy preservation. Conditional privacy means that while others are unable to identify vehicles based on their transmitted messages, it should still be possible to trace vehicles if necessary. Furthermore, given the coverage limitations of RSUs and the traffic volume within these areas, data compression becomes a critical issue that must be addressed in VANET systems to maintain efficient communication [10].

1.2 Traffic Sign Recognition (TSR!)

Traffic Sign Recognition (TSR) is a critical component of intelligent transportation systems (ITS) and autonomous vehicle technology. Its primary goal is to identify and interpret traffic signs to aid driver decision-making or to enable autonomous vehicles to navigate roads safely and efficiently. By accurately recognizing signs such as speed limits, stop signs, and warnings, TSR systems enhance road safety and contribute to the seamless integration of automated driving technologies [11].

The importance of TSR extends beyond autonomous vehicles. **ADAS!** (**ADAS!**) also rely heavily on TSR to provide real-time feedback to human drivers, reducing accidents caused by missed or misinterpreted traffic signs. For instance, TSR systems can alert drivers about an upcoming speed limit change or detect stop signs even in adverse weather conditions [12].

Developing robust TSR systems poses unique challenges, given the vast diversity of traffic signs worldwide, as well as the influence of environmental factors such as poor lighting, occlusion, and weather-related impairments. As a result, TSR has become a prominent research area in computer vision and machine learning, driving advancements in algorithms and models that strive to match human-level accuracy and reliability [13].

1.2.1 Datasets for TSR: The GTSRB Dataset

The **GTSRB!** (**GTSRB!**) dataset has been selected as the primary dataset for this study. Widely recognized in the field of traffic sign recognition, GTSRB serves as a standard benchmark for developing and evaluating machine learning models. Its comprehensive collection of real-world traffic sign images provides the foundation for building robust TSR systems [14].

The German Traffic Sign Recognition Benchmark (GTSRB) dataset is a pivotal resource in the field of computer vision, particularly for traffic sign detection and recognition tasks. Introduced during the International Joint Conference on Neural Networks (IJCNN) in 2011, the GTSRB dataset was designed to facilitate research and development in traffic sign recognition systems, which are crucial for intelligent transportation systems and autonomous vehicles [15]. The dataset encompasses a diverse collection of over 50,000 images, categorized into 43 distinct classes of traffic signs, including various speed limits, warning signs, and informational signs [16].

Figure 3: Representatives of the 43 traffic sign classes in the GTSRB dataset.

This extensive classification allows for comprehensive training and evaluation of machine learning models, particularly **CNNs!** (**CNNs!**), which have demonstrated significant efficacy in recognizing traffic signs [17].

The GTSRB dataset is particularly notable for its real-world applicability, as it includes images captured under varying conditions, such as different lighting and weather scenarios, which enhances the robustness of the models trained on it [18]. The dataset's structure supports both supervised learning and evaluation methodologies , making it an ideal benchmark for comparing the performance of various algorithms and architectures [19]. For instance, researchers have utilized GTSRB to validate innovative approaches, such as multi-column deep neural networks and attention mechanisms, which have shown promising results in improving classification accuracy [16, 20].

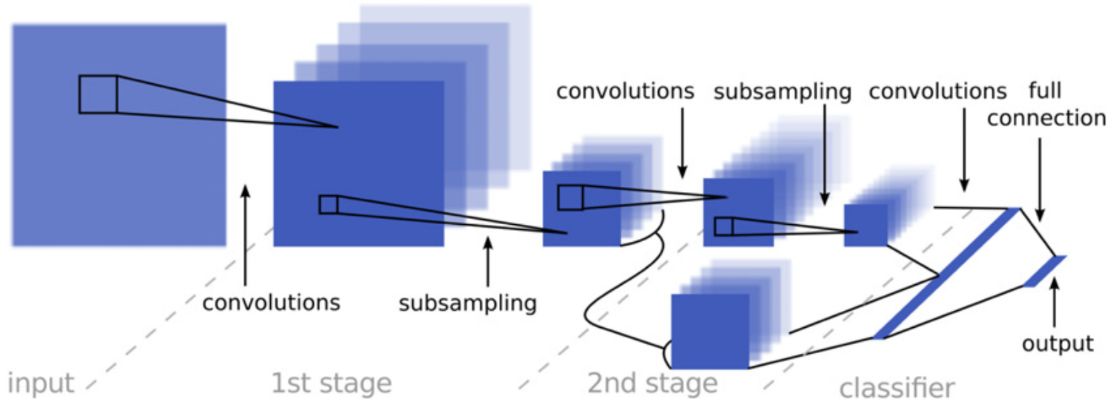

Figure 4: Example of a CNN architecture used for TSR

Moreover, the GTSRB dataset has been instrumental in advancing the understanding of adversarial attacks and model robustness in traffic sign recognition systems. Studies have employed GTSRB to evaluate the resilience of CNNs against various forms of adversarial perturbations, highlighting the importance of security in autonomous driving applications [21, 22]. The dataset's widespread use in the academic community underscores its significance as a benchmark for both foundational and cutting-edge research in traffic sign recognition [23].

In summary, the GTSRB dataset serves as a critical benchmark in the domain of traffic sign recognition, providing a rich resource for training and evaluating machine learning models. Its comprehensive collection of images, diverse classes, and real-world applicability make it an invaluable tool for researchers aiming to enhance the accuracy and reliability of traffic sign recognition systems.

2 Methodology

In this study, a deep learning model was developed for traffic sign recognition using the pretrained MobileNetV2 architecture. The application of transfer learning allowed the model to be fine-tuned effectively for the classification of German traffic signs, which significantly enhanced both training efficiency and performance on the target dataset. Transfer learning is particularly beneficial in scenarios where computational resources are limited, as it leverages existing knowledge from pretrained models to improve learning outcomes in specific tasks [24]. The MobileNetV2 architecture, known for its efficiency and low computational complexity, is well-suited for real-time applications in traffic sign recognition [25].

Following the training and optimization of the model, a Vehicular Ad Hoc Network (VANET) system was simulated to evaluate the model's real-world applicability in dynamic vehicular environments. VANETs are characterized by their ability to facilitate communication between vehicles, which is crucial for applications such as traffic sign recognition and real-time traffic management [26]. The dynamic nature of VANETs, due to the movement of vehicles, presents unique challenges for traffic prediction and management, making the integration of deep learning models essential for enhancing the accuracy and reliability of traffic sign recognition systems [27].

Finally, the model's accuracy was rigorously tested under various conditions, both within and outside the simulated VANET system, to assess its robustness and reliability across diverse contexts. The evaluation of traffic sign recognition systems in varying environmental conditions, such as different lighting and weather scenarios, is critical for ensuring their effectiveness in real-world applications [28]. The results from these tests provide valuable insights into the model's performance and its potential for deployment in intelligent transportation systems [29].

2.1 Deep Learning Model Training

A pretrained MobileNetV2 model was selected and fine-tuned by modifying its classifier layer to adapt it for traffic sign recognition. The model was subsequently trained on the prepared dataset to optimize its performance for accurate classification.

2.1.1 MobileNetV2

MobileNetV2 is a state-of-the-art CNN architecture specifically designed for mobile and edge devices, emphasizing computational efficiency and performance. It builds upon its predecessor, MobileNetV1, by introducing several enhancements that address the limitations of earlier models, particularly in terms of non-linearities and bottlenecks in narrow layers [30, 31]. The architecture is characterized by its use of depthwise separable convolutions and an inverted residual structure, which significantly reduces the number of parameters and computational cost while maintaining high accuracy in image classification tasks [31, 32].

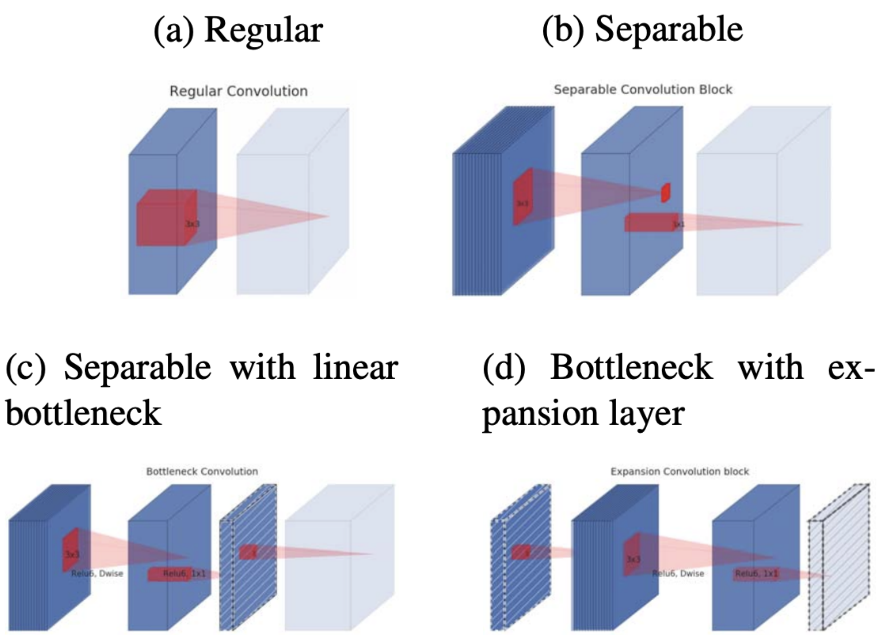

Figure 1: The difference between residual block and inverted residual.

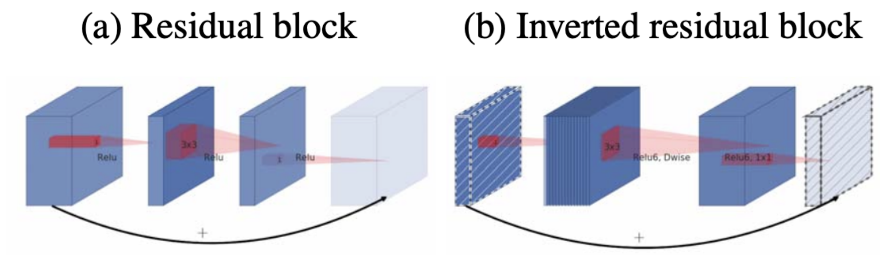

Figure 2: Evolution of separable convolution blocks

The structure of MobileNetV2 consists of 53 layers organized into 17 blocks, with a total of approximately 2.25 million parameters [33, 34]. Each block employs a linear bottleneck design, which allows for the reduction of channels before applying depthwise convolutions and subsequently expanding them again through pointwise convolutions. This design minimizes the introduction of non-linearities that could hinder performance, thus enhancing the model's ability to learn complex features while remaining lightweight [32, 35]. The architecture also incorporates global average pooling at the end, converting the spatial input into a fixed-size vector suitable for classification tasks [31, 35].

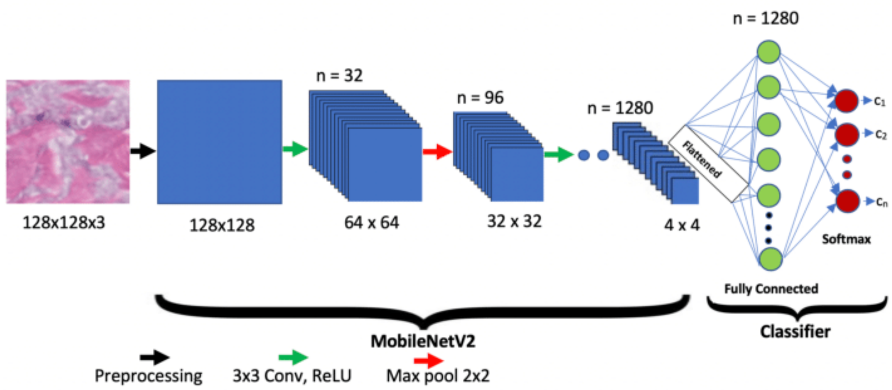

Figure 3: MobileNetV2 network architecture [33].

MobileNetV2 has demonstrated remarkable versatility across various applications, including image classification, feature extraction, and real-time inference on devices with limited computational resources [36, 37]. Its efficient design principles make it particularly well-suited for deployment in mobile applications, where speed and accuracy are critical [37, 38]. Furthermore, the architecture supports transfer learning, allowing it to leverage pre-trained weights for improved performance in specific tasks, even with smaller datasets [33, 37, 39].

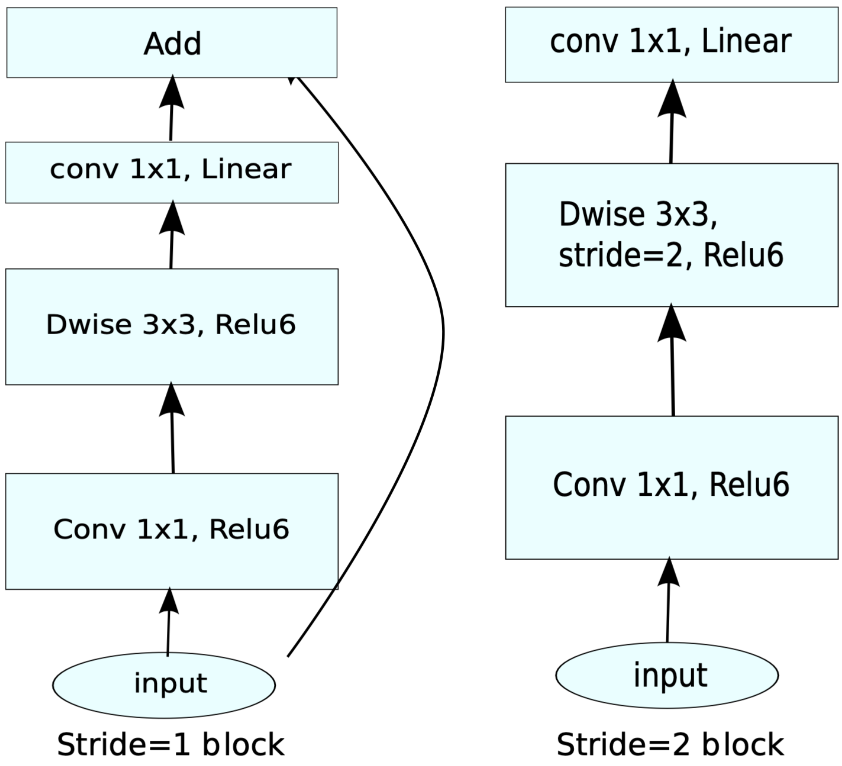

Figure 4: MobileNetV2 layers.

2.1.2 Pretrained Model Selection

The MobileNetV2 architecture was selected as the foundational model for traffic sign recognition due to its lightweight design and computational efficiency. MobileNetV2 is characterized by its use of depthwise separable convolutions and inverted residual structures, which significantly reduce the number of parameters and computational cost compared to standard convolutional networks [31]. This architectural choice is particularly advantageous for mobile and edge devices, where computational resources are limited, allowing for efficient processing without sacrificing performance [37]. The lightweight nature of MobileNetV2 makes it highly optimized for real-time applications, which is crucial for deployment in vehicular environments where low-latency processing is essential [37].

Moreover, initializing the model with pretrained weights on the ImageNet dataset allows MobileNetV2 to leverage rich feature representations, thereby expediting the convergence process during training [40]. This practice is supported by evidence that pretrained models can achieve better performance in various tasks, including image classification, by utilizing learned features from large datasets [40, 41]. The effectiveness of this approach is further underscored by studies demonstrating that pretrained networks can enhance model robustness and generalization capabilities, particularly in scenarios with limited labeled data [42].

In summary, the selection of MobileNetV2 for traffic sign recognition is justified by its efficient architecture, suitability for real-time applications, and the benefits derived from transfer learning through pretrained weights on the ImageNet dataset.

2.1.3 Dataset Partitioning

The German Traffic Sign Recognition Benchmark (GTSRB) dataset was partitioned to facilitate effective training, validation, and testing. The training dataset was split using an 80/20 ratio, where 80% of the data was allocated for training and 20% for validation. This stratified partitioning ensures that the model generalizes well by validating its performance on unseen data during training. The test dataset was kept entirely separate to provide an unbiased assessment of the model's final performance.

2.1.4 Image Preprocessing and Data Augmentation

Robust image preprocessing and data augmentation techniques were applied to enhance the model's generalization capabilities. Training images underwent a series of transformations, including resizing to 224x224 pixels, random horizontal flipping, random rotations up to 15 degrees, and color jittering to simulate variations in lighting conditions. These transformations artificially increased the diversity of the dataset, mitigating overfitting. All images were subsequently converted to tensors and normalized using the ImageNet mean and standard deviation values to standardize the input distribution. Testing images were subjected to resizing and normalization only, ensuring consistent evaluation conditions.

2.1.5 Fine-Tuning Strategy

To adapt MobileNetV2 for traffic sign recognition, the final classification layer was replaced with a fully connected layer comprising 43 output nodes, corresponding to the number of traffic sign classes in the GTSRB dataset. All preceding convolutional layers were frozen to preserve the pretrained feature extraction capabilities, while the new classifier layer was fine-tuned to specialize in distinguishing between traffic sign categories. This selective fine-tuning strategy balances computational efficiency with task-specific learning.

2.1.6 Training Configuration

The model was trained over 10 epochs with a batch size of 16, utilizing the Adam optimizer with a learning rate of 0.001. Cross-entropy loss was employed as the objective function to handle the multi-class classification task. A validation split of 20% within the training set enabled continuous performance monitoring and adjustment during training. These hyperparameters were chosen to provide an optimal trade-off between model convergence speed and generalization performance, ensuring efficient and stable learning across the dataset.

2.2 VANET Simulation

This VANET (Vehicular Ad-Hoc Network) simulation models how vehicles dynamically communicate with one another in a decentralized network. Each vehicle operates as a node capable of transmitting and receiving information within a specific communication range. This setup mirrors real-world vehicle-to-vehicle (V2V) communication used in intelligent transportation systems to improve traffic safety and efficiency.

2.2.1 Network Interface Structure

- **Transmission:** Each vehicle is equipped with a *Network Interface* that allows it to broadcast messages (such as alerts or detected hazards) to nearby vehicles.

Transmission is limited by a *communication range*, ensuring that only vehicles within proximity receive the information.

- **Reception Queue:** Vehicles maintain a *reception queue* where incoming messages are stored. This queue ensures that vehicles can process multiple messages in an orderly fashion.
- **Processing Messages:** Vehicles periodically check their reception queue and act upon the received information, allowing them to adapt to traffic conditions or hazards.
- **Distance Calculation:** A key feature is the simulation of distance on a *looped or circular road*, where the shortest path between vehicles is calculated, even if they are near the start and end of the road. This ensures continuous communication without boundary disruptions.

2.2.2 Vehicle Initialization and Communication

- **Unique Vehicle IDs:** Each vehicle is assigned a unique identifier, distinguishing it within the network.
- **Vehicle Positioning:** Vehicles are placed at specific positions along a simulated road, which is critical for calculating communication ranges and message dissemination.
- **Dynamic Interaction:** As the simulation progresses, vehicles move and continually check for nearby vehicles to send or receive messages, simulating real-world traffic flow.

2.2.3 Simulation Flow

1. Vehicles broadcast messages to nearby vehicles within their communication range.
2. Vehicles receive and store incoming messages in their reception queue.
3. Vehicles process messages to make decisions or pass the information along.

4. The simulation repeats these steps over multiple iterations, reflecting continuous vehicle interaction.

2.2.4 Communication Efficiency

- The simulation evaluates how effectively vehicles can transmit and receive data in dynamic traffic conditions.
- By simulating road constraints (looped roads) and communication ranges, it models real-world challenges in VANET systems.

2.3 Simulating Adverse Conditions

The primary goal of training the model and developing the VANET simulation system was to evaluate its performance under various conditions and assess its resilience.

For each condition, two sets of tests were conducted:

1. **Independent Vehicle Test:** A single vehicle operated without any connection or communication with other vehicles, relying solely on its own recognition capabilities.
2. **VANET System Test:** The same conditions were applied, but this time vehicles communicated and shared information. Predictions were made collaboratively through a consensus mechanism, allowing vehicles to collectively interpret traffic signs.

The environmental conditions used for testing were:

1. Normal (no alterations)
2. Brightness variation
3. Motion blur
4. Angle and rotation adjustments

2.3.1 Normal (no alterations)

To initiate the tests and simulations, both the system and the standalone vehicle were evaluated using the original dataset without any modifications or simulated conditions. This was a standard test conducted on the dataset's test set. Regardless of being part of the VANET system or operating independently, all vehicles utilized the same model to generate predictions on the identical dataset.

2.3.2 Brightness Variation

Considering that vehicles are equipped with headlights of varying intensities, producing different levels of brightness and illuminating varying distances ahead, we assumed that this could lead to differences in recognition performance. Each vehicle, having a unique set of headlights, may perceive and interpret its surroundings differently under such conditions.

Additionally, environmental lighting plays a significant role, as vehicles may be positioned at varying distances from streetlights, resulting in different visibility conditions for each vehicle.

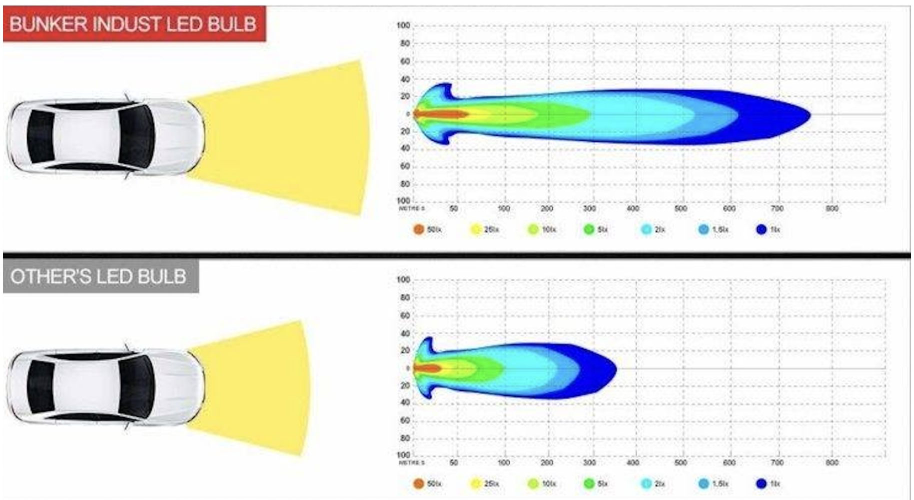

Figure 5: Different luminosity and light coverage on different vehicles.

To simulate this condition adjusting image brightness was used which involves scaling pixel intensity by a factor to simulate various lighting conditions. A factor of 1.0 keeps brightness unchanged, values above 1.0 brighten the image, and values below

1.0 darken it. This technique is commonly used in data augmentation to improve model robustness in computer vision tasks. The factor was chosen randomly for each car in the range of 0.3 to 1.3. This randomness ensures that the condition is correctly simulated. On the standalone case, the test dataset has been preprocessed with the same approach of multiplying the pixels by a random factor and then the model was tested in an isolated environment using the newly-created test dataset.

(a) Low brightness **(b)** Normal brightness **(c)** High brightness

Figure 6: Different luminosity and light coverage at random.

2.3.3 Motion Blur

Applying motion blur simulates the effect of object or camera movement by introducing a Gaussian blur to the image. This technique randomly selects a blur radius between 0.5 and 3.0, where lower values create mild blur and higher values produce more intense blur. The Gaussian blur smooths image details, mimicking the visual distortion caused by rapid movement or camera shake. This augmentation improves model robustness by enabling it to recognize objects and traffic signs under blurred or dynamic conditions. The approach for testing this effect on the systems was the same as the approach about brightness adjustment. It goes without saying that the idea behind this simulation comes from a normal situation in which different cars with various ranges of velocity pass by a traffic sign resulting into varied motion blur levels.

(a) Mild motion blur **(b)** Moderate motion blur **(c)** Severe motion blur

Figure 7: Simulation of different motion blur intensities on vehicle perception.

2.3.4 Angle and Rotation Adjustments

Rotation adjustment simulates the effect of a camera capturing images from various angles by randomly rotating the image within a specified range. In this case, the image is rotated by a random angle between -45° and 45°. This transformation, applied using bicubic resampling, maintains image quality while altering its orientation. This augmentation helps models become more robust in recognizing objects and traffic signs from different perspectives and tilted viewpoints.

(a) Slight rotation **(b)** Moderate rotation **(c)** Severe rotation

Figure 8: Simulation of various rotation angles affecting object recognition.

3 Results

This section provides a detailed report of the results obtained from the trained model. It presents the outcomes of various tests and performance metrics without engaging in evaluation or discussion. The purpose of this section is to systematically document the model's results, setting the foundation for deeper analysis in subsequent sections.

3.1 Model's Performance

The following sections offer a comprehensive report on the performance of the trained model. This includes results from multiple assessment methods and performance metrics. Key visualizations, such as the Calibration Curve, which illustrates the relationship between predicted probabilities and actual outcomes, and the Confusion Matrix, which details the model's classification accuracy across all classes, are presented. These results provide a clear and organized summary of the model's performance, which will be further evaluated and discussed in later sections.

3.1.1 Calibration Curve

The calibration curve provides insights into how well the model's predicted probabilities align with the actual outcomes. This analysis is crucial for understanding whether the model's confidence in its predictions is justified.

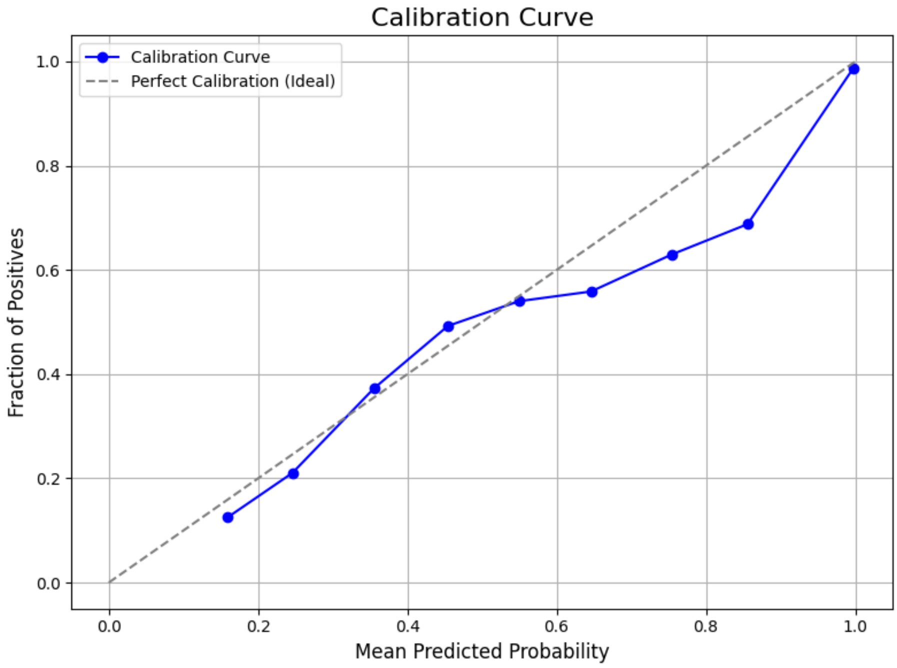

Figure 1: Calibration Curve showing the relationship between predicted probabilities and the fraction of positives.

Below the Diagonal (Underconfidence)

For lower predicted probabilities (approximately 0.1 to 0.4), the calibration curve lies **below** the ideal diagonal line. This behavior indicates that the model is **underconfident** in its predictions, predicting lower probabilities than the actual observed outcomes.

Near the Diagonal (Well-Calibrated)

In the mid-range probabilities (around 0.5 to 0.7), the curve aligns closely with the ideal line. This suggests that the model's predictions in this range are **well-calibrated**, with predicted probabilities accurately reflecting the true outcomes.

Above the Diagonal (Overconfidence)

For higher predicted probabilities (approximately 0.8 to 0.9), the curve rises **above** the diagonal. This indicates that the model becomes **overconfident**, predicting higher probabilities than what is observed in reality.

Sharp Rise at 1.0

At the highest predicted probability (1.0), the curve nearly touches the ideal calibration line. This suggests that predictions made with full confidence are relatively accurate.

3.1.2 Confusion Matrix

The analysis of the confusion matrix reveals comprehensive insights into the model's performance across various classes. The **overall accuracy** stands at approximately **95.61%**, confirming the model's high capability in accurately classifying the input data. This strong performance indicates effective generalization and robustness across the 43 distinct classes.

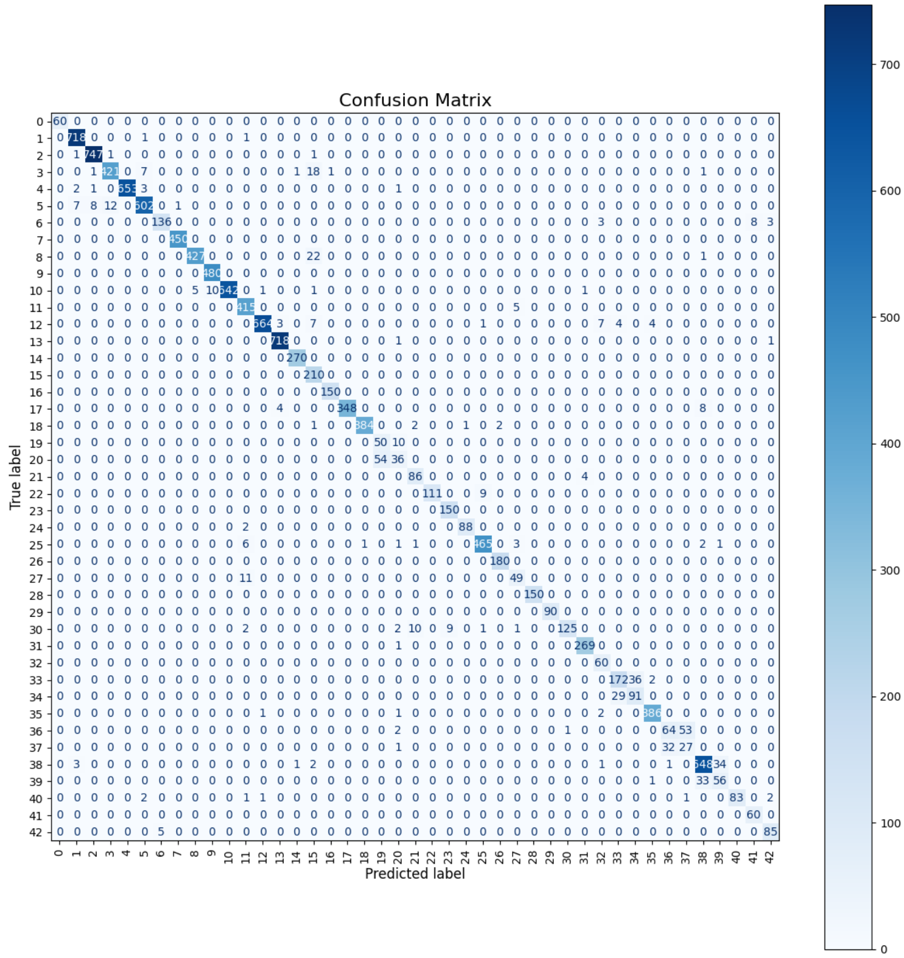

Figure 2: Confusion Matrix.

Further examination of the model's behavior is reflected in the class-wise metrics. The **average precision** is **90.69%**, indicating that when the model makes a positive prediction for a class, it is correct over 90% of the time. The **average recall** is slightly higher at **91.36%**, meaning the model successfully identifies more than 91% of actual instances for each class. This balance between precision and recall results in an **average F1-score** of **90.72%**, highlighting the model's consistent performance across various categories.

The confusion matrix also highlights misclassification trends. While the majority of predictions align with true labels, some off-diagonal elements reveal instances where certain classes are confused. These misclassifications could result from visual similarities between specific classes or limited training samples for those categories.

3.2 Tests

The following table summarizes the accuracy results of the model in both standalone and VANET simulation modes across different test conditions:

Condition	**Standalone Model Accuracy (%)**	**VANET Simulation Accuracy (%)**	**Difference (%)**
Normal (No Alteration)	95.68	95.68	0.00
Brightness Adjustment	95.51	95.46	-0.05
Motion Blur	64.08	66.37	+2.29
Rotation	62.15	77.66	+15.51

Table 1: Accuracy comparison between standalone model and VANET simulation across various conditions.

- **Normal Condition:** Both the standalone model and the VANET simulation achieved an accuracy of 95.68%. This identical performance is due to the deterministic nature of the model, which produces the same output when given the same input, rendering the consensus mechanism ineffective in this case.
- **Brightness Adjustment:** The standalone model achieved 95.51% accuracy, very close to the original performance. This robustness is attributed to the model being trained on images with varying brightness levels. The VANET simulation achieved 95.46%, with a negligible difference of 0.05%, indicating minimal impact due to the model's high baseline performance.
- **Motion Blur:** The standalone model's accuracy dropped to 64.08%. However, the VANET simulation improved the accuracy to 66.37%, reflecting a modest 2.29% enhancement due to the consensus mechanism, which helped correct some misclassifications.
- **Rotation:** This condition showed the most significant improvement. The standalone model achieved 62.15% accuracy, while the VANET simulation achieved 77.66%, marking a substantial 15.51% increase. This improvement highlights the effectiveness of the consensus mechanism in handling more challenging distortions.

4 Discussion

The results of the conducted tests underscore the significant impact of integrating traffic sign recognition within a Vehicular Ad-hoc Network (VANET), contingent upon varying situational contexts. This study aimed to evaluate how this integration could enhance vehicular communication and overall road safety by facilitating more reliable and accurate traffic sign detection.

Under optimal conditions—such as scenarios with high-resolution images and without any data interference or distortion—the VANET-based system did not present a distinct advantage over the standalone traffic sign recognition model. Both systems performed equivalently, indicating that in controlled environments where external factors are minimized, the addition of VANET does not yield measurable improvements.

However, in more dynamic and realistic driving conditions where variability among vehicles is a critical factor, the VANET system demonstrated a pronounced advantage. These conditions may include discrepancies in sensor calibration, diverse lighting and weather environments, varying vehicle speeds, and differences in computational resources across vehicles. In such complex and unpredictable scenarios, the collaborative nature of VANET proved to be instrumental in improving detection accuracy and system reliability.

This performance boost is primarily attributed to the system's ability to enable real-time data sharing and consensus-based decision-making among multiple vehicles. Through the proposed consensus and communication mechanism, vehicles collectively analyze and interpret traffic sign information. This cooperative strategy mitigates individual errors, reduces false positives and negatives, and allows for faster and more accurate recognition of traffic signs. For example, if one vehicle experiences sensor obstruction or misinterpretation due to environmental interference, other vehicles in the network can compensate by providing corroborative data.

Moreover, the VANET system enhances resilience against data anomalies and potential cyber threats by distributing the decision-making process across multiple nodes. This decentralized approach not only improves fault tolerance but also strengthens the system's resistance to malicious attacks aimed at compromising individual vehicles' perception systems.

The findings from this study emphasize the practical benefits of deploying VANET-integrated traffic sign recognition in real-world applications. The collective intelligence and communication fostered by this network significantly contribute to safer and more efficient driving experiences. By enabling vehicles to share and validate critical traffic information, the VANET system facilitates proactive responses to road conditions, reduces the likelihood of accidents, and optimizes traffic flow.

Future research should focus on further enhancing the communication protocols within VANETs to handle high-density traffic scenarios and ensuring scalability for widespread deployment. Additionally, exploring the integration of advanced machine learning algorithms and sensor fusion techniques could further improve the robustness and adaptability of the system across diverse driving environments.

In conclusion, while standalone traffic sign recognition systems perform adequately under ideal conditions, their integration within a VANET framework offers substantial advantages in more variable and challenging environments. The cooperative recognition and decision-making processes enabled by VANET significantly improve the reliability, accuracy, and safety of traffic sign detection, underscoring the value of collaborative vehicular communication systems in modern transportation networks.

5 Conclusions

using v2x doesn't have negative impact, but it has lots of positive impact.

the results vary from little non to lots.

Bibliography

[1] A. Yeola, C. Adak, S. Chattopadhyay, and S. Chanda. "Enhancing Traffic Sign Recognition: A Deep Learning Approach for Occluded Environments". In: *2024 IEEE International Conference on Computer Vision and Machine Intelligence (CVMI)*. 2024, pp. 1–6. DOI: `10.1109/CVMI61877.2024.10782104`.

[2] N. S. Pearre and H. Ribberink. "Review of research on V2X technologies, strategies, and operations". In: vol. 105. 2019, pp. 61–70. DOI: `https://doi.org/10.1016/j.rser.2019.01.047`.

[3] L. L. Avant, K. A. Brewer, A. A. Thieman, and W. F. Woodman. "Recognition errors among highway signs". In: vol. 1027. 1986, pp. 42–45.

[4] W. Tong, A. Hussain, W. X. Bo, and S. Maharjan. "Artificial Intelligence for Vehicle-to-Everything: A Survey". In: vol. 7. 2019, pp. 10823–10843. DOI: `10.1109/ACCESS.2019.2891073`.

[5] D. K. Nayak, A. Singh, S. R. V. Reddy, S. D. Roy, and S. Kundu. "Performance Analysis of NR 5G Cellular Vehicle to Everything Communication in Mode 3". In: vol. 137. 1. 2024, pp. 27–48. DOI: `10.1007/s11277-024-11295-w`.

[6] M. L. Chen, F. Ke, Y. Lin, M. J. Qin, X. Y. Zhang, and D. W. K. Ng. "Joint Communications, Sensing, and MEC for AoI-aware V2I Networks". In: 2024, pp. 1–1. DOI: `10.1109/TCOMM.2024.3519539`.

[7] A. Ahmad, M. N. Sial, M. Z. Awan, and S. Ullah. "Coverage probability of C-V2X network with full duplex communication on BSs over shared channels". In: vol. 87. 4. 2024, pp. 1167–1182. DOI: `10.1007/s11235-024-01220-8`.

[8] R. Miucic, A. Sheikh, Z. Medenica, and R. Kunde. "V2X Applications Using Collaborative Perception". In: *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. 2018, pp. 1–6. DOI: `10.1109/VTCFall.2018.8690818`.

[9] G. Sidorenko. “Cooperative Automated Driving for Enhanced Safety and Ethical Decision-Making”. In: Halmstad, 2024.

[10] N. Pakniat and Z. Eslami. “Security Analysis and Improvement of an Intelligent Transportation System based on Certificateless Aggregate Signature”. In: vol. 8. 1. Imam Hussein University, 2020, pp. 25–33. eprint: `https://ecdj.ihu.ac.ir/article_204746_3a7cab69ce0f1a3eb6d111d72dafeb51.pdf`.

[11] D. Thakur, D. Gholap, C. Kale, and B. Suryawanshi. “Automatic Self Driving Car”. In: 2024.

[12] J. Zhang, X. Zou, L.-D. Kuang, J. Wang, R. S. Sherratt, and X. Yu. “CCTSDB 2021: A More Comprehensive Traffic Sign Detection Benchmark”. In: vol. 12. 23. 2022. DOI: `10.22967/HCIS.2022.12.023`.

[13] A. Turquet, A. Wuestefeld, G. K. Svendsen, F. K. Nyhammer, E. L. Nilsen, A. P. Persson, and V. Refsum. “Automated Snow Avalanche Monitoring and Alert System Using Distributed Acoustic Sensing in Norway”. In: vol. 5. 4. 2024, pp. 1326–1345. DOI: `10.3390/geohazards5040063`.

[14] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel. “Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition”. In: *Neural Networks* 32 (2012). Selected Papers from IJCNN 2011, pp. 323–332. DOI: `https://doi.org/10.1016/j.neunet.2012.02.016`.

[15] A. Møgelmose, M. M. Trivedi, and T. B. Moeslund. “Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey”. In: *Ieee Transactions on Intelligent Transportation Systems* (2012). DOI: `10.1109/tits.2012.2209421`.

[16] D. Cireşan, U. Meier, J. Masci, and J. Schmidhuber. “Multi-Column Deep Neural Network for Traffic Sign Classification”. In: *Neural Networks* (2012). DOI: `10.1016/j.neunet.2012.02.023`.

[17] J. Zhang, M.-C. Huang, X. Jin, and X. Li. “A Real-Time Chinese Traffic Sign Detection Algorithm Based on Modified YOLOv2”. In: *Algorithms* (2017). DOI: `10.3390/a10040127`.

[18] “Traffic Sign Recognition Based on Up-Sampling Convolution”. In: (2019). DOI: `10.18178/wcse.2019.06.020`.

[19] T. Wang, H. Shen, Y. Xue, and Z. Hu. "A Traffic Signal Recognition Algorithm Based on Self-Paced Learning and Deep Learning". In: *Ingénierie Des Systèmes D Information* (2020). DOI: `10.18280/isi.250211`.

[20] Y. Garg, K. S. Candan, and M. L. Sapino. "SAN : Scale-Space Attention Networks". In: (2020). DOI: `10.1109/icde48307.2020.00079`.

[21] K. N. Kumar, V. Chalavadi, R. Mitra, and C. K. Mohan. "Black-Box Adversarial Attacks in Autonomous Vehicle Technology". In: (2020). DOI: `10.1109/aipr50011.2020.9425267`.

[22] X. Huang. "NeuronInspect: Detecting Backdoors in Neural Networks via Output Explanations". In: (2019). DOI: `10.48550/arxiv.1911.07399`.

[23] S. Kolouri, A. Saha, H. Pirsiavash, and H. Hoffmann. "Universal Litmus Patterns: Revealing Backdoor Attacks in CNNs". In: (2020). DOI: `10.1109/cvpr42600.2020.00038`.

[24] D. Tabernik and D. Skočaj. "Deep Learning for Large-Scale Traffic-Sign Detection and Recognition". In: *Ieee Transactions on Intelligent Transportation Systems* (2020). DOI: `10.1109/tits.2019.2913588`.

[25] R. Hashim, R. P. Singh, and M. Mehra. "Road Sign Detection System Using Neural Networks and Tensor Flow". In: *International Journal for Research in Applied Science and Engineering Technology* (2022). DOI: `10.22214/ijraset.2022.40672`.

[26] A. Priyadharshini. "An Efficient Key Agreement and Anonymous Privacy Preserving Scheme for Vehicular Ad-hoc Networks With Handover Authentication". In: *Concurrency and Computation Practice and Experience* (2023). DOI: `10.1002/cpe.7979`.

[27] S. S. Sepasgozar and S. Pierre. "Network Traffic Prediction Model Considering Road Traffic Parameters Using Artificial Intelligence Methods in VANET". In: *Ieee Access* (2022). DOI: `10.1109/access.2022.3144112`.

[28] W. Sun, H. Du, X. Zhang, and X. He. "Traffic Sign Recognition Method Integrating Multi-Layer Features and Kernel Extreme Learning Machine Classifier". In: *Computers Materials & Continua* (2019). DOI: `10.32604/cmc.2019.03581`.

[29] P. Perez-Murueta, A. Gómez-Espinosa, C. Cárdenas, and M. González-Mendoza. "Deep Learning System for Vehicular Re-Routing and Congestion Avoidance". In: *Applied Sciences* (2019). DOI: 10.3390/app9132717.

[30] M. H.-M. Khan, A. Makoonlall, N. Nazurally, and Z. Mungloo-Dilmohamud. "Identification of Crown of Thorns Starfish (COTS) Using Convolutional Neural Network (CNN) and Attention Model". In: *Plos One* (2023). DOI: 10.1371/journal.pone.0283121.

[31] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks". In: (2018). DOI: 10.1109/cvpr.2018.00474.

[32] H. Lahiani. "A Comparative Study of Two Deep Learning Architectures for Gesture Recognition on ArSL2018 Dataset". In: (2024). DOI: 10.21203/rs.3.rs-4523549/v1.

[33] M. Akay, Y. Du, C. L. Sershen, M. Wu, T. Y. Chen, S. Assassi, C. Mohan, and Y. M. Akay. "Deep Learning Classification of Systemic Sclerosis Skin Using the MobileNetV2 Model". In: *Ieee Open Journal of Engineering in Medicine and Biology* (2021). DOI: 10.1109/ojemb.2021.3066097.

[34] K. M. Sünnetci, E. Kaba, F. B. Çeliker, and A. Alkan. "Comparative Parotid Gland Segmentation by Using ResNet-18 and MobileNetV2 Based DeepLab V3+ Architectures From Magnetic Resonance Images". In: *Concurrency and Computation Practice and Experience* (2022). DOI: 10.1002/cpe.7405.

[35] M. K. Alam Mazumder. "A Robust and Light-Weight Transfer Learning-Based Architecture for Accurate Detection of Leaf Diseases Across Multiple Plants Using Less Amount of Images". In: *Frontiers in Plant Science* (2024). DOI: 10.3389/fpls.2023.1321877.

[36] N. S. Mor. "A Comparative Analysis of Convolutional Neural Network and Vision Transformer Embeddings on a Novel Domain-Specific Task". In: (2024). DOI: 10.21203/rs.3.rs-4496133/v1.

[37] I. M. Subrata Sandhiyasa. "Real Time Face Recognition for Mobile Application Based on Mobilenetv2". In: *Jurnal Multidisiplin Madani* (2023). DOI: 10.55927/mudima.v3i9.5924.

[38] A. Chaudhary. "Road Surface Quality Detection Using Light Weight Neural Network for Visually Impaired Pedestrian". In: *Evergreen* (2023). DOI: 10.5109/6792818.

[39] O. J. Jidan. "A Comprehensive Study of DCNN Algorithms-Based Transfer Learning for Human Eye Cataract Detection". In: *International Journal of Advanced Computer Science and Applications* (2023). DOI: 10.14569/ijacsa.2023.01406105.

[40] D. Le, M. N. Alam, C. K. Yao, J. I. Lim, Y.-T. Hsieh, R. V. Paul Chan, D. Toslak, and X. Yao. "Transfer Learning for Automated OCTA Detection of Diabetic Retinopathy". In: *Translational Vision Science & Technology* (2020). DOI: 10.1167/tvst.9.2.35.

[41] A. Mathis, T. Biasi, S. Schneider, Y. Mert, B. Rogers, M. Bethge, and M. W. Mathis. "Pretraining Boosts Out-of-Domain Robustness for Pose Estimation". In: (2019). DOI: 10.48550/arxiv.1909.11229.

[42] S. Rajaraman, F. Yang, Z. Liang, and Z. Xue. "Uncovering the Effects of Model Initialization on Deep Model Generalization: A Study With Adult and Pediatric Chest X-Ray Images". In: (2023). DOI: 10.1101/2023.05.31.23290789.