2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) | 979-8-3315-1250-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSCDS65426.2025.11167123

Integrated Traffic Incident Classification using SegFormer and Faster R-CNN: A Multi-Stage Approach for Enhanced Detection and Analysis

Sankar Ganesh Karuppasamy
Department of CSE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
prof.sankarganesh@gmail.com

Sajeev Ram Arumugam
Department of CSE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
imsajeev@gmail.com

Sheela Gowr P
Department of CSE
Vels Institute of Science, Technology & Advanced Studies
Chennai, India
psheelagowr85@gmail.com

Divya Muralitharan
Centre for Artificial Intelligence
Department of Artificial Intelligence and Data Science
Rajalakshmi Institute of Technology
Chennai, India
divyamuralitharan@gmail.com

Tamilselvi S
Department of Networking and Communications
SRM Institute of science and Technology
Chennai, India
tamilselvy@gmail.com

Gowthami K
Department of Artificial Intelligence and Data Science
Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College
Chennai, India
gowthami2730@gmail.com

***Abstract*—A traffic incident refers to any incidence or situation that interrupts the regular movement of car traffic or presents a hazard to individuals using the road. These incidents comprise a wide range of accidents involving vehicles, including rear-end crashes, head-on accidents, or side-impact crashes, which can lead to trauma, destruction, or property damage. This also includes vehicle failures caused by mechanical failures, such as engine problems or flat tires, which render cars immobile. The proposed paper has a state-of-the-art method for categorizing traffic incidents, which combines two sophisticated computer vision models: SegFormer and Faster R-CNN. SegFormer, an advanced semantic segmentation feature, creates detailed pixel-by-pixel classification maps of traffic scenes. This feature facilitates the clear distinction of various traffic actions and components. The final segmentation result offers a broad understanding of the spatial arrangement and corresponding information within the image. In addition, a Faster R-CNN for object detection has outstanding performance in detecting and categorizing distinct items such as automobiles and pedestrians. In the segmented regions, faster R-CNN improves the detection and classification of specific traffic-related items. The efficiency of the proposed approach is measured using a collection of traffic images, showing improved ability in identifying and classifying several forms of traffic hazards. The findings demonstrate that this hybrid method greatly surpasses conventional single-model techniques, providing a more robust and complete solution for the study and control of traffic incidents.**

Keywords—Traffic Incident, CNN, YOLO, SVM, SegFormer, Faster R-CNN, SVM, RNN, LSTM.

I. INTRODUCTION

Traffic management and conserving road safety have grown more problematic as metropolitan areas continue to grow and as the number of vehicles on the road increases. Traffic events, such as accidents, vehicle breakdowns, and obstruction, prevent traffic movement, resulting in delays, economic losses, and increased risks to public safety [1]. The World Health Organization approximates that traffic accidents take the lives of almost 1.3 million people each year, ranking them among the leading causes of death worldwide [2]. Statistics from the National Highway Traffic Safety Administration [3] in the United States show that more than 6 million road accidents take place per year, resulting in around 40,000 deaths and countless injuries. Moreover, occurrences like accidents and car breakdowns worsen traffic obstruction, resulting in an annual cost of more than $87 billion to the U.S. economy through decreased productivity and unnecessary fuel consumption. The growth in the global number of vehicles worsens these problems, especially in metropolitan regions where traffic-related instabilities are most rigorous. These data points highlight the immediate requirement for more efficient traffic incident detection and classification technologies to reduce their negative impact on road safety and fuel economy. Conventional traffic management systems [4] typically depend on manual monitoring and limited sensor data, which can be slow, ineffective, and prone to malfunctions. The development of complex computer vision techniques has the potential to optimize and automate traffic incident identification and classification, allowing for quick and detailed reactions. Integrating advanced models such as SegFormer [5] for semantic segmentation and Faster R-CNN [6] for object recognition is an essential approach for addressing the necessities of traffic incident management. Existing approaches for traffic event identification and categorization often rely on manual surveillance and inadequate sensor systems, resulting in delayed reactions and inaccurate results. Such conventional methods face difficulties in delivering the level of specificity and accuracy needed to efficiently handle complicated traffic situations, especially in real time. An essential condition exists for more complicated and automated systems capable of accurately identifying and categorizing different forms of traffic accidents, therefore allowing more effective traffic control and enhanced road safety. The objectives of the proposed paper are as follows:

- A novel approach is proposed to improve traffic event identification and categorization by encompassing SegFormer and Faster R-CNN models. Employing SegFormer's semantic segmentation to achieve exact pixel-level categorization of traffic scenes.
- The objective is to utilize Faster R-CNN for accurate object recognition and categorization in traffic scenes, with a specific prominence on automobiles and other relevant objects.

- The goal is to test the proposed model's performance on actual traffic statistics and compare its effectiveness with conventional single-model methods.
- The aim is to enhance the precision and efficiency of traffic incident detection, resulting in improved traffic management and faster rescue efforts.

The proposed work introduces a new method for classifying traffic incidents by combining two advanced computer vision models, namely SegFormer and Faster R-CNN. The main role is to combine the advantages of various models to develop a system that is robust, with the ability to understand circumstances in incredible detail and accurately identify objects. The proposed method not only improves the precision and efficiency of traffic event detection, but it also enables the classification of different types of incidents in complex traffic situations.

II. Related Work

The detection and classification of traffic incidents have been the subject of extensive study over the years. The need to enhance traffic management systems, reduce congestion, and boost road safety has motivated this research. Historical methods for detecting traffic incidents mostly depended on infrastructure-based devices, such as loop detectors, cameras, and radar systems. These approaches' dependence on stationary installations and their inability to provide real-time, comprehensive incident analysis have limited their effectiveness. People commonly employ closed-circuit television (CCTV) cameras for real-time traffic monitoring. Despite its effectiveness in delivering visual data, conventional video surveillance requires continuous human supervision, making it labour-intensive and prone to human error. Recent advancements in automated video analysis systems have aimed at addressing this problem.

Advanced machine learning techniques such as Support Vector Machines (SVM) and decision trees [7] were used to identify traffic incidents through analyzing patterns of traffic flow and data collected from sensors. These models were capable of categorizing traffic conditions using predetermined characteristics, but their dependence on manually designed features and limited capacity to generalize well across various traffic situations frequently constrained them. Implementing automated traffic incident detection is critical for ensuring road safety and optimizing traffic management. However, the task of creating efficient models becomes difficult when confronted with inadequate and limited datasets. Hence, obtaining broad annotated datasets for the purpose of training machine learning models presents difficulties and involves significant costs, particularly in instances in which incidences are few or diversified. Using techniques for adding to data, like oversampling the minority class or making fake data with tools like SMOTE (Synthetic Minority Oversampling Technique), can help balance a dataset. Moreover, advanced techniques, such as ensemble methods or transfer learning, can be used to enhance the performance of the model on limited and imbalanced datasets. Implementing these techniques improves the model's capacity for accurately recognizing traffic accidents, even under difficult circumstances, which results in more reliable and efficient traffic monitoring systems [8].

Ensemble learning techniques, like Random Forests and Gradient Boosting, make event detection models more resilient by combining several decision trees to cut down on overfitting and improve accuracy. Convolutional Neural Networks (CNNs) have improved traffic incident detection by enabling the automatic learning of characteristics from extensive datasets. Convolutional Neural Networks (CNNs) have shown remarkable effectiveness in the analysis of traffic images and videos, resulting in significant improvements in the identification of traffic incidents. The integration of Convolutional Neural Networks (CNNs) with a genetic algorithm in a deep learning model presents a highly promising method for reducing serious traffic accidents. CNNs' excellent performance in processing and analyzing visual data makes them well-suited for identifying patterns and characteristics in traffic images or video. The model may adjust the architecture and hyperparameters of the CNN by including a genetic algorithm, ensuring that the network is optimized for its best performance. This integration enables the system to acquire knowledge from large volumes of traffic data, revealing subtle indications that occur before major incidents, such as driver conduct, vehicle paths, and environmental circumstances [9].

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks [10] have been used to analyze sequential data, such as traffic videos, to capture temporal dependencies. This enhanced approach improves the detection of incidents that occur over time, such as the increasing growth in traffic congestion. YOLO (You Only Look Once) [11], along with comparable object identification algorithms, have had an important impact on the field of traffic incident detection. These models provide the capability to detect and categorize many items simultaneously in real-time, making them very suitable for the identification of vehicles, pedestrians, and other components involved in traffic incidents. Semantic segmentation models such as UNet [12] have been employed to give pixel-level categorization of traffic scenes, a critical aspect in comprehending the context of traffic hazards.

The proposed work attempts to bring in some important gaps and constraints in existing research, despite significant advances in traffic incident detection using computer vision and machine learning. An underlying drawback lies in the management of imbalanced datasets, which are characterized by rare accidents and breakdowns in comparison to conventional traffic flow. The existing approaches face difficulties with properly capturing these important occurrences, resulting in reduced detection accuracy and a higher number of false negatives [7]. Moreover, existing approaches frequently lack the capacity to efficiently manage real-time data in dynamic traffic environments, where factors such as weather conditions, lighting fluctuations, and road layouts can significantly affect model performance. Moreover, the dependence of several existing methodologies on manually designed characteristics or outdated detecting techniques limits their development and modification to other geographical areas or traffic situations. This work aims to close these gaps and enhance the precision and resilience of traffic incident detection systems by combining deep learning models such as SegFormer and Faster R-CNN, real-time detection optimization, and the implementation of strategies to better handle imbalanced data.

III. METHODOLOGY

The proposed work integrates SegFormer and Faster R-CNN to create a robust system for real-time traffic incident detection and categorization. First, the SegFormer model will semantically divide traffic scenes and provide pixel-wise categorization maps based on road environment context and spatial information. This segmentation identifies vehicles, pedestrians, and road infrastructure. SegFormer's output will be used with Faster R-CNN's object detection to recognize and classify traffic events' specific objects, such as vehicles in accidents or obstructing traffic. Data augmentation will address imbalanced datasets. The integrated model will be trained and tested on available datasets using accuracy, precision, recall, and F1-score to determine its capacity to detect important traffic incidents.

A. Data Collection

The Cityscapes dataset [8] is ideal for identifying traffic incident situations and other traffic accidents in urban areas since it offers detailed observations for vehicles, pedestrians, road signs, and road surfaces. The dataset has a wide set of images in different lighting and weather conditions, which helps the process apply across conditions and detect problems in real time throughout traffic circumstances. The Cityscapes collection has 5,000 well-annotated images. The images from 50 cities range a diversity of weather, seasons, and urban landscapes, making them useful for urban traffic analysis and identifying incidents.

B. Augmentation

Data augmentation approaches can increase the robustness of training data for Cityscapes dataset traffic incident detection models. Random cropping and scaling [9] impressionist camera angles and distances, allowing the model to generalize across perspectives. In bad weather, Gaussian noise [10] may replicate camera sensor activity, and blurring can approximate vehicle motion blur. First, the image is randomly cropped to simulate multiple viewpoints or zoomed-in portions. Randomly scaling the image up or down after cropping helps the model adjust to different object sizes and camera distances. Finally, Gaussian noise improves generalization by enabling the model to manage noisy situations like low-light or cloudy images. People often obstruct vehicles in traffic incidents. These augmentations make the model more adaptive, enhancing its incident detection.

C. Model Architecture

The proposed work employs a hybrid architecture that integrates SegFormer and Faster R-CNN to detect and categorize traffic incidents. Each model has a unique function in analyzing the visual input, allowing precise semantic segmentation and object detection for immediate identification of traffic incidents. SegFormer is a conceptual segmentation model that utilizes a hierarchical transformer-based architecture. It employs multi-head self-attention to analyze image areas, capturing both long-range dependencies and multi-scale context. Fig. 1. shows the working of the proposed architecture.

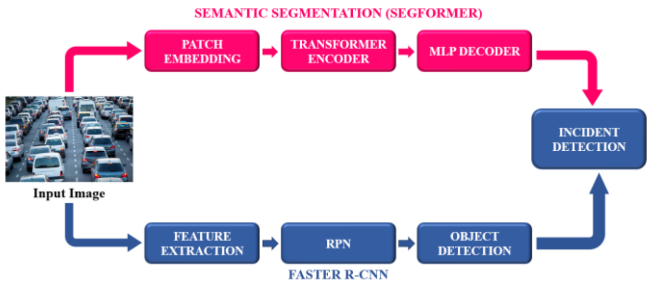

Fig.1. Proposed Architecture

The input image $I \in R^{N\times D}$, where $N = \frac{H}{P}\times\frac{W}{P}$ is the number of patches and D is the embedding dimension.

$$I' = PatchEmbed(I) \qquad (1)$$

Where $I' \in R^{\frac{H}{P}\times\frac{W}{P}\times D}$

In multi-head special attention, each transform block for input $X \in R^{N\times D}$, the query Q, the key K, and the value V matrices are projected:

$$Q = XW_Q \qquad (2)$$

$$K = XW_K \qquad (3)$$

$$V = XW_V \qquad (4)$$

The attention mechanisms are computed as

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{D^k}}\right)V \qquad (5)$$

Where D^k is the dimension of the key vectors, andW_Q, W_K and W_V are the learned projection matrices. Following the attention layer, a feed-forward network analyzes the output data, which comprises two linear layers with a non-linear activation function (GELU):

$$FFN(X) = W_2\sigma(W_1X + b_1) + b_2 \quad (6)$$

Where W_1 and W_2 are the weights, b_1 and b_2 are biases, and σ is the GELU activation. At each stage, the hierarchical transformer encoder reduces the resolution of the feature maps X_1, X_2, X_3, X_4, resulting in the generation of multi-scale features. Their combination yields comprehensive contextual information for segmentation. The final segmentation map is predicted by employing a lightweight multi-layer perceptron (MLP) decoder to process the multi-scale information.

$$Segmentation\ Map = MLP(X1, X2, X3, X4) \qquad (7)$$

The result is a categorization of the image at the pixel level into different traffic-related clusters, such as cars, pedestrians, road signs, and so on. Combining region proposal networks (RPN) with a classification and regression network based on convolutional neural networks (CNNs), the Faster R-CNN model is a way to find objects. The input image $I \in R^{H\times W\times 3}$ is initially processed by a convolutional neural network (CNN), such as ResNet, in order to generate a feature map $F \in R^{H_f\times W_f\times 3}$, where H_f and W_f are the spatial dimensions of the feature map and C is the cluster of channels.

$$F = CNN(I) \qquad (8)$$

The Region Proposal Network (RPN) generates candidate object proposals by iteratively moving a tiny window across the feature map F. The RPN operation generates a set of anchor boxes with related classification scores. The classification score denotes the probability of an object being present at the i^{th} anchor. Statistical bounding box regression modifies the anchor boxes to optimize their fit with the given objects:

$$B_i = Anchor_i + \triangle B_i \tag{9}$$

Where $\triangle B_i$ is the predicted adjustment for the i^{th} anchor box. The ROI pooling technique extracts fixed-size features from the feature map for each area proposal. The aggregated features are subsequently fed into fully connected layers to perform classification and bounding box regression. Object Classification and Regression involves classifying each proposed region into one of the object categories, such as vehicle or pedestrian, and predicting a refinement of the bounding box.

$$\widehat{B_\iota} = FC(ROI\ Pooling(F, B_i) \tag{10}$$

$$\widehat{C_\iota} = Softmax(FC(ROI\ Pooling(F, B_i)) \tag{11}$$

Where $\widehat{B_\iota}$ is the refined bounding box and $\widehat{C_\iota}$ is the predicted class label. This hybrid architecture combines accurate object detection with fine-grained scene knowledge to improve traffic incident detection and enable robust identification of various traffic occurrences. SegFormer generates the segmentation map, and Faster R-CNN produces the bounding boxes and labels. In contrast to Faster R-CNN, which offers object-level data such as bounding boxes around vehicles or pedestrians, the segmentation map provides pixel-level context about the image (e.g., road sections, automobiles). The system looks for abnormalities or inconsistencies that could point to a traffic incident by analyzing the segmentation map and the things it has observed. The proposed architecture takes the best parts of semantic segmentation (using SegFormer) and object identification (using Faster R-CNN) and combines them to find traffic events more accurately by looking at both the whole image and individual objects. The proposed architecture takes the best parts of semantic segmentation (using SegFormer) and object identification (using Faster R-CNN) and combines them to find traffic events more accurately by looking at both the whole image and individual objects. Fig. 2. Shows the sample output for traffic incident classification.

Fig.2 Traffic Incident Classification [19]

D. Need for Faster R-CNN

The Faster R-CNN was chosen for object detection due to its superior performance in identifying multiple objects under occlusion, which is common in real-world traffic incidents. Unlike single-shot detectors like YOLO or SSD, Faster R-CNN uses a Region Proposal Network (RPN) to generate candidate object regions before classification and regression. This allows for more accurate detection, especially in complex or congested scenes where vehicles and pedestrians overlap or are partially occluded.

IV. Results and Discussions

A series of studies using both quantitative and qualitative measures evaluated the efficiency of the proposed traffic incident detection system. An evaluation of the system was conducted using a benchmark dataset, which included SegFormer for segmentation and Faster R-CNN for object detection. The accuracy was calculated by examining its ability to correctly identify and classify incidents in traffic situations. Precision reduces false positives, which refer to the wrong identification of traffic incidents. Recall reduces false negatives, which refer to unrecognized incidents. Experiencing a high recall rate is impotant in traffic incident identification, as the incapability to identify actual incidents, such as accidents, could result in safety threats or interruptions in emergency response.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

$$Precision = \frac{TP}{TP+FP} \tag{13}$$

$$Recall = \frac{TP}{TP+FN} \tag{14}$$

A model makes a true positive (TP) prediction when it correctly recognizes a traffic incident. False positives (FP) occur when the model incorrectly calculates the presence of an incident that is not actually present. True Negative (TN) indicates the cases in which the model accurately calculates the absence of a traffic incidence. When the model accurately identifies the absence of a positive instance, it achieves a true negative. False negatives (FN) are positive instances, such as traffic incidents, that the model failed to correctly identify. By combining SegFormer for semantic segmentation and Faster R-CNN for object detection, the proposed hybrid method offers big improvements and new ways of doing things when it comes to finding traffic accidents. Some traffic incident detection systems exclusively rely on convolutional neural networks (CNNs) [11] for the purpose of object detection and categorization. Typically, these approaches prioritize the detection of automobiles, pedestrians, and traffic aspects, but they generally lack an extensive understanding of the surrounding environment. Traditional convolutional neural networks (CNNs) [12] primarily concentrate on bounding boxes and object detection. However, the proposed method integrates both segmentation and object detection, therefore enhancing the contextual comprehension of traffic situations.

Designed for pixel-wise categorization and visual comprehension, unsupervised segmentation models like UNet [13] or PSPNet [14] lack the ability to distinguish objects. The combination of Faster R-CNN with SegFormer segmentation allows for both pixel-level perception and object-level detection. This allows more accurate identification of traffic incidents, mostly in vital traffic environments. Although UNet segmentation models offer definite pixel-wise segmentation, it do not show the same level of object detection and classification as Faster R-CNN. YOLO [15] is a high-speed algorithm that holds both segmentation (SegFormer) and object detection (Faster R-CNN) to boost the accuracy of incident detection. This feature allows the system to classify incidents even in complex situations. The YOLO algorithm focuses on speed and manages smaller section information. The proposed method is more effective in identifying complex traffic incidents as it combines pixel-wise context with object detection.

A high-performance computer environment is required in the experimental setup for the proposed traffic incident detection system to effectively manage the complexity of the deep learning methodologies. The equipment comprises an Intel Xeon 32-core processor, 128GB RAM, and 2TB SSD storage designed for efficient processing of huge databases and rapid data retrieval. Software-wise, the system utilizes deep learning models implemented in PyTorch 1.9.1, CUDA 11.1, and cuDNN 8.0 to enhance GPU-based efficiency. A hybrid model for semantic segmentation and object identification is trained using a dataset of 3500 images from the Cityscapes collection. The images are categorized into training, testing, and validation sets within a 7:2:1 ratio. In all, a total of 2450 images are utilized for training, 700 images for testing, and 350 images for validation. Random cropping, scaling, and Gaussian noise are employed as data augmentation methods to enhance the resilience of the model. The Adam optimizer with early halting is used to train the models, and their performance is evaluated using accuracy, precision, and recall evaluation criteria. An evaluation was conducted on the test dataset to evaluate the performance of the hybrid system that combines SegFormer for semantic segmentation and Faster R-CNN for object detection. The evaluation was carried out using the following metrics: accuracy, precision, and recall with the existing methods. Table 1. Shows the comparative analysis of the proposd system with the existing approach.

TABLE I. COMPARATIVE ANALYSIS OF THE EXISTING APPROACH WITH THE PROPOSED WORK

Techniques	***Accuracy***	***Precision***	***Recall***
GAN and TSSAE [16]	90.62%	NA	NA
SASYNO-RF-RSKNN [17]	95.7%	96.7%	NA
YOLOv5 and DeepSORT [1]	98.9%	NA	NA
K-NN and DTW [18]	90%	NA	NA
Proposed Work	**98.7%**	**97.4%**	**93.6%**

([NA] The Performance metric was not calculated/Mentioned)

The training accuracy is dependent upon the particular dataset and experimental methodology employed in the proposed work. Training loss in models such as SegFormer and Faster R-CNN measures the difference of the model's predictions from the actual labels in the training data. When combined, these losses would be responsible for the entire training loss in such a hybrid model, representing both the segmentation and detection performance. A lower training loss implies that the model exhibits better acquisition of information from the training data and enhances its accuracy in prediction. Fig. 3. Shows the training accuracy and Fig. 4 training loss for the proposed work.

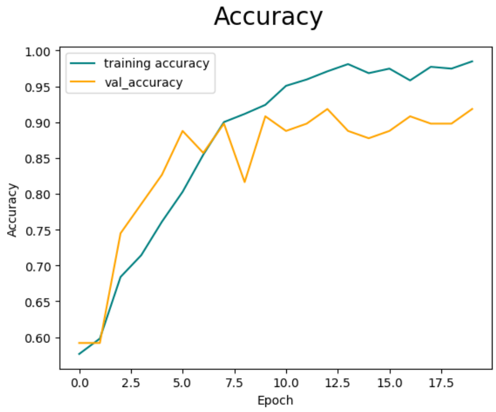

Fig.3. Training Accuracy

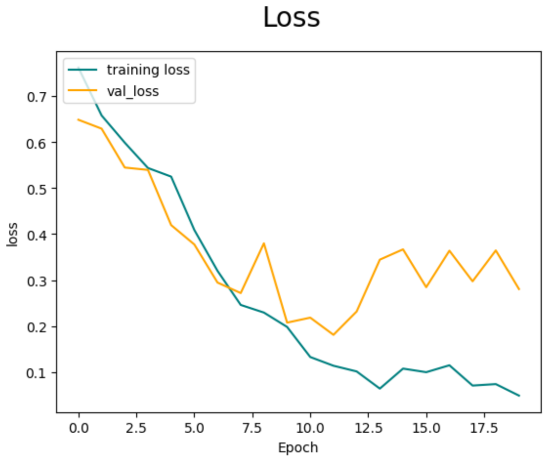

Fig.4. Training Loss

The hybrid technique is evaluated using key metrics like segmentation accuracy and object detection precision and recall. Effective integration seeks to optimize incident categorization by capitalizing on the advantages of both models, potentially improving overall performance in comparison to using each model independently. However, the issues include managing computational complexity, ensuring robustness in various traffic conditions, and determining limits in handling occlusions or overlapping incidents. Future research may investigate the optimization of the integration process, conduct experiments using supplementary data sources, and improve the model to effectively manage complex and diverse traffic circumstances.

Object Detection Enhancement Techniques:

- **Anchor Box Optimization**: Custom anchor box scales and aspect ratios were tuned
- **High-Resolution Feature Maps**: The backbone CNN used in Faster R-CNN was configured to retain higher resolution feature maps, enhancing detection accuracy of small or overlapping objects.
- **Improved ROI Pooling**: Adaptive ROI Align was used instead of basic ROI Pooling to preserve spatial alignment.

- **Hard Negative Mining**: Training involved focusing on difficult-to-classify background areas.
- **Augmentation Against Occlusion**: Simulating occluded and blurred conditions during training using Gaussian for better robustness.
- **Multi-scale Testing**: During inference, images were resized at different scales to capture objects of varying sizes, increasing detection sensitivity.

V. Conclusion

The proposed work demonstrates the potential for improving traffic incident classification by merging sophisticated segmentation and object identification models. By combining SegFormer's exact segmentation capabilities with Faster R-CNN's robust object identification, the proposed method provides a more complete and precise categorization of traffic incidents. The findings emphasize the potential of this hybrid approach in enhancing the effectiveness of traffic control applications. In order to make incident detection and classification even better in real-life traffic situations, future developments should focus on improving model performance, combining different data sources, and getting rid of certain problems. The proposed work attains a precision rate of 98.7%.

References

[1] Basheer Ahmed, Mohammed Imran, Rim Zaghdoud, Mohammed Salih Ahmed, Razan Sendi, Sarah Alsharif, Jomana Alabdulkarim, Bashayr Adnan Albin Saad, Reema Alsabt, Atta Rahman, and Gomathi Krishnasamy. "A real-time computer vision based approach to detection and classification of traffic incidents." *Big data and cognitive computing* 7, no. 1 (2023): 22.

[2] World Health Organization, "Road traffic injuries," *Fact Sheets*, [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries. Accessed: Sep. 03, 2024.

[3] A. W. Block and S. Walker, "2007 motor vehicle occupant safety survey: Driver education and graduated driver licensing," *Teen Driver Safety and Graduated Driver Licenses*, no. Dec., pp. 73–85, 2011.

[4] N. Nigam, D. P. Singh, and J. Choudhary, "A review of different components of the intelligent traffic management system (ITMS)," *Symmetry*, vol. 15, no. 3, p. 583, 2023.

[5] W. M. Elmessery, R. M. Elhaddad, R. M. Basha, M. R. Elfar, and A. E. Hassanien, "Semantic segmentation of microbial alterations based on SegFormer," *Frontiers in Plant Science*, vol. 15, p. 1352935, 2024.

[6] Y. Liu, S. Zhou, X. Liu, C. Hao, B. Fan, and J. Tian, "Unbiased Faster R-CNN for single-source domain generalized object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2024, pp. 28838–28847.

[7] C. Dong and N. Chang, "Overview of the identification of traffic accident-prone locations driven by big data," *Digital Transportation and Safety*, vol. 2, no. 1, pp. 67–76, 2023.

[8] S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, and S. Roth, "The Cityscapes dataset for semantic urban scene understanding," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016.

[9] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022.

[10] E. Goceri, "Evaluation of denoising techniques to remove speckle and Gaussian noise from dermoscopy images," *Computers in Biology and Medicine*, vol. 152, p. 106474, 2023.

[11] K. Wang, J. Gao, and X. Lei, "MTC: A multi-task model for encrypted network traffic classification based on transformer and 1D-CNN," *Intell. Autom. Soft Comput.*, vol. 37, no. 1, 2023.

[12] Z. Shi, N. Luktarhan, Y. Song, and G. Tian, "BFCN: A novel classification method of encrypted traffic based on BERT and CNN," *Electronics*, vol. 12, no. 3, p. 516, 2023.

[13] D. Muralitharan, S. Tamilselvi, J. D. Joann, S. R. Arumugam, S. G. Karuppasamy and C. S. Shylaja, "Exploring the Efficacy of U-Net-CNN Architecture for Automated Detection of Psoriatic Arthritis from Medical Imaging Data," *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, Erode, India, 2024, pp. 227-232, doi: 10.1109/ICSSAS64001.2024.10760526.

[14] X. Jie and R. Fan, "A road extraction method based on improved PSPNet," in *Proc. Int. Conf. Remote Sensing, Surveying, and Mapping (RSSM)*, SPIE, 2023, pp. 98–105.

[15] S. R. Arumugam, S. G. P, G. Ponnaian, M. K, S. G. Karuppasamy and D. Muralitharan, "Synergizing Generative Adversarial Networks and Pseudo-Labeling for Improved Knee Osteoarthritis Detection," *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, Coimbatore, India, 2024, pp. 1022-1030, doi: 10.1109/ICSCSS60660.2024.10625608.

[16] L. Li, Y. Lin, B. Du, F. Yang, and B. Ran, "Real-time traffic incident detection based on a hybrid deep learning model," *Transportmetrica A: Transp. Sci.*, vol. 18, no. 1, pp. 78–98, 2022.

[17] T. Xie, Q. Shang, and Y. Yu, "Automated traffic incident detection: Coping with imbalanced and small datasets," *IEEE Access*, vol. 10, pp. 35521–35540, 2022, doi: 10.1109/ACCESS.2022.3161835.

[18] P. Zißner, P. H. L. Rettore, B. P. Santos, J. F. Loevenich, and R. R. F. Lopes, "DataFITS: A heterogeneous data fusion framework for traffic and incident prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 10, pp. 11466–11478, 2023, doi: 10.1109/TITS.2023.3281752.

[19] https://ijirt.org/publishedpaper/IJIRT178816_PAPER.pdf.