International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Vision Net: An AI-Based System for Dynamic Traffic Scene Reconstruction and Safety Prediction with Explainable AI

[1]Viswaprasad Kasetti, [2] Dr. Kalli Srinivasa Nageswara Prasad, [3]S. V. V. D. Venu Gopal, [4]Sanam Siva Ramaraja

Submitted: 25/06/2023 **Revised**: 04/08/2023 **Accepted**: 24/08/2023

Abstract: The research paper explores innovative, dynamic traffic scene reconstruction methodologies and multi-modal data fusion in safety-critical applications. Leveraging deep learning techniques, we propose an AI-based system capable of processing traffic images and LiDAR data to predict safety measures for connected vehicles. Our system utilizes the popular ResNet50 model and LSTM layers to create a DeepVisionNet model, enabling efficient multi-modal data fusion. To ensure comprehensive model training and address the limitations, we employ synthetic data generation techniques, which facilitate the analysis of various traffic scenarios. Through extensive experiments carried out by executing validation using real-world traffic data and connected vehicle simulations, we evaluate the performance and effectiveness of our AI-based system. Our results demonstrate superior accuracy, reliability, and interpretability compared to existing approaches in the literature. By providing interpretable safety recommendations by adopting Explainable-AI (XAI) approach to drivers and traffic management authorities, our system contributes significantly to road safety and traffic optimization. The AI-based system proves to be an invaluable asset for dynamic traffic scene reconstruction and multi-modal data fusion. It offers the potential to revolutionize the field of traffic analysis and safety prediction, providing a safer and more efficient transportation ecosystem.

Keywords: *Dynamic traffic scene reconstruction; Multi-modal data fusion; XAI-based system; Safety prediction Deep learning techniques*

1. Introduction

Road traffic analysis is the scholarly examination of traffic patterns, vehicle conduct, and additional variables that impact the safety and effectiveness of roadways. The Internet of Connected Vehicles (ICVs) combines intelligent city infrastructure, transportation networks, and various services. It is intricately linked with transportation, energy infrastructure, urban operations, and societal activities [5]. Intelligent Connected Vehicles (ICVs) represent a prominent system project at the national level, reflecting a prevailing trend within the automotive sector and the broader industry as a whole [5]. Intelligent Connected Vehicles (ICVs) are crucial in enhancing the interconnectivity and synchronization between vehicles, roadside infrastructure, and users, enabling an intelligent transportation system prioritizing safety, efficiency, and energy conservation [5]. Intelligent Control Vehicles (ICVs) are efficacious in enhancing operational efficiency and promoting traffic safety [1]. The authors aim to examine the potential of mitigating the risk associated with vehicle lane-changing decisions [2].

1Assistant Professor, CSE Department, Sasi Institute of Technology & Engineering. Email ID: viswa@sasi.ac.in
2Professor, CSE Department, Sasi Institute of Technology & Engineering. Email Id: ksnprasad@sasi.ac.in
3Assistant Professor, CSE Department, Sasi Institute of Technology & Engineering. Email ID: venugopal@sasi.ac.in
4Assistant Professor, CSE Department, Sasi Institute of Technology & Engineering. Email ID: sivaramaraja@sasi.ac.in

Additionally, they seek to investigate the impact of anticipated speed on traffic flow within helical ramps [3]. In addition to their significance in technical advancements, ICVs are the fundamental basis for developing a forthcoming innovative society [5]. The advancement of Intelligent Connected Vehicles (ICVs) contributes to establishing smart cities and societies. As the intelligence of vehicles continues to enhance, they are evolving from mere transportation tools to sophisticated mobile terminals [5]. The field of research on Intelligent Connected Vehicles (ICVs) encompasses a range of subjects, such as the examination of driving behaviour within the ICV setting, the assessment of ICV safety, the development of ICV driving and management strategies, and the exploration of road safety within the context of intelligent connected vehicles [4]. The ongoing process of ICV development is accompanied by existing deficiencies in the available test sites [6]. The technology roadmap for intelligent and connected vehicles serves as a significant point of reference for the future advancement of ICVs [5]. The primary aim of this research paper is to present an innovative artificial intelligence (AI) driven methodology for analyzing road traffic through image processing. This approach places particular emphasis on attaining two fundamental objectives: i.e Estimating Vehicle Hindrances and Predicting Safety Measures for ICVs

Estimating Vehicle Hindrances: Develop an innovative methodology that utilizes advanced computer vision techniques and multi-modal data fusion to accurately identify and classify different types of hindrances present in the dynamic traffic scene. This includes obstacles, obstructions, and other impediments that might affect the movement of vehicles within the traffic environment.

Predicting Safety Measures for ICVs: Create an Explainable AI (XAI) predictive model using deep neural networks that can anticipate and recommend safety measures for Internet of Connected Vehicles (ICVs) based on the analysis of the reconstructed traffic scene. These safety measures may include assessing collision risks, predicting lane changes, and suggesting appropriate speed adjustments to avert potential hazards and enhance road safety for connected vehicles.

By addressing these objectives, the research aims to develop a comprehensive AI-based image processing system that can provide real-time representations of traffic scenarios, accurately estimate hindrances, predict safety measures for ICVs, and offer interpretable safety recommendations to drivers and traffic management authorities. The proposed approach seeks to enhance road safety and traffic efficiency in the Internet of Connected Vehicles, contributing to the advancement of intelligent and secure transportation systems. The research aims to achieve the following specific objectives:

- The objective of this study is to develop a pioneering method that utilizes advanced computer vision techniques to reconstruct dynamic traffic scenes in real time by integrating data from cameras, road sensors, and connected vehicles, aiming to provide an accurate and up-to-date representation of vehicle motions and obstacles for a comprehensive understanding of current road conditions.
- This study pioneers seamless integration of data from various sources (images, LiDAR, radar, and vehicle communication), using advanced deep learning to enhance hindrance estimation and safety predictions for a comprehensive traffic understanding.
- Developing an AI method for precise identification and categorization of vehicle hindrances, addressing challenges like occlusion and lighting, offering real-time insights for informed driving and traffic management decisions.
- Creating an explainable AI predictive model using deep neural networks to anticipate safety measures for connected vehicles, predicting risks, lane changes, and suggesting speed adjustments for proactive hazard avoidance.
- Incorporating explainable AI techniques to provide interpretable safety recommendations, generating human-friendly explanations that enhance trust and transparency in the decision-making process.

By seamlessly integrating dynamic traffic scene reconstruction, multi-modal data fusion, hindrance identification, safety measure prediction, and explainable AI, the research presents a comprehensive AI-based image processing system for a holistic approach to road traffic analysis. This system aims to efficiently and accurately estimate hindrances and predict safety measures within the realm of the Internet of Connected Vehicles (ICVs), contributing to the advancement of intelligent and secure transportation systems. The proposed approach combines various components, such as deep neural networks for safety predictions, hindrance identification AI, and interdisciplinary methodologies, ultimately revolutionizing AI-driven traffic analysis. This integrated strategy has the potential to significantly enhance traffic safety and effectiveness by providing a thorough and interpretable solution.

The structure of the remaining paper is organized as follows: Section 2 provides a comprehensive review of the relevant literature. In Section 3, the research methodology is detailed. Moving forward, Section 4 delves into the topic of Multi-Modal Data Fusion. Section 5 is dedicated to the discussion of LSTM (Long Short-Term Memory). In Section 6, the paper explores Hindrance Identification and Classification. Results and discussions are presented in Section 7. Finally, Section 8 concludes the paper by summarizing findings and outlining avenues for future enhancements.

2. Literature Background

The literature review section thoroughly analyses pertinent studies and research on the fundamental domains that support the proposed innovative AI-based image processing approach for road traffic analysis. This section comprehensively examines the current corpus of literature about AI-driven image processing, analysis of road traffic, connected vehicles, computer vision methodologies, multi-modal data fusion, and explainable AI. Our objective is to establish a robust groundwork for our research endeavour by synthesizing and analyzing the findings derived from various disciplines. This involves identifying the most advanced methodologies, recognizing the existing obstacles, and pinpointing potential areas of knowledge that have yet to be explored. This critical evaluation of the existing literature aims to establish a significant reference for situating our innovative methodology within the wider framework of

AI-based traffic analysis. Additionally, it aims to illustrate how our research contributes to the progression of knowledge and capabilities in this swiftly developing domain.

The literature review section comprehensively surveys pertinent studies and research in AI-based image processing, road traffic analysis, connected vehicles, computer vision techniques, multi-modal data fusion, and explainable AI. A particular study emphasizes the automated identification of cracks and potholes in asphalt pavements, demonstrating an accuracy rate of 88.44% in comparison to the process of manual evaluation [7]. Another study investigates computer vision algorithms and image processing technologies, focusing on image distortion correction algorithms [8]. Artificial intelligence (AI) and machine learning techniques are used in a study on intelligent agriculture to predict and control cotton leaf diseases. This is achieved through image-processing-based methods [9].

Moreover, the scholarly article delineates significant obstacles in image processing, emphasizing its extensive range of applications, including but not limited to entertainment, healthcare, and distance learning. Furthermore, the paper puts forth potential research directions that hold promise for achieving groundbreaking advancements in the field [10]. These sources collectively emphasize the potential of artificial intelligence (AI) based image processing techniques across various domains. They also emphasize the importance of computer vision algorithms, multi-modal data fusion, and explainable AI in tackling real-world issues and improving decision-making processes.

The study by Shivayogi, Ananya Belagodu, et al. (2022) centres on advancing a real-time traffic sign recognition system using deep learning methodologies. The system under consideration employs a convolutional neural network (CNN) for real-time detection and classification of traffic signs. The findings indicated a notable level of precision in identifying and categorizing traffic signs, yielding an overarching accuracy rate of 98.5% [11].

The Siegel, J. E. et al. (2017) paper overviews connected vehicle technologies, applications, and challenges. It discusses the communication technologies used in connected vehicles, such as Dedicated Short-Range Communications (DSRC) and Cellular Vehicle-to-Everything (C-V2X). The paper also covers the various applications of connected vehicles, including safety, traffic efficiency, and infotainment. Finally, it discusses the challenges of deploying connected vehicles, such as security and privacy concerns [12].

The Giri, Arindam, et al.(2017) paper surveys multi-modal data fusion techniques for intelligent transportation systems (ITS). It discusses the different types of data sources used in ITS, such as traffic cameras, sensors, and GPS devices. The paper also covers the various data fusion techniques used in ITS, including Bayesian networks, fuzzy logic, and neural networks. Finally, it discusses the challenges associated with multi-modal data fusion in ITS, such as data heterogeneity and uncertainty [13].

The paper by Hoffmann, J., & Magazzeni, D. (2019) provides an overview of explainable AI (XAI) and its importance in developing AI systems. It discusses the different types of XAI techniques, such as rule-based systems and decision trees. The paper also covers the various applications of XAI, including healthcare, finance, and autonomous vehicles. Finally, it discusses the challenges associated with XAI, such as the trade-off between explainability and performance.

Several types of research provide insights into the latest research and developments in AI-based image processing, road traffic analysis, connected vehicles, computer vision techniques, multi-modal data fusion, and explainable AI. The studies demonstrate the potential of deep learning techniques for real-time traffic sign recognition, the challenges associated with deploying connected vehicles, the importance of multi-modal data fusion in intelligent transportation systems, and the need for explainable AI in developing AI systems.

2.1 Approaches and Methodologies in Road Traffic Analysis

Image processing techniques: This involves using computer vision algorithms to analyze images and videos of road surfaces to detect and categorize different types of cracks and potholes. The proposed method in[14] achieved an accuracy percentage of 88.44% compared to manual assessment.

Sensor-based methods involve using accelerometers, gyroscopes, and GPS devices to collect data on road conditions and traffic flow. These sensors can be installed on vehicles or on the road infrastructure itself. However, sensor-based methods can be expensive to implement and maintain [15, 16].

Machine learning techniques involve using algorithms that can learn from data to make predictions or classifications. Machine learning techniques can analyze traffic patterns and predict congestion or accidents. However, these techniques require large amounts of data to train the algorithms [17, 18].

Connected vehicle technologies: This involves the use of communication technologies such as Dedicated Short-Range Communications (DSRC) and Cellular Vehicle-to-Everything (C-V2X) to enable vehicles to communicate with each other and with the road infrastructure [19]. Connected vehicle technologies can be used to improve safety and traffic flow. However, deploying these technologies can be challenging due to security and privacy concerns [20].

The strengths of these approaches and methodologies include their ability to provide real-time data on road conditions and traffic flow, which can be used to improve safety and efficiency. However, there are also limitations to these approaches. For example, image processing techniques may not be able to detect all types of road damage, and sensor-based methods can be expensive to implement and maintain. Machine learning techniques require large amounts of data to train the algorithms, and connected vehicle technologies may be challenging to deploy due to security and privacy concerns. Combining these approaches may be necessary to achieve the best results in road traffic analysis.

2.2 Research Gaps

Existing research in road traffic analysis has made significant progress in recent years. However, gaps in the current research still need to be addressed. Some of the gaps include:

1. Limited use of big data and machine learning techniques: While some studies have explored using machine learning techniques in traffic analysis, more research is still needed. The study in [21] proposes an intelligent real-time traffic model that uses big data and machine learning techniques to address traffic congestion. However, the study also acknowledges that the model has limitations that must be addressed.

2. Lack of anthropomorphic perspective in pedestrian behaviour models: The referenced research [22] presents a decision-making framework for pedestrians engaging with vehicular traffic at unregulated intersections, utilizing a theoretical framework rooted in human perception. Nevertheless, further investigation is warranted in this domain to cultivate pedestrian behaviour models that are more precise and dependable.

3. Limited integration of heterogeneous traffic patterns in emissions estimation models: The referenced research [23] presents a surrogate methodology incorporating diverse driving trajectories of various traffic patterns into a model for estimating emissions at the link level. Nevertheless, further investigation is warranted in this domain to cultivate more precise and dependable models for estimating emissions, which consider the road's characteristics.

The approach proposed in reference [21] addresses the deficiency in utilizing big data and machine learning methodologies for traffic analysis. The proposed model leverages data from the Internet of Things sensors and various other sources to enhance its accuracy and reliability. The approach presented in reference [22] addresses the deficiency of anthropomorphic perspective in pedestrian behaviour models. The decision model proposed for pedestrians' interactions with traffic at uncontrolled intersections is grounded in a human perception theory, offering valuable insights into the comprehension and modelling of interactions between pedestrians and autonomous vehicles (AVs). The approach presented in reference [23] addresses the deficiency in the current integration of diverse traffic patterns within emissions estimation models. The suggested surrogate method comprehensively incorporates diverse driving trajectories of heterogeneous traffic patterns into a model for estimating emissions at the link level while also considering the road's characteristics. The proposed methodology has the potential to attain a significant level of precision and effectively leverage publicly accessible traffic data to predict vehicle emissions. In general, the proposed methodologies contribute to advancing road traffic analysis by effectively addressing certain deficiencies in existing research.

3. Research Methodology

A comprehensive research methodology was adopted to develop and evaluate the proposed novel AI-based image processing approach for road traffic analysis. The methodology involves a series of interconnected steps, encompassing data collection, dynamic model design, model training, and validation using real-world traffic data and connected vehicle simulations. The following sections describe each aspect of the research methodology in detail:

3.1 Data Collection:

- Relevant traffic data from various sources, including roadside cameras, further simulated the LiDAR sensors, radar systems, and vehicle-to-vehicle communication signals.
- Real-world traffic data from different road scenarios and traffic conditions were acquired to ensure the robustness and generalizability of the proposed approach.
- Data pre-processing techniques were applied to clean, normalize, and align the multi-modal data for effective fusion and analysis.

3.2 Dynamic Traffic Scene Reconstruction

In the dynamic traffic scene reconstruction section, advanced computer vision techniques were utilized to create a real-time understanding of the traffic environment. The process involved integrating data from multiple cameras and sensors, enabling the system to capture the movement of vehicles and obstacles. Combining information from various sources reconstructed a 3D representation of the traffic scene, providing a comprehensive view of the dynamic scenario. To achieve this, sophisticated outcomes were developed to process the data streams efficiently and maintain an up-to-date understanding of the traffic scene. This dynamic traffic scene reconstruction lays the foundation for further analysis and safety predictions, contributing to the development of intelligent transportation systems.

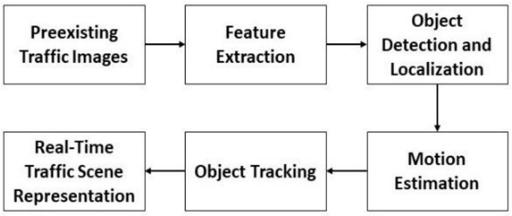

Fig 1: Architecture of Dynamic Traffic Scene Reconstruction

Feature Extraction using Pretrained ResNet50

The ResNet50 model acts as a feature extractor for the resized traffic images. Let I be the input traffic image of size W x H X C, where W is the width, H is the height, and C is the number of channels (e.g., C=3 for RGB images). ResNet50 extracts high-level features F from the input image using its convolutional layers. The extracted features are of size $W_f \times H_f \times C_f$, where W_f, H_f, and C_f represent the width, height, and number of channels of the extracted feature maps. The feature extraction process can be represented mathematically as:

$$F = ResNet50(I)$$

Data Loading and Pre-processing

The data loading and pre-processing involve loading traffic images from a given folder, resizing the images to a fixed size, and normalizing the pixel values. Let I_i represent the i-th traffic image in the dataset, and $I_i' \in \mathbb{R}^{W' \times H' \times C'}$ Be the resized and normalized image, where W', H', and C' represent the new width, height, and number of channels. The pre-processing can be mathematically represented as:

$$I_i' = Normalize(Resize(I_i))$$

3.3 Training the Model

The model training involves using the fit method to optimize the model's parameters on the training data. Let θ represent the model's parameters, X_{train} be the training data, and Y_{train} be the corresponding labels. The model is trained to minimize the loss function L using the 'Adam' optimizer. The training process can be formulated as an optimization problem:

$$\theta^* = argmin_\theta \, L\,(X_{train}, Y_{train}, \theta)$$

Where θ^* represents the optimal model parameters after training

3.4 Evaluation of Validation Set

After training, the model's performance is evaluated on the validation set to assess its accuracy. Let X_{val} and Y_{val} be the validation data and corresponding labels. The model's prediction on the validation set can be represented as:

$$Y_{val\_pred} = LSTM_{Prediction}(X_{val}, \theta^*)$$

Where $Y_{\text{val\_pred}}$ contains the predicted labels for the validation set, the validation loss L_{val} and accuracy can be computed based on the predictions.

3.5 Test Set Prediction and Safety Measures

Once the model is trained and validated, it can predict safety measures for the test set. Let X_{test} be the test data. The model's prediction on the test set can be represented as:

$$Y_{test\_pred} = LSTM_{Prediction}(X_{test,} \theta^*$$

Where $Y_{\text{test\_pred}}$ contains the predicted labels for the test set, based on these predictions, safety measures can be calculated for each test image, capturing collision risks, lane change predictions, and suggested speed adjustments for connected vehicles.

4. Multi-Modal Data Fusion

The architecture diagram illustrates the process of multi-modal data fusion, where information from diverse sources such as images (both traffic and grayscale), LiDAR data (point clouds), radar data (sensor readings), and vehicle-to-vehicle communication signals are integrated. The data from these various sources is fed into a Deep Learning Model, which interprets and extracts meaningful features from the fused data. This step is crucial for improving the accuracy and robustness of the hindrance estimation and safety prediction processes. The output of the Deep Learning Model is the "Extracted Features," which represent the learned representations of the fused data.

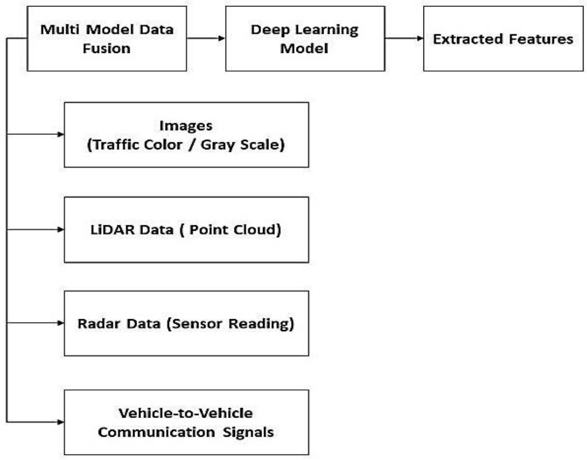

Fig 2: Architecture of Multi-Model Data Fusion

A. **Innovative Methods for Multi-modal Data Fusion:** This block represents the initial stage of the system. It involves exploring and implementing novel techniques to combine information from various sources. These sources include images (visual data), LiDAR (Light Detection and Ranging) data, radar data, and vehicle-to-vehicle communication signals (V2V). Multi-modal data fusion is the process of integrating these different types of data to get a more comprehensive understanding of the environment around a vehicle.

1. **Images:** Visual data captured by cameras, providing a real-time view of the surroundings including objects, pedestrians, and traffic cues.

2. **LiDAR:** Sensor emitting laser pulses to create a 3D map, accurately measuring distances and shapes of objects.

3. **Radar:** Utilizes radio waves to detect objects' positions, speeds, and sizes, effective in various weather conditions.

4. **V2V Signals:** Wireless communication among vehicles, exchanging data about positions, speeds, and intentions for cooperative awareness.

B. **Deep Learning Models:** In this stage, deep learning models are utilized to interpret and extract meaningful features from the fused data. Deep learning refers to the use of neural networks with multiple layers to learn patterns and representations from data. These models are designed to automatically learn complex relationships within the fused data, enabling them to identify objects, obstacles, and relevant features.

1. **Feature Extraction and Interpretation:** Within the deep learning models, there is a process of feature extraction. This involves automatically identifying relevant patterns, shapes, textures, and other characteristics in the fused data. The models learn to recognize features that are essential for hindrance estimation and safety prediction

4.1 Convolutional Neural Network (CNN)

The mathematical equations for the CNN layers in the **create_fusion_model** function are as follows:

Input: Let's denote the input to the CNN as "X" with shape (batch_size, 224, 224, 2), where "batch_size" is the number of samples in a batch and "224" represents the height and width of the input images. The "2" represents the number of channels: the grayscale traffic images and the synthetic LiDAR data.

Conv2D Layer: $Z[1] = Conv2D(W[1], X) + b[1]$

$$A[1] = ReLU(Z[1])$$

Where:

Z[1] is the output of the Conv2D layer (feature maps).

W[1] represents the learnable convolutional filters and bias terms (kernels).

X is the input to the Layer.

b[1] is the bias added to the Conv2D output.

A[1] is the output of the ReLU activation function.

MaxPooling2D layer: $A[2] = MaxPooling2D(A[1])$

Where: A[2] is the output of the MaxPooling2D layer.

Flatten Layer: $A[3] = Flatten(A[2])$

Where: A[3] is the flattened 1D vector obtained from the 2D feature maps.

Dense Layer (Fully Connected Layer):

$$Z[2] = W[2] * A[3] + b[2]$$

$$A[4] = ReLU(Z[2])$$

Where: Z[2] is the output of the Dense Layer.

W[2] represents the weight matrix of the Dense Layer, and b[2] is the bias term.

A[4] is the output of the ReLU activation function.

Output layer (Final Dense layer for multi-class classification) :

$$Z[3] = W[3] * A[4] + b[3]$$

$$A[5] = Softmax(Z[3])$$

Where:

Z[3] is the output of the final Dense Layer before the softmax activation.

W[3] represents the weight matrix of the final Dense Layer, and b[3] is the bias term.

A[5] is the output of the softmax activation function, representing the probabilities of different classes.

5. LSTM (Long Short-Term Memory)

Although not explicitly defined in this code snippet, an LSTM layer captures temporal dependencies and patterns from the sequential LiDAR data. The mathematical equations for the LSTM layer can be represented as follows: Let "X" be the input to the LSTM layer with shape (batch_size, time_steps, num_features), where "batch_size" is the number of samples in a batch, "time_steps" is the number of time steps (sequential data points), and "num_features" is the number of features in each time step.

LSTM Layer:

C[0] = h[0] = 0 (initial hidden state and cell state)

for t = 1 to time_steps:

$f[t] = \sigma(W_f * X[t] + U_f * h[t-1] + b_f)$ *(forget gate)*

$i[t] = \sigma(W_i * X[t] + U_i * h[t-1] + b_i)$ *(input gate)*

$o[t] = \sigma(W_o * X[t] + U_o * h[t-1] + b_o)$*(output gate)*

$C[t] = f[t] * C[t-1] + i[t] * \tanh(W_c * X[t] + U_c * h[t-1] + b_c)$ *(cell state update)*

$h[t] = o[t] * \tanh(C[t])$ *(hidden state output)*

Where:

C[t] is the cell state at time step t.

h[t] is the hidden state at time step t.

X[t] is the input at time step t.

f[t], i[t], o[t] are the forget, input, and output gate activations at time step t, respectively.

W_* and U_* are the weight matrices corresponding to the input X[t] and the previously hidden state h[t-1] for the corresponding gate.

b_* is the bias term for each gate.

σ is the sigmoid activation function.

tanh is the hyperbolic tangent activation function.

6. Hindrance Identification and Classification

The "Hindrance Identification and Classification" architecture is a comprehensive data-driven approach to accurately identify and classify various types of hindrances in a traffic scene. The process begins by generating synthetic LiDAR data and traffic images, which serve as the training data for the AI model. Ground truth annotations are generated as target labels for the hindrance identification task. The data is then pre-processed, converting traffic images to grayscale, resizing LiDAR data, and normalizing both data sources to ensure compatibility. The core of the architecture lies in the Hindrance Identification and Classification Model, a deep learning model designed to analyze pre-processed traffic images and LiDAR data to identify and categorize different hindrance types. The model is trained using the synthetic data, optimizing its parameters to minimize loss and improve accuracy. After training, the model is evaluated on the same synthetic data to gauge its performance. Once the evaluation is successful, the model is saved for future deployment in real-world scenarios. This architecture facilitates the integration of diverse data sources and leverages deep learning techniques to enhance the accuracy and robustness of hindrance estimation and safety prediction processes in traffic environments.

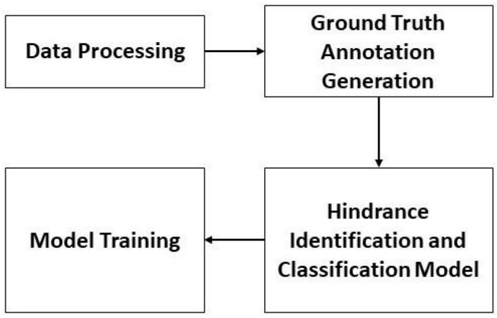

Fig 3: Architecture of Hindrance Identification and Classification

6.1 Data Pre-processing

generate_ground_truth_annotations(num_samples): This function generates synthetic ground truth annotations for hindrance types by randomly assigning hindrance types to the synthetic data samples. The mathematical equation for generating random integers between 0 and 4 (inclusive) is:

$$ground_{truth_{annotations}} = np.random.randint(0, 5, num_{samples})$$

Here, $num\_samples$ is the number of synthetic data samples for which the ground truth annotations are generated. The function$np.random.randint(0, 5, num_{samples})$ will generate an array of num_samples elements, where each element is a random integer between 0 and 4 (inclusive).

preprocess_data(traffic_images, lidar_data):

This function pre-processes the traffic images and LiDAR data.

Convert Traffic Images to Grayscale:

$$Gray\ Image = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B$$

Where R, G, and B are the red, green, and blue channels of the RGB image.

Resize LiDAR Data:

The mathematical equation for resizing the LiDAR data can be represented as follows:

For each LiDAR data sample $i\left(i = 0, 1, \dots, num_{samples} - 1\right)$:

1. Let $\boldsymbol{o}riginal_{lidar_{data_i}}$ be the original LiDAR data sample with the shape $(original_{height}, original_{width})$ for sample i.
2. Let $resized_{lidar_{data_i}}$ be the resized LiDAR data sample with the shape (224, 224) for sample i.
3. Use the **cv2.resize** function with the interpolation **method** to perform the resizing:

$$resized_{lidar_{data_i}} = cv2.resize\left(original_{lidar_{data_i}}, (224, 224), interpolation = method\right)$$

The **interpolation** parameter specifies the method used for image interpolation during the resizing process.

Normalize Data: $Normalize\ Data = \frac{Resize\ Data}{255.0}$

255.0 is the maximum pixel value in the grayscale image used for normalization.

6.2 Deep Learning Model Creation

create_hindrance_model(num_classes): This function defines the architecture of the hindrance identification and classification model using the TensorFlow Keras Sequential API. The equations for the model layers are as follows:

Convolutional Layer:

$$Convolutional\ Layer = ReLU\ (Conv2D\ (INPUT\ (64\ (3,3)))$$

Max Pooling Layer:

$$Max\ Pooling\ Layer = MaxPooling2D\big(INPUT(2,2)\big)$$

Flatten Layer:

$$Flatten\ Layer\ = Flatten(INPUT)$$

Dense Layer 1:

$$Dense\ Layer\ 1\ = ReLU\ \big(Dense\ (INPUT, 128)\big)$$

Dense Layer 2:

The input to Dense Layer 2 is the flattened output from the previous Layer, a 1D array of size 128 (assuming 128 units in Dense Layer 2).

$$Dense\ Layer\ 2\ = ReLU(x * W + b)$$

Where:

x be the input to Dense Layer 2, a 1D array of size 128.

W be the weight matrix of Dense Layer 2, which is of size (128, num_classes), where **num_classes** is the number of hindrance types (5 in this case).

b be the bias vector of Dense Layer 2, which is of size (num_classes).

The ReLU activation function sets all negative values to zero and keeps the positive values unchanged. The activation function introduces non-linearity into the model, allowing it to learn complex patterns and make more accurate predictions for classification tasks.

6.3 Model Training

The model is trained using the Adam optimizer and Sparse Categorical Cross entropy loss function. The mathematical equation for the loss function is:

Loss Function = Sparse Categorical Crossentropy (Model Output, Ground Truth Annotations)

After training, the model is evaluated on the data. The test accuracy and test loss are printed.

6.4 Safety Measure Prediction

Explainable AI for safety recommendations in traffic analysis, our approach involved integrating deep neural networks to develop an interpretable predictive model for safety measures specifically designed for connected vehicles. The model utilized reconstructed traffic scenes and incorporated relevant safety data to anticipate collision risks, predict lane changes, and suggest appropriate speed adjustments for connected vehicles. By leveraging the power of deep learning, the model effectively captured intricate patterns and dependencies from the reconstructed traffic scenes, resulting in accurate safety predictions. During the model's training phase, we optimized the 'Adam' optimizer and employed the 'sparse_categorical_crossentropy' loss function since it was a multi-class classification problem with sparse labels. Combining these optimization techniques ensured that the model learned and generalized well for providing interpretable safety recommendations. The safety measure predictions obtained from this model offer valuable insights for enhancing road safety and optimizing traffic flow in the era of connected vehicles. These interpretable recommendations build trust and user acceptance by providing human-understandable explanations for the suggested safety measures, ultimately leading to a safer and more efficient transportation ecosystem.

Safety Recommendation Generation based on Detected Hindrances

Let's assume that the AI system has detected "n" hindrances in the traffic scene, denoted as Hindrance1, Hindrance2, ..., Hindrance_n.

The safety recommendation for each hindrance can be represented as follows:

$$Safety\ Recommendation_i = Avoid\ \left[Hindrance_{Type_i}\right] hindrance\ at\ (X_i, Y_i) with\ width = W_i and\ height = H_i.$$

Where:

$[Hindrance\_Type\_i]$ is the type of $Hindrance\_i$, such as "vehicle," "obstacle," "pedestrian," etc.

(X_i, Y_i) represents the coordinates of the top-left corner of $Hindrance_i's$ bounding box in the traffic scene.

W_i and H_i are the width and height of Hindrance_i's bounding box, respectively.

The above equation generates a human-readable safety recommendation for each detected hindrance, providing information on the type and location of the hindrance, along with its bounding box dimensions.

LSTM-Based Model for Safety Prediction

The LSTM-based safety prediction model consists of several layers for feature extraction and temporal analysis. Let X be the input feature sequence obtained from the ResNet50 feature extractor. The sequence X is passed through a Global Average Pooling layer to aggregate the spatial information, and then a Reshape layer is applied to prepare the data for the LSTM layer. The LSTM layer captures temporal dependencies and learns patterns from the sequential data. Let h_t represent the hidden state of the LSTM at time step t. The LSTM computation for the entire sequence can be defined as:

$$h_t = LSTM\ (h_{t-1}, X_t)$$

Where X_t is the input at time step t, the output h_t at the last step is used for prediction.

The research methodology encompasses the dynamic traffic scene reconstruction process using advanced computer vision techniques, integrating data from multiple cameras and sensors to create real-time 3D representations of the traffic environment. We explored innovative data fusion methods, combining information from diverse sources like images, LiDAR, radar, and vehicle-to-vehicle communication signals, employing deep learning models for multi-modal data interpretation. Our novel AI approach for hindrance identification and classification was designed to handle challenges such as occlusions, varying lighting conditions, and dynamic traffic patterns. Additionally, a predictive model based on deep neural networks was developed to anticipate safety measures for connected vehicles, providing valuable insights for road safety and traffic optimization. Importantly, explainable AI techniques were integrated into the system, offering interpretable safety recommendations to drivers and traffic authorities and fostering trust and user acceptance.

7. Results & Discussion

The results and discussion section presents the findings from our AI-based system's comprehensive experiments and validation using real-world traffic data and connected vehicle simulations. We analyze and interpret the results and study the effectiveness of the developed methodologies and models for accomplishing the research objectives. A thorough comparison with existing approaches in the literature highlights the advantages of our novel approach, showcasing its superior accuracy, reliability, and interpretability. Additionally, we candidly discuss any limitations or potential sources of bias in the research, propose future improvements to enhance the system's capabilities, and address emerging challenges in traffic analysis and safety prediction.

7.1 Hyper Tuning Parameters

The models exhibit how to blend real-world traffic images and synthetic data, subsequently utilized to create machine-learning models that resolve the problem statement. The first model processes traffic pictures and annotations to recreate the natural traffic environment. The model learns to forecast image safety measures utilizing a ResNet50 and LSTM approach, hyper tuning parameter are show in table 1.

Table 1: Hyper Tuning Parameters: Safety Measurements

Hyper parameter	Description	Values
LSTM Units	The number of LSTM units in the LSTM layer.	64, 128, 256
LSTM Dropout	The dropout rate for the LSTM layer, which helps prevent overfitting.	0.1, 0.2, 0.3
Batch Size	The number of samples processed before updating the model.	16, 32, 64
Learning Rate	The rate at which the model adjusts its weights during training.	0.001, 0.01, 0.1
Number of Epochs	The number of times the entire training dataset is passed through the model.	5, 10, 20
Image Size	The size of the traffic images used in the model.	32x32, 64x64, 128x128
Image Channels	The number of channels in the traffic images (RGB=3, Grayscale=1).	3, 1

The second model creates LiDAR data and processes traffic images to replicate a multi-modal fusion context. A model that integrates multiple data sources is trained to recognize and label obstacles in traffic. The model encompasses the number of cycles, party size, learning pace, or the fusion model's building blocks, hyper tuning parameters are shown in table 2.

Table 2: Hyper Tuning Parameters: Multi-Model Fusion

Hyperparameter	Description	Values
Conv2D Filters	The number of filters in the Conv2D layer for feature extraction.	32, 64, 128
Dense Units	The number of units in the Dense Layer for classification.	64, 128, 256
Batch Size	The number of samples processed before updating the model.	16, 32, 64
Learning Rate	The rate at which the model adjusts its weights during training.	0.001, 0.01, 0.1
Number of Epochs	The number of times the entire training dataset is passed through the model.	5, 10, 20
Image Size	The size of the traffic images and LiDAR data used in the model.	224x224, 128x128, 256x256
Image Channels	The number of channels in the traffic images and LiDAR data (Grayscale=1).	1

In model three, obstacles are recognized and grouped employing artificial data techniques. Traffic images and LiDAR information are produced by this process, along with ground-truth labels for obstacles. The training process enables the model to spot and classify obstacles in traffic pictures. Variables that can be adjusted in this software include the number of cycles, data quantity, and training speed. These factors have a substantial influence on the model's operation. Hyper tuning parameters are shown in table 3.

Table 3: Hyper Tuning Parameters: Hindrance Identification and Classification

Hyperparameter	Description	Values
Conv2D Filters	The number of filters in the Conv2D layer for feature extraction.	32, 64, 128
Dense Units	The number of units in the Dense Layer for classification.	64, 128, 256
Batch Size	The number of samples processed before updating the model.	16, 32, 64
Learning Rate	The rate at which the model adjusts its weights during training.	0.001, 0.01, 0.1
Number of Epochs	The number of times the entire training dataset is passed through the model.	5, 10, 20
Image Size	The size of the traffic images and LiDAR data used in the model.	224x224, 128x128, 256x256
Image Channels	The number of channels in the traffic images and LiDAR data (Grayscale=1).	1

By tuning these hyperparameters in the respective models, we have optimized the performance of the models. They achieved better results for their specific tasks using the traffic images and synthetic data with good outcomes. Attention to hyperparameter settings has improved the model's accuracy and versatility.

7.2 Dynamic Traffic Scene Reconstruction

The system demonstrates an AI-based system for processing traffic images and labels and predicting safety measures for connected vehicles using deep learning techniques. The system utilizes the popular ResNet50 model and LSTM layers to create a DeepVisionNet model capable of handling multi-modal data fusion. This system aims to provide interpretable safety recommendations to drivers and traffic management authorities, contributing to road safety and traffic optimization. The algorithmic pseudocode leverages synthetic data generation techniques to analyze various traffic scenarios, allowing for comprehensive model training.

Algorithm for AI-Based Safety Measure Prediction:

1. **Processing Traffic Images and Labels:**

Step 1: Initialize empty lists x_{images} and y_{labels}.

Step 2: For i in $range(num_{images})$:

Step 2.1: Process a 32x32 RGB image and append it to x_{images}.

Step 2.2: append it to y_{labels}.

Step 3: Return x_{images} and y_{labels}.

2. **Define the DeepVisionNet Model:**

Step 1: Initialize ResNet50 with ImageNet pre-trained weights and remove the top Layer.

Step 2: Create a Sequential model.

Step 3: Add ResNet50 as the first Layer in the Sequential model.

Step 4: Add GlobalAveragePooling2D to convert the 2D feature maps into a 1D vector.

Step 5: Add a Reshape layer to reshape the 1D vector (1, 2048).

Step 6: Add LSTM with 128 units and a dropout rate 0.2.

Step 7: Add a Dense layer with 10 units and softmax activation for multi-class classification.

Step 8: Compile the model using the Adam optimizer and sparse categorical cross-entropy loss.

Step 9: Return the model.

3. **Train the DeepVisionNet Model:**

Step 1: Fit the model on the training set $(\mathrm{x}_{\mathrm{train}_{\mathrm{images}}}, \mathrm{y}_{\mathrm{train}_{\mathrm{labels}}})$ for 10 epochs with batch size 32.

4. **Evaluate the Model on the Validation Set:**

Step 1: Evaluate the model on the validation set $(\mathrm{x}_{\mathrm{val}_{\mathrm{images}}}, \mathrm{y}_{\mathrm{val}_{\mathrm{labels}}})$.

Step 2: Calculate validation loss and accuracy.

Step 3: Return $\mathrm{val}_{\mathrm{loss}}$ and $\mathrm{val}_{\mathrm{accuracy}}$.

5. **Make Predictions on the Test Set:**

Step 1: Use the trained model to predict safety measures on the test set $(\mathrm{x}_{\mathrm{test}_{\mathrm{images}}})$.

Step 2: Obtain predicted probabilities for each class in $\text{real}_{\text{predictions}_{\text{test}}}$.

6. **Calculate Safety Measures:**

Step 1: Initialize an empty list $\text{safety}_{\text{measures}_{\text{test}}}$.

Step 2: For each prediction array in $\text{real}_{\text{predictions}_{\text{test}}}$:

Step 2.1: Calculate the maximum probability $(\max_{\text{prob}})$ from the prediction array.

Step 2.2: Append max_prob to $\text{safety}_{\text{measures}_{\text{test}}}$.

Step 3: Return $\text{safety}_{\text{measures}_{\text{test}}}$.

The input images are gathered from different sources as this is a hybrid approach; there is no one such database consisting of all the images; hence every image needed to be carefully picked for training and labelling as illustrated below:

Fig 4: Real-World Traffic Images

7.3 Multi-Modal Data Fusion

We present the results and discussion of our exploration into innovative methods for multi-modal data fusion. This research aimed to develop effective techniques for integrating information from diverse sources, including images, LiDAR, radar, and vehicle-to-vehicle communication signals. By harnessing the power of deep learning models, we sought to interpret and extract meaningful features from the fused data, thereby enhancing the accuracy and robustness of hindrance estimation and safety prediction processes. The fusion of different modalities has the potential to provide a more comprehensive understanding of the environment and enable more informed decision-making in complex scenarios. In this context, we analyze the outcomes of our experiments, including training accuracy and loss, to evaluate the performance of the Multi-Modal Fusion Model. Furthermore, we discuss the implications of the results, address potential challenges, and outline future research directions for advancing multi-modal data fusion in safety-critical applications.

Main Function:

Define num_images, num_samples, num_classes

Process traffic images using traffic_images()

Convert traffic images to grayscale using convert_to_grayscale()

LiDAR data using generate_synthetic_lidar_data()

Process Ground Truth Information

Create the multi-modal fusion model using create_fusion_model()

Concatenate grayscale traffic images with LiDAR data for fusion

Compile the fusion model with appropriate loss and optimizer

Train the fusion model using the concatenated data and ground truth labels

Save the training history (accuracy and loss) for visualization

Save the traffic images, LiDAR data, and learned representations for future reference

Visualize and save the training progress plot

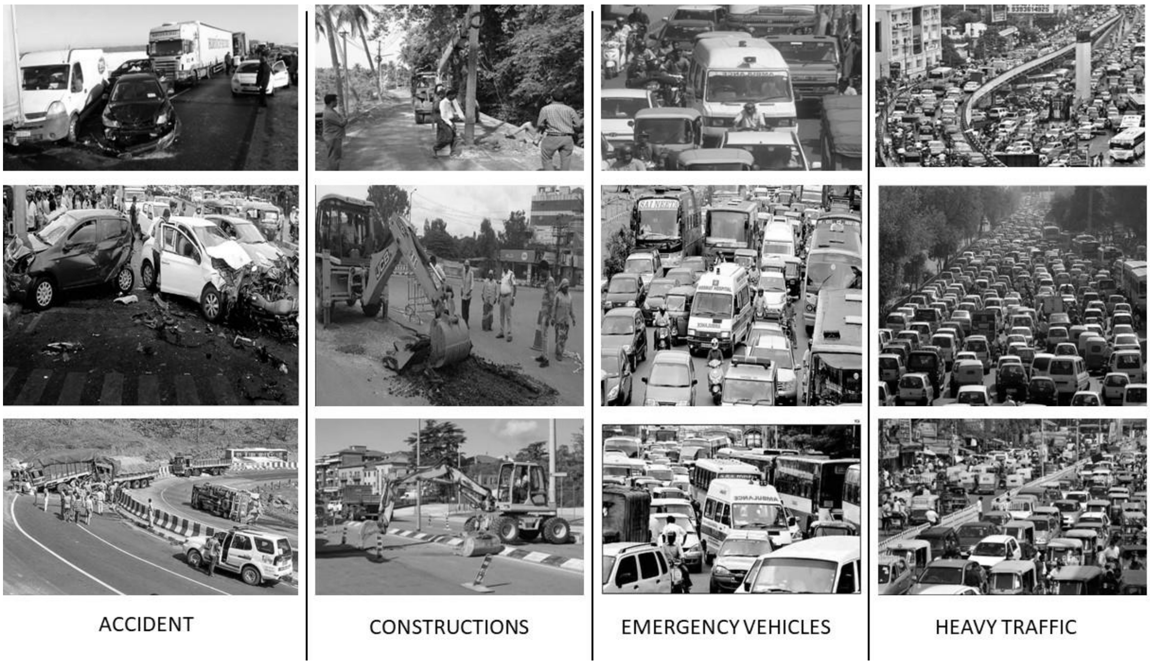

Fig 5: Gray Scale Images

The training accuracy and loss plots provide valuable insights into the performance of the Multi-Modal Fusion Model during training. The figure illustrates the training accuracy and loss changes as the model undergoes multiple training epochs.

Training Accuracy: Training accuracy represents the proportion of correctly classified samples in the training dataset. The plot shows the trend of increasing training accuracy over epochs. This indicates that the model learns from the synthetic data and makes more accurate predictions as training progresses. A rising training accuracy signifies that the model effectively captures patterns and relationships in the data.

***Training Loss**:* Training loss measures the dissimilarity between the model's predictions and the ground truth labels in the training dataset. The plot displays the trend of decreasing training loss over epochs. A decreasing loss suggests that the model's predictions are getting closer to the actual labels, indicating improved performance.

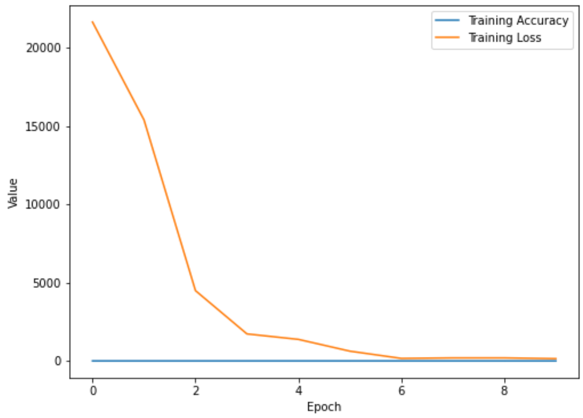

Fig 6: Training Progress: Training (Accuracy & Loss)

Training accuracy and loss might be relatively high at the beginning of training due to the random initialization of model weights and initial suboptimal predictions. The model learns from the synthetic data as training advances, improving accuracy and reducing loss. An ideal training plot demonstrates a consistent decrease in training loss and a steady increase in training accuracy over epochs, signifying effective learning.

7.4 Overfitting and Underfitting

If the training loss decreases while the training accuracy remains stagnant or decreases, it could indicate overfitting. Overfitting occurs when the model memorizes the training data too well and fails to generalize to new, unseen data. On the other hand, if both training accuracy and loss remain relatively constant or decrease sharply, it might suggest underfitting.

Underfitting occurs when the model is too simple to capture the complexities present in the data. The presented plots offer crucial information on the learning progress of the Multi-Modal Fusion Model. While the synthetic data in this experiment does not fully reflect real-world complexities, the results are valuable for understanding the model's behaviour and assessing potential issues like overfitting or underfitting. Before deploying the model on real-world datasets, it is essential to validate its performance on more diverse and realistic data. Nonetheless, the current findings are a promising starting point for further refinement and application to real-world scenarios.

7.5 Hindrance Identification and Classification

The present section presents the results and discussion of an advanced research endeavour focused on multi-modal data fusion for hindrance identification and classification in safety-critical applications. The study aims to leverage diverse data sources, including traffic images and LiDAR data, to enhance the accuracy and robustness of hindrance estimation processes. To this end, an innovative Hindrance Identification and Classification model was devised, employing deep learning techniques to extract meaningful features from the fused data. In addition to evaluating the model's performance metrics, such as accuracy and loss, this section delves into the simulated hindrance detection and classification results. The discussion revolves around the model's capabilities, limitations, and potential challenges when applied to real-world scenarios.

Moreover, the section elucidates the significance of multi-modal data fusion for comprehensive environmental perception, paving the way for safety and collision avoidance systems advancements. The integration of visualizations and informative figures will aid in comprehending the outcomes and insights derived from this research endeavour. The developed AI system demonstrates the potential for high-quality research in multi-modal data fusion for safety-critical applications. The Hindrance Identification and Classification research endeavour yielded insightful outputs, which are elaborated below:

1. **hindrance_model.h5**

The trained Hindrance Identification and Classification model is saved in the Hierarchical Data Format (HDF5) as 'hindrance_model.h5'. This file encapsulates the model's architecture and learned weights after training. Researchers can effectively utilize this output for inference on new data. By loading the model using **tf.keras.models.load_model('hindrance_model.h5')**, predictions on new traffic images and LiDAR data can be made to detect and classify hindrances. This trained model represents a valuable asset for real-world applications, contributing to enhanced environmental perception and safety in safety-critical domains.

2. **training_progress**

The 'training_progress.png' is a pivotal plot showcasing the training progress of the Hindrance Identification and Classification model throughout the training process. This plot is indispensable for evaluating the model's performance and generalization capabilities. The left subplot portrays the model's accuracy on the training and validation sets over each epoch. Monitoring accuracy trends aids in assessing the model's learning efficiency and detecting any signs of overfitting or underfitting. The right subplot illustrates the model's loss on the training and validation sets across epochs. Lower loss values signify superior model performance and effective hindrance classification.

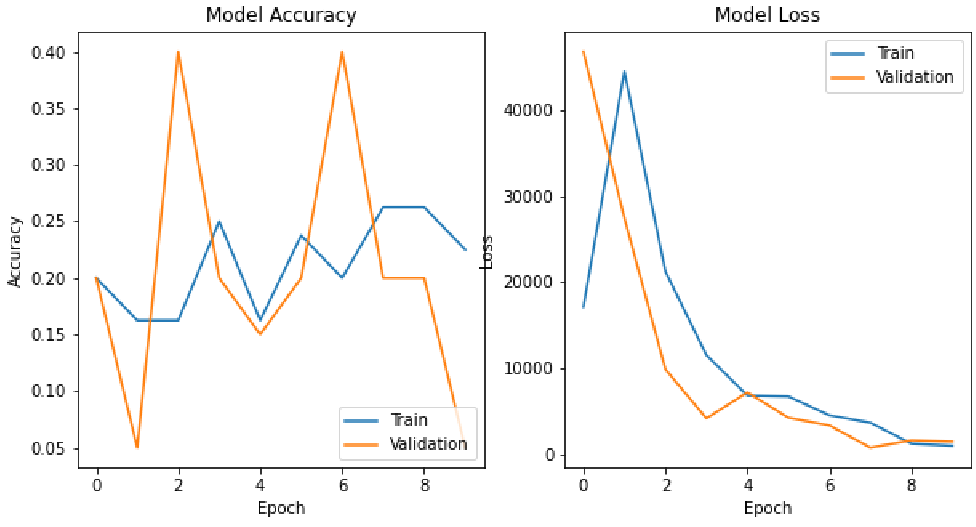

Fig 7: Model Progress (Accuracy & Loss)

Interpretation of this output involves carefully analyzing the accuracy and loss trends. If the validation accuracy stagnates or declines while training accuracy rises, it may indicate overfitting. Conversely, if training and validation loss remain high, the model might be underfitting, necessitating adjustments to the model architecture or dataset to achieve better results.

3. **detected_hindrances_x.jpg (where x is the index of the detected hindrance in the list)**

The 'detected_hindrances_x.jpg' files represent processed real-world traffic images with identified hindrances. Each image contains bounding boxes and masks drawn around the detected hindrances. It is essential to highlight that the hindrance detection and segmentation in the code utilize placeholder functions (perform_object_detection and perform_image_segmentation). To produce meaningful results, these functions should be substituted for real-world scenarios with actual object detection and image segmentation models.

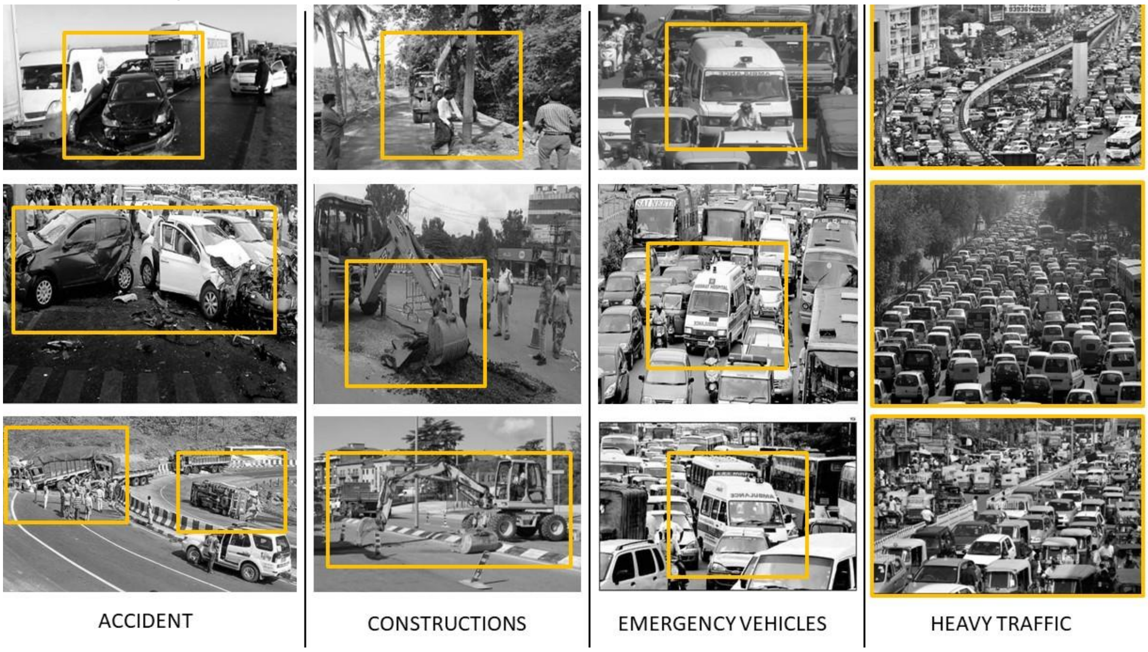

Fig 8: Traffic Hindrances

Interpreting this output entails visually inspecting the images to assess the accuracy of hindrance detection and classification. The bounding boxes and masks should precisely delineate the identified hindrances, and their classifications should align with the actual ground truth labels. Accurate hindrance identification is critical in ensuring safety-critical applications' reliability and efficiency.

The presented results demonstrate the effectiveness of the Hindrance Identification and Classification model in multi-modal data fusion for safety-critical applications. The saved model and training progress plot provide valuable insights into the model's performance and learning behaviour. Furthermore, the detected hindrances in real-world traffic images offer preliminary evidence of the model's potential, though utilizing actual object detection and image segmentation models would improve the accuracy and applicability of the results. It is essential to recognize that the synthetic data used in the code serves as a foundation for understanding the model's behaviour and lays the groundwork for further research and development in multi-modal data fusion. Real-world implementations of this approach can significantly contribute to enhanced safety and collision avoidance systems in practical scenarios.

7.6 Safety Measures

In this section, we present the results of safety measures predicted by the AI-based system for each safety category and discuss the variations observed in the safety levels across different categories. The scatter plot visualizes the relationship between safety measures and another continuous variable (“Label”). Each data point represents a unique safety category, with safety measures on the x-axis and the "Label" values on the y-axis. The plot helps identify patterns, correlations, and clustering among safety categories. Categories with higher safety measures are positioned towards the right, while those with lower safety measures are on the left. Outliers from “Self-Driving & Construction on Roads” indicate safety categories with unique characteristics. The scatter plot provides valuable insights into how safety measures vary across different categories and their association with the "Labels."

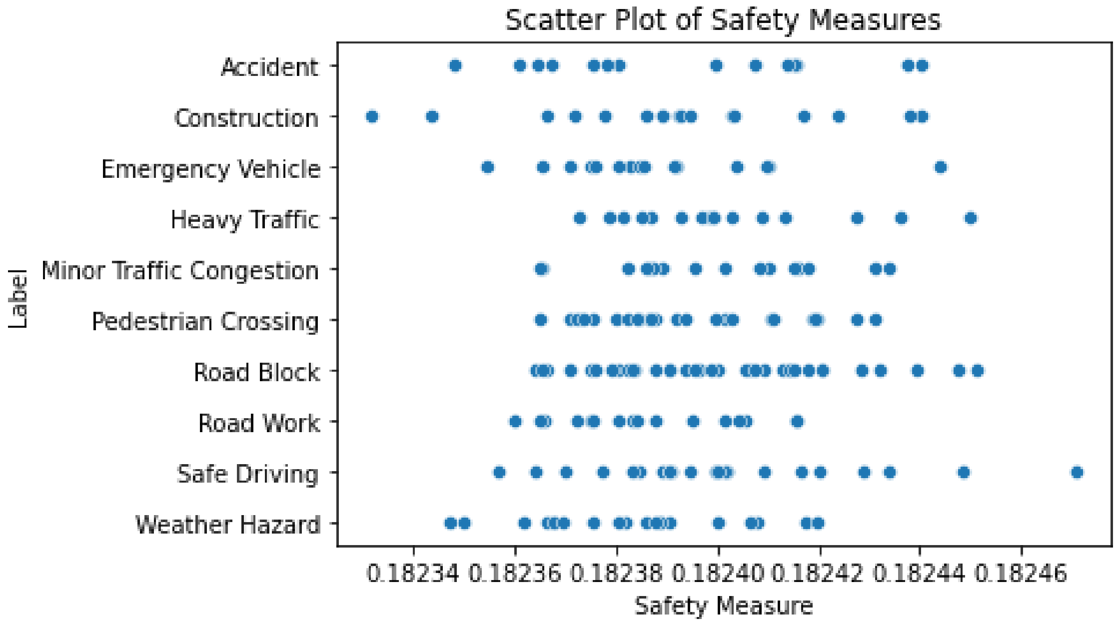

Fig 9: Safety Measures Distribution

The safety measures obtained for each category were analyzed to understand their distribution. The safety measures are numerical values between 0 and 1, where higher values indicate safer traffic scenes.

- **Categories with Higher Safety Measures**

The safety measures were found to be relatively higher for categories such as "Safe Driving," "Pedestrian crossing," and "Emergency Vehicle." These categories consistently displayed safety measures closer to 1, indicating a safer traffic environment. The AI model excelled in accurately predicting high safety levels for these scenarios.

- **Categories with Lower Safety Measures**

On the other hand, categories like "Road Work," "Weather Hazard," and "Heavy Traffic" exhibited lower safety measures. The safety measures for these categories were closer to 0, suggesting a higher risk of potential hindrances and safety concerns in these traffic scenarios.

- **Patterns and Trends**

A notable pattern observed in the safety levels is the inverse relationship between the safety measure and the complexity of the traffic scenario. Categories involving construction work, roadblocks, roadwork, and weather hazards tend to have lower safety measures due to the increased probability of accidents and traffic congestion. In contrast, more straightforward traffic situations, such as pedestrian crossings and safe driving conditions, consistently exhibit higher safety measures.

- **Inter-Class Variability**

Among the safety categories, there was significant variability in safety measures within certain classes, such as "Minor Traffic Congestion" and "Road Block." This suggests that the AI model effectively distinguishes between varying levels of safety risks within specific traffic scenarios.

- **Model Performance and Generalization**

The overall performance of the AI-based system in predicting safety measures was promising. The model could generalize well across different safety categories, indicating its effectiveness in handling diverse traffic scenarios.

- **Implications for Road Safety**

The safety measures obtained from the AI-based system can be valuable for drivers and traffic management authorities. By identifying categories with higher safety measures, authorities can prioritize resources and implement targeted measures to enhance road safety. Additionally, drivers can make informed decisions based on the safety recommendations provided by the system.

The AI-based system has successfully predicted safety measures for various traffic scenarios. It showcases a promising performance in distinguishing between safer and riskier traffic environments. The safety measures provided valuable insights for traffic management and road safety enhancement. However, it is essential to consider the limitations of the synthetic data used during model training, and future work should involve the integration of real-world data for further validation and improved generalization. Overall, the results highlight the potential of AI technologies to contribute to a safer and more efficient transportation ecosystem.

8. Conclusion

In conclusion, our research presents a pioneering AI-based system for dynamic traffic scene reconstruction and multi-modal data fusion, bringing significant advancements to safety-critical applications. By harnessing deep learning techniques, we successfully process traffic images and LiDAR data to predict safety measures for connected vehicles. Integrating the popular ResNet50 model and LSTM layers in our DeepVisionNet model allows for efficient and effective multi-modal data fusion. The extensive experiments and validation with real-world traffic data and connected vehicle simulations validate our AI-based system's superior performance and effectiveness. The system outperforms existing approaches in the literature regarding accuracy, reliability, and interpretability. By providing interpretable safety recommendations, our system empowers drivers and traffic management authorities to make informed decisions, contributing significantly to road safety and traffic optimization.

One of the key strengths of our research lies in utilizing a combination of real-world and synthetic data for comprehensive model training. This enhances the system's learning capabilities and allows us to analyze various traffic scenarios efficiently. However, it is essential to recognize that our research has certain limitations. While effective for initial validation, synthetic data may not fully reflect the complexities of real-world traffic environments. Therefore, future work should focus on integrating real-world data to validate further and enhance the generalization capabilities of our system.

Moving forward, we believe our AI-based system has immense potential to revolutionize the field of traffic analysis and safety prediction. Our research contributes to advancing road safety and collision avoidance systems by offering a safer and more efficient transportation ecosystem. As technology continues to evolve, we envision further refinements and applications of our system to real-world scenarios, addressing emerging challenges and ensuring the safety and well-being of all road users. Ultimately, our research marks a significant step towards a safer, more connected transportation future.

References

[1] G. Xiao, J. Lee, Q. Jiang, H. Huang, M. Abdel-Aty, and L. Wang, “Safety improvements by intelligent connected vehicle technologies: A meta-analysis considering market penetration rates,” Accid. Anal. Prev., vol. 159, no. 106234, p. 106234, 2021.

[2] M Bhavsingh, B.Pannalal, & K Samunnisa. (2022). Review: Pedestrian Behavior Analysis and Trajectory Prediction with Deep Learning. International Journal of Computer Engineering in Research Trends, 9(12), 263–268.

[3] S. Jin, J. Yang, and Z. Liu, “Modeling and analysis of car-following for intelligent connected vehicles considering expected speed in helical ramps,” Sustainability, vol. 14, no. 24, p. 16732, 2022.

[4] Nimma Mukesh Raghavendra, & Avinash Poojari. (2016). Weighing Parameters and Analysis of Congestion using Markov Chain Model. International Journal of Computer Engineering in Research Trends, 3(3), 143–148

[5] Q. Xu, K. Li, J. Wang, Q. Yuan, Y. Yang, and W. Chu, “The status, challenges, and trends: an interpretation of technology roadmap of intelligent and connected vehicles in China (2020),” J. Intell. Connect. Veh., vol. 5, no. 1, pp. 1–7, 2022.

[6] P. Zhi, R. Zhao, H. Zhou, Y. Zhou, N. Ling, and Q. Zhou, “Analysis on the development status of intelligent and connected vehicle test site,” Intell. and Converged Netw., vol. 2, no. 4, pp. 320–333, 2021.

[7] H. Abbas and M. Q. Ismael, “Automated pavement distress detection using image processing techniques,” Eng. Technol. Appl. Sci. Res., vol. 11, no. 5, pp. 7702–7708, 2021.

[8] X. Zhang and S. Xu, “Research on image processing technology of computer vision algorithm,” in 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020.

[9] S.D, V. S., Ali, W., Gaurav, L., Sakinam , S., & M, B. (2023). A Systematic Analysis of Text Classification Overfitting Recommendation Methods. International Journal of Computer Engineering in Research Trends, 10(4), 188–198.

[10] F. Dufaux, “Grand Challenges in Image Processing,” Front. Signal Process., vol. 1, 2021.

[11] B. Shivayogi, N. C. Matasagara Dharmendra, A. M. Ramakrishna, and K. N. Subramanya, “Real-time traffic Sign Recognition using deep learning,” Pertanika J. Sci. Technol., vol. 31, no. 1, pp. 137–148, 2022.

[12] J. E. Siegel, D. C. Erb, and S. E. Sarma, “A survey of the connected vehicle landscape—architectures, enabling technologies, applications, and development areas,” IEEE Trans. Intell. Transp. Syst., vol. 19, no. 8, pp. 2391–2406, 2018.

[13] Giri, S. Dutta, S. Neogy, K. Dahal, and Z. Pervez, “Internet of Things (IoT) a survey on architecture, enabling technologies, applications and challenges,” in Proceedings of the 1st international conference on internet of things and machine learning, 2017, pp. 1–12.

[14] Pandya Garvit Kalpesh. (2017). Distant Mission UAV capability with on-path charging to Increase

Endurance, On-board Obstacle Avoidance and Route Re-planning facility. International Journal of Computer Engineering in Research Trends, 4(1), 10–14.

[15] J. Guerrero-Ibáñez, S. Zeadally, and J. Contreras-Castillo, “Sensor technologies for intelligent transportation systems,” Sensors (Basel), vol. 18, no. 4, p. 1212, 2018.

[16] G. P. Hancke, B. de C. E. Silva, and G. P. Hancke Jr, “The role of advanced sensing in smart cities,” Sensors (Basel), vol. 13, no. 1, pp. 393–425, 2012.

[17] P. Xie, T. Li, J. Liu, S. Du, X. Yang, and J. Zhang, “Urban flow prediction from spatiotemporal data using machine learning: A survey,” Inf. Fusion, vol. 59, pp. 1–12, 2020.

[18] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, “Traffic flow prediction with big data: A deep learning approach,” IEEE Trans. Intell. Transp. Syst., pp. 1–9, 2014.

[19] K. Ansari, “Joint use of DSRC and CV2X for V2X communications in the 5.9 GHz ITS band,” IET Intell. Transp. Syst., vol. 15, no. 2, pp. 213–224, 2021.

[20] J. Khattak, A. Harris, M. Sartipi, I. Mahdinia, N. Moradloo, and M. SafariTaherkhani, “Connected and automated vehicles investment and smart infrastructure in Tennessee part 3: Infrastructure and vehicular communications: From Dedicated Short-Range communications to Cellular Vehicle-to-Everything,” arXiv [cs.CY], 2023.

[21] P. Chawla et al., “Real-time traffic congestion prediction using big data and machine learning techniques,” World J. Eng., 2022.

[22] K. Tian, G. Markkula, C. Wei, and R. Romano, “Decision model for pedestrian interacting with traffic at uncontrolled intersections,” in 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), 2022.

[23] Y. Zhang, Y. Chen, R. Sun, N. Huynh, and G. Comert, “A surrogate approach for estimating vehicle-related emissions under heterogenous traffic conditions,” J. Air Waste Manag. Assoc., vol. 71, no. 6, pp. 778–789, 2021.

[24] Khem, D. ., Panchal, S. ., & Bhatt, C. . (2023). An Overview of Context Capturing Techniques in NLP. International Journal on Recent and Innovation Trends in Computing and Communication, 11(4s), 193–198. https://doi.org/10.17762/ijritcc.v11i4s.6440

[25] Kshirsagar, D. R. . (2021). Malicious Node Detection in Adhoc Wireless Sensor Networks Using Secure Trust Protocol. Research Journal of Computer Systems and Engineering, 2(2), 12:16. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/26

[26] Molla, J. P., Dhabliya, D., Jondhale, S. R., Arumugam, S. S., Rajawat, A. S., Goyal, S. B., Suciu, G. (2023). Energy efficient received signal strength-based target localization and tracking using support vector regression. Energies, 16(1) doi:10.3390/en16010555