Enhancing Road Safety: Road Friction Object Detection

Ravishankar S Bevinal
Computer Science and Engineering
KLE Technological University
Hubballi, India
ravishankarbevinal10@gmail.com

Sujatha C
Computer Science and Engineering
KLE Technological University
Hubballi, India
sujata_c@kletech.ac.in

Pratiksha S Tigadi
Computer Science and Engineering
KLE Technological University
Hubballi, India
pratikshatigadi@gmail.com

Yashasvi Kattimani
Computer Science and Engineering
KLE Technological University
Hubballi, India
kattimaniyashasvi@gmail.com

Gokul G Gadkar
Computer Science and Engineering
KLE Technological University
Hubballi, India
gokulgadkar8501@gmail.com

Shankar Gangishetty
IIIT Hyderabad
Hyderabad, India
shankar.gangisetty@ihub-data.iiit.ac.in

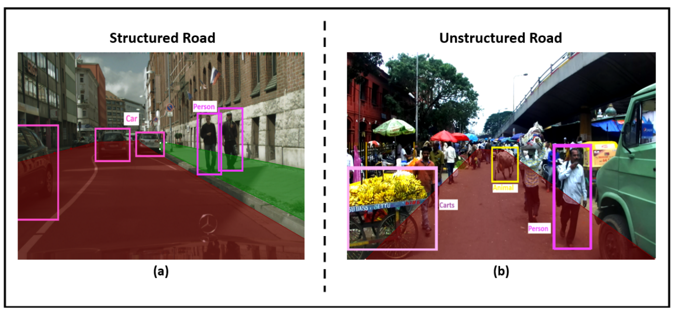

Fig. 1. (a) A snapshot of organized roads in Western countries, featuring distinct lane markings (*red*) and pedestrian walkways (*green*). The pedestrians are walking on the sidewalks, hence not causing hindrance to the flow of traffic. (b) An image capturing the chaotic essence of Indian streets, where unstructured roads lack defined lanes and organized pathways. Common disruptions to the flow of traffic - a street vendor with a fully loaded cart, a cow, pedestrians wandering leisurely on the street(*red*)due to the absence of sidewalks.

***Abstract*—Road safety in India presents a multifaceted challenge due to the dynamic nature of road environments. Diverse elements such as vehicles, pedestrians, and livestock, considered as road friction elements, frequently disrupt traffic flow and demand efficient management. Additionally, the roads often lack clear markings or divisions, making navigation difficult for Advanced Driver Assistance Systems (ADAS). To address these challenges, we propose a computer vision-based solution for detecting and classifying critical road friction elements on Indian roads. From the Indian Driving Dataset (IDD), which includes 34 annotated classes, we focus on seven main classes: animals, motorcycles, bicycles, trucks, cars, persons, and riders. Additionally, two new classes—carts and garbage—are introduced, making a total of nine classes considered as road friction elements. To locate these friction elements and enhance understanding of road environments, the road label is utilized for road segmentation. The YOLOv5 model handles the detection of friction elements, while road segmentation is managed by the U-Net model. Experimental results indicate a high mAP of 0.92 for the YOLOv5 model trained on common labels and 0.89 for the model focused on less common labels. We establish ground truth by annotating images for binary classification of detected objects into friction or nonfriction elements. Using an AdaBoost classifier, we achieve 94.8% accuracy in classifying these elements as either friction or nonfriction. This method demonstrates significant promise for advancing road safety, traffic management, and autonomous driving in India.**

***Index Terms*—Object detection, Road segmentation, Friction elements detection, Indian road scenarios.**

979-8-3503-7952-5/24/$31©2024 IEEE

I. Introduction

In many Asian streets, particularly in India, street vendors line the roads with goods such as vegetables and fruits, while pedestrians and animals navigate through these crowded areas. These roads usually lack clear markings that pose challenges for the drivers. On the contrary, foreign roads boast predictable traffic flow, clear lanes, and strict rule adherence, with designated speed lanes and pedestrian zones ensuring efficient traffic flow as depicted in Fig.1. This contrast between the two environments highlights the challenges faced by autonomous driving systems (ADAS), which rely on clear road markings and consistent traffic patterns, both often absent on India's dynamic streets. To address this, systems are needed to alert drivers of obstacles, known as road friction elements, which include pedestrians, animals, street vendors, and potholes. To better understand the impact of these friction elements, a study introduces the roadside friction index (RSFI) as a crucial parameter for measuring the influence of side friction on urban traffic flow. This study focuses on elements like pedestrian crossings, on-street parking, slow-moving vehicles, and street vendors [7]. However, less than 10% of the studies have explored other frequently encountered road friction elements, such as animal movement and land-use activity [2]. Although there is a theoretical understanding of these challenges, practical implementation remains limited. Additionally, previous studies have primarily concentrated on individual friction elements, such as street vendor detection [8], animal detection [10], and illegal parking detection [12], and have often overlooked the consideration of multi-class friction elements. It is thus essential to detect these road friction elements to improve overall road safety, as this ensures timely alerts for drivers and helps avoid potential hazards. Therefore, efficient object detection systems are necessary to accomplish this. In line with this, recent studies have investigated YOLOv3's speed advantages for real-time tasks in road safety systems [6], with different YOLO versions being tailored for specific needs. However, it is not enough to merely identify these road friction elements, understanding their positions relative to the road is also critical. Thus, road segmentation becomes necessary, which can be achieved using models such as U-Net [14] and DeepLabv3+ [17]. Towards this, our contributions are as follows:

1) Built the dataset using the Indian Driving Dataset (IDD) [3] Segmentation dataset, with two additional classes - garbage and carts, which were not previously present. This is crucial due to the dataset's focus on India's diverse traffic scenarios, including unstructured road infrastructures, ensuring to address the unique challenges posed by Indian roads.
2) Propose a methodology for detecting and classifying multi-class road friction elements, including *nonmotorized vehicles, parked vehicles, garbage, pedestrian groups, and livestock* which involves establishing the logic for categorizing objects as friction elements by discerning their spatial relationship with the road, which requires precise road segmentation.
3) Generate ground truth for the binary classification of the detected elements as friction or nonfriction.

The remainder of the paper is organized as follows: Section 2 briefly presents related work. The proposed road friction element detection is described in Section 3. Experiments and results are presented and discussed in Section 4. Finally, the paper is concluded in Section 5.

II. Related Work

From the Literature survey, we observe that attention has been directed towards understanding the impact of road friction elements, advancing object detection techniques, and road segmentation methods. Studies in these areas have utilized various computer vision and deep learning approaches to enhance detection accuracy and improve road safety.

A. Detection of Roadside friction elements

The presence of various obstacles and activities occurring on or near the roadway significantly impacts traffic flow and safety, particularly in unstructured environments. These are referred to as road friction elements. These elements are prevalent in many developing regions where roads are commonly occupied by vehicles, pedestrians, and animals, often without clear demarcations or traffic rules. Understanding and managing these friction elements is crucial for improving traffic efficiency and safety.

Studies have examined road friction elements in unstructured environments, including one on urban traffic performance in Tanzania, Africa. Through empirical and statistical analyses, researchers have pinpointed key factors influencing traffic speed and capacity, which have informed traffic management strategies aimed at improving efficiency and safety [1].In the Indian context, research primarily focuses on the impact of roadside friction elements on traffic flow. Studies reveal that both roadside friction and traffic volume contribute to a 9% reduction in roadway capacity [11]. Additionally, another study introduced the Road Side Friction Index (RSFI), which is calculated using friction element counts and weight factors. This index is directly proportional to the Percentage Speed Reduction (PSR) [7].

Conclusively, these studies highlight five major road friction elements: slow-moving vehicles, parked vehicles, garbage, pedestrian groups, and livestock. While these studies focus on measuring the roadside friction index and its impacts, they lack practical implementations. Hence, we propose a computer vision-based solution for detecting these objects.

The integration of computer vision has significantly advanced the detection of road friction elements. For example, locally normalized Histograms of Oriented Gradients (HOG) have demonstrated strong performance in this area [10]. Additionally, a combined approach using a Single Shot MultiBox Detector (SSD) and support vector machines (SVM) proved effective for detecting road friction elements [12]. Another study utilized MobileNet-SSD for detection and VGG-16 for classification, providing a valuable dataset for further research [8].

All these previous methodologies primarily focused on detecting single objects on the road that could potentially hinder traffic flow. In contrast, our approach aims to detect multiple classes of road friction elements using the YOLOv5 model.

B. Road Segmentation

Road segmentation involves identifying and outlining road networks from images or remote sensing data. This task is vital for determining the location of drivable roads, which in turn is essential for identifying road friction elements. A study explores cutting-edge deep learning model architecture for semantic segmentation, focusing on the extraction of road segments from satellite images. Their experimentation utilizes Pan-Sharpened RGB input data sourced from the SpaceNet Roads Dataset and scrutinizes the performance of U-Net, SegNet, and ResNet models. Through experiments the paper compares these models and concludes that U-Net achieves the highest F1 score (0.526) using the least trainable parameters across all areas of interest, establishing its superiority in road segmentation tasks [14]. We hence propose carrying out road segmentation using the U-Net model comprising 23 convolutional layers over IDD segmentation(I & II) dataset [3].

III. Dataset Description

The dataset utilized for the study was borrowed from IDD segmentation (I& II) [15], with a total of 16,097 images used for training and testing.IDD segmentation is a publicly available state-of-the-art dataset for comprehending road scenes within unstructured environments. Most autonomous naviga-

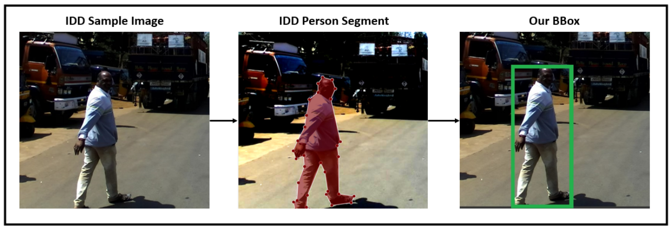

Fig. 2. The input image is from the IDD segmentation dataset, where the label ”Person” has been segmented using the segmentation data and subsequently drawn into a bounding box.

tion datasets focus on structured roads, while the Indian Driving Dataset (IDD) captures unstructured, unpredictable road conditions, including diverse traffic making it more representative of real-world Indian roads. The dataset contains 10,000 images labeled with 34 classes from 182 drive sequences on Indian roads. Captured by a front-facing car camera, the images were taken around Hyderabad and Bangalore, mostly in 1080p, with some in 720p. Out of the 34 available classes, we selected 7 distinct labels that include *animals, bicycles, cars, motorcycles, persons, riders, and trucks.* Furthermore, to enhance our comprehension of road scenes, we utilized the label,*road* for road segmentation, comprising 8088 images sourced from IDD segmentation. The existing dataset was presenting two challenges that required attention: (i)The IDD segmentation dataset had segmented data which provides finer-grained information about the contents of an image as shown in Fig. 2, resulting in fewer pixels intersecting with the road segment making our estimation of whether it is a roadside friction element or not inaccurate. (ii) The road segmentation aspect of the dataset offers detailed masks for roads in images. However, these masks may contain jagged edges, Fig. 3, which can hinder our analysis of the interaction between detected road friction elements and the segmented road.

Hence, we overcome these challenges by :

- **(i)** Segmentation data of the IDD segmentation dataset is converted into detection boxes to obtain the region of influence of the object, as shown in Fig.2. As a result, converting to bounding boxes allows us to capture the object's area of interest better.
- **(ii)** We use techniques like convolution average filtering and binary thresholding to smooth road mask edges, ensuring no irregularities. This enhancement aids in a more precise analysis of detected road friction elements and segmented roads. In Fig. 3, the impact of these techniques on the U-Net segmentation model can be seen, improving road area delineation with smoother, accurate boundaries.

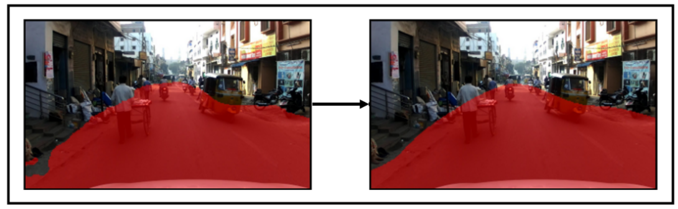

Fig. 3. The left image is the irregular road masking mapped on the original image and the right image is the smoother binary mask output on the road by using convolution average filtering and binary thresholding.

A. Data Annotation

Recognizing the need for additional road friction elements, we introduced two more labels: Garbage and Carts. These labels were manually annotated from the IDD segmentation I & II datasets, as they were not included in the original 34 annotations. Annotations were performed using the CVAT Annotation tool [16].

Fig. 4. The left images are the input images consisting of labels *carts* and *garbage* and the right images show both the labels annotated.

Ground truth annotations were performed on 413 images to establish the binary classification of road elements as friction (yes) or nonfriction (no). This process involved examining each image for road friction elements, including animals, bicycles, cars, motorcycles, persons, riders, and trucks, as well as the additional labels for garbage and carts.

B. Data Statistics

In the preprocessed dataset, we utilized the IDD segmentation dataset, which provides detailed pixel-wise counts for fully annotated 34 labels. From this comprehensive list, we identified a subset of 7 crucial labels including animals, motorcycles, bicycles, trucks, cars, persons, and riders.

Additionally, we annotated two other labels which were not present in the IDD dataset, namely garbage and carts, categorizing these 9 labels as fixated labels. Fixated labels are those that are deemed critical for the primary objectives of our analysis and require focused detection and monitoring due to their high relevance to road safety and traffic flow. The

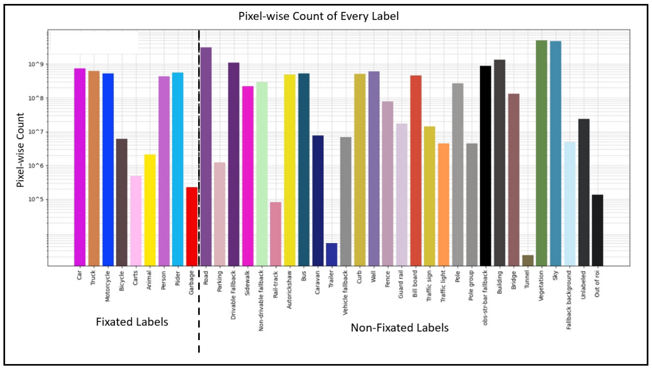

Fig. 5. The pixel-wise count of every label in the IDD dataset and includes the distinction of fixated and not fixated labels.(Note: the labels *carts* and *garbage* are manually added).

remaining 27 labels are categorized as nonfixated labels. These nonfixated labels, while still part of the dataset, are considered less critical for our specific use case and thus do not receive the same level of emphasis in our detection efforts. Fig. 5 presents the pixel-wise count of all these labels, providing a comprehensive view of the dataset.

Referring to Fig.6, for conducting object detection on the identified fixated labels, which are the road friction elements under consideration, we analyze the instance count of each of these labels.

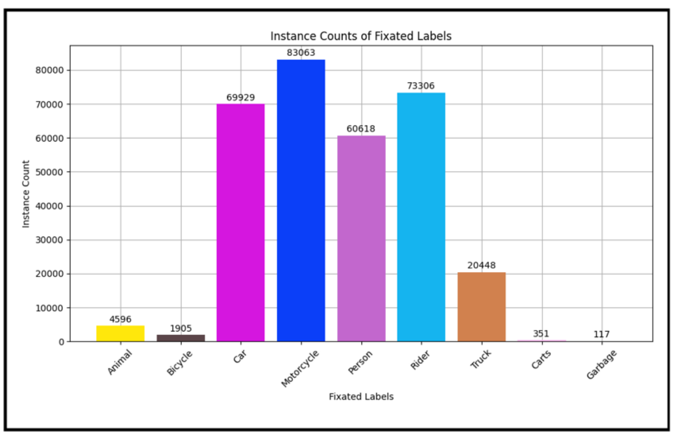

Fig. 6. The instance count of the fixated labels in the acquired dataset

IV. Methodology

This section outlines our methodology for detecting and classifying road friction elements begins with preprocessing to prepare input images. During object detection, two YOLOv5l model variants identify high-frequency and less frequent labels. For the road segmentation task, a U-Net performs pixel-wise road segmentation to map object distribution. Next is the elimination section under which distant and nonrelevant labels are eliminated based on bounding box analysis. Intersection analysis with segmented roads follows, and an AdaBoost classifier is used to classify road elements into friction or not. Each step ensures efficient and accurate detection in dynamic road environments. The pre-processing step is only done during the training of the models and not during the actual detection. Pre-processing includes filtering out the labels to create a dataset containing only the 9 required labels out of the 34 from the IDD dataset. Next, these segmentations are converted to the required bounding box format for the object detection models.

A. Object Detection

For object detection, we have selected the YOLOv5l model after comparing the detection results as discussed in subsection.V-A. The task of object detection can be divided into two sub-tasks, detection of high-frequency labels and less frequent labels so we have employed two YOLOv5l models for the task of object detection. Fig. 7(a) shows the YOLOv5l model [2] trained on high-frequency labels such as animal, person, car, truck, motorcycle, bicycle, and biker. Fig. 7(b) illustrates the second YOLOv5l model [2] specifically trained on lower-frequency labels like "garbage" and "cart," focusing on accurately detecting less common objects. Using two separate YOLOv5l models, each trained on distinct sets of labels based on their frequency within the dataset ensures accuracy across all object categories. Segregating training data based on label frequency tailors the models to prioritize accurate identification of both common and less frequent objects, enhancing the overall performance and reliability of the object detection system.

B. Road Segmentation

Once we have the object detection results, the next step is to segment the road as shown in the Fig. 7(c). For this task of road segmentation, the U-net model is employed for pixel-wise segmentation of the road region within the image. U-net model was selected after comparing the results with the other existing models as discussed in subsection.V-A. This segmentation is crucial for a precise understanding of object distribution within the road environment [14]. The U-Net model comprises 23 convolutional layers, utilizing repeated 3x3 convolutions and max pooling for down-sampling and up-sampling, followed by convolutional layers for feature extraction and reconstruction.

C. Elimination of distant and nonstationary objects

To enhance the efficiency of object detection in autonomous vehicles, we eliminate detected objects based on the area of their bounding boxes, emphasizing the importance of nearby friction elements. Prioritizing the elimination of distant and nonrelevant objects helps optimize computational resources as shown in Fig. 7(d). As the vehicle approaches these elements, they can be re-evaluated and processed accordingly, ensuring the vehicle remains responsive to its dynamic surroundings. After this elimination, We specifically address the exclusion of motorcycles that are not acting as road friction elements. If a motorcycle is associated with a rider, it is likely in motion and thus less relevant as a friction element. As shown in Fig. 7(e) we analyse the intersection percentage between the

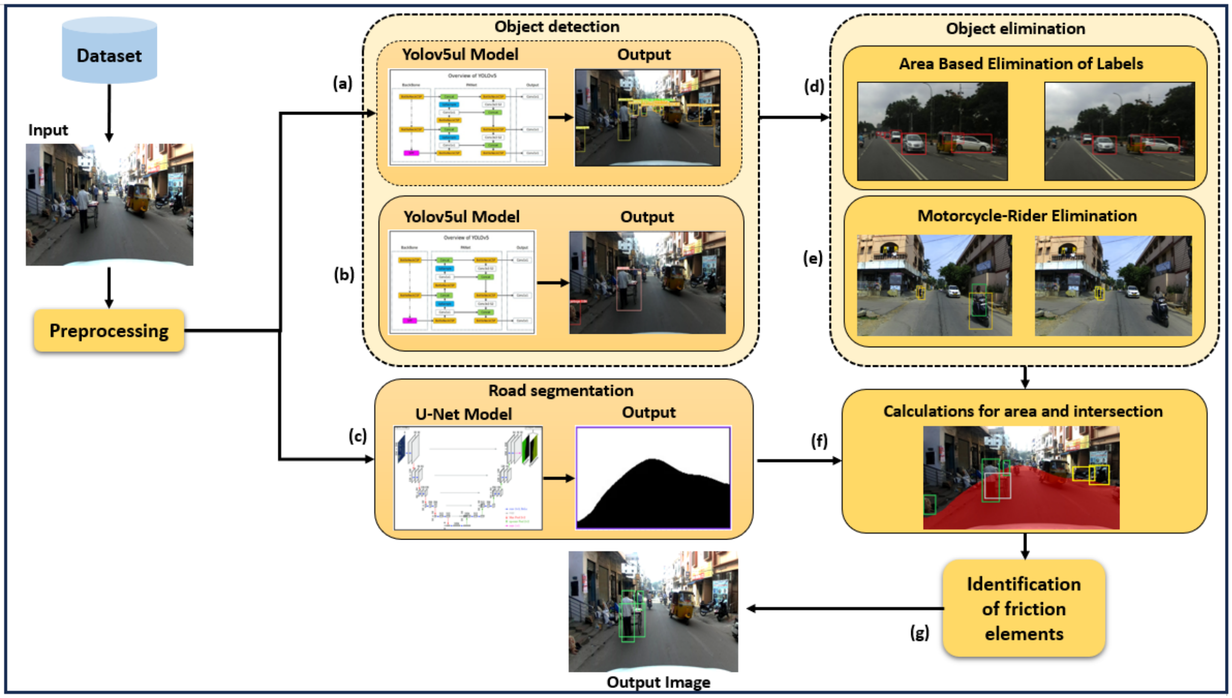

Fig. 7. The proposed framework for Road Friction Object Detection. The input is pre-processed and then the following steps are followed :(a) An object detection model detects highly frequent roadside friction elements in the input image. (b) Another object detection model detects infrequently occurring labels in the input image. (c) A segmentation model performs road segmentation, providing the output with road masking. (d) Area-based elimination is applied to the input image, showing the result. (e) Motorcycle-biker elimination process on the input image, displaying the result. (f) Bounding boxes and segmented road areas are used to calculate their intersection. (g) The model for identifying the friction elements classifies detected objects as road friction elements or not.

bounding boxes of a rider with their respective motorcycle to come to a conclusion of whether or not to eliminate a detected motorcycle, essentially refining the label "Motorcycle" i.e. one of the considered road friction elements in our case. The metrics we used for this purpose of elimination are:

- **Intersection Percentage:** Calculates the percentage of the object's area that overlaps with the segmented road region or other detected objects, providing insights into the object's proximity to the road.

$$\text{Intersection Percentage} = \left(\frac{\text{Area}_{\text{int}}}{\text{Area}_{\text{BBox}}}\right) \times 100 \quad (1)$$

where, $Area_{\text{int}}$ is the area of the intersection and $Area_{\text{bbox}}$ is the area of the motorcycle bounding box.
- **Area:** Measures the overall area of each detected object in pixels, aiding in quantifying the size and scale of identified elements.

$$\text{Area of Detected Object} = \text{Width}_{\text{BBox}} \times \text{Height}_{\text{BBox}} \quad (2)$$

where, $Width_{\text{BBox}}$ is the width of the bounding box and $Height_{\text{BBox}}$ is the height of the bounding box.

D. Identification of Friction Elements

Intersection Between Bounding Boxes and Segmented Road : Calculating the intersection between the bounding box and segmented roads provides valuable insights into an object's influence on road traffic As shown in Fig. 7 (f). Each detected label is categorized into three scenarios: on-road, off-road, and roadside. The intersection of the road with the detected object's bounding box determines the percentage of the object's influence on road traffic. Depending on the label, the percentage threshold varies to classify the object as a friction or nonfriction element.

Binary Classification of Road Friction Elements: The final and crucial step in this pipeline is the classification of roadside friction elements using a model which is trained on the ground truth, area, and the intersection percentage of each class. An AdaBoost classifier was then trained to perform the binary classification of the detected objects in the image, labelling them as either "Friction" (yes) or "nonfriction" (no) elements. As shown in Fig. 7 (f), ground truth data was prepared by annotating 287 images for labels such as car, truck, person, motorcycle, bicycle, and rider, and around 126 images for labels like garbage and carts, using the CVAT annotation tool [16] from the test images of the IDD segmentation dataset.

V. Results and Discussions

In this section, we detail the experimental setup and results obtained from the three models utilized in our study: YOLOv5 for object detection, U-Net for road segmentation, and AdaBoost for classification of detected objects. We employed two YOLOv5 models for detecting road friction elements. The first model was trained on common labels such as animals, motorcycles, bicycles, trucks, cars, persons, and riders. This model demonstrated successful learning over 50 epochs, with well-converged loss curves indicating stable training. Its precision around 0.7 indicates accurate detection of frequent objects, paving the way for improved performance. The second YOLOv5 model focused on the less frequent labels, garbage and carts. This model showed significant improvement over the course of 50 epochs. The learning trajectory, evidenced by the loss function graphs, highlights the model's potential for refinement. It effectively classified the less frequent labels, delivering commendable performance for garbage detection and satisfactory results for carts.

For road segmentation, we utilized the U-Net model with specific parameters: three epochs, a batch size of 16, a learning rate of 0.0001, and image dimensions of 160x240 pixels. The model was optimized using binary cross-entropy loss and the Adam optimizer with the specified learning rate. This setup enabled accurate segmentation, yielding promising results that were crucial for subsequent object detection and classification tasks.

Lastly, we implemented an AdaBoost classifier to categorize detected objects as friction elements or nonfriction elements. The classifier was configured with 2000 N_estimators and a learning rate of 0.275. This model effectively classified detected objects, contributing significantly to the recognition and classification of friction elements, which was integral to the study's success. The results highlight the potential of these models to enhance road safety, traffic management, and autonomous driving in challenging environments.

A. Comparison with Existing Detection and Segmentation Models

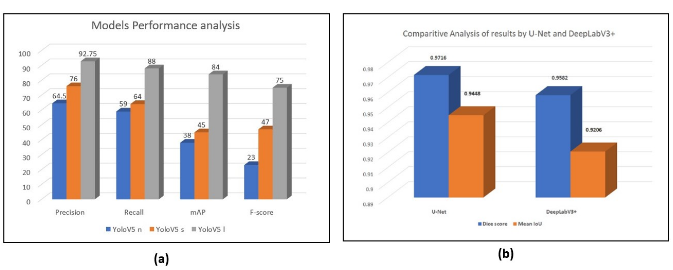

Fig. 8. (a)Comparison of YOLOv5 Models based on precision, recall, mAP, and F-score. YOLOv5l (large version) consistently outperforms YOLOv5n (normal version) and YOLOv5s (small version) across all metrics, achieving the highest precision (92.75%), recall (88%), mAP (84%), and F-score (75%). (b) Comparison of Dice Score and mIOU between U-Net and DeepLabV3+ for Road Segmentation.

This subsection explores and compares various state-of-the-art models for object detection and road segmentation. By evaluating these models, we identified the most suitable approach for our solution.For object detection, we utilized YOLOv5s, YOLOv5l, and YOLOv5n models. In Figure 8(a), YOLOv5l showed exceptional precision (92.75%) and recall (88%), with a mean Average Precision (mAP) of 84%. YOLOv5n performed well, enhancing the spectrum between YOLOv5s and YOLOv5l. YOLOv5s exhibited commendable performance with precision, recall, and mAP scores of 76%, 64%, and 45%, respectively.

In conclusion, YOLOv5l is the ideal choice for object detection, demonstrating better performance across key metrics with finely balanced precision, recall, and mAP, aligning perfectly with our objectives.

As part of the road segmentation task initially, two models, U-Net and DeepLabV3+, were employed. The performance of these models was evaluated based on several metrics, including pixel-wise accuracy, dice score, and mean Intersection over Union (IoU).Fig 8(b) presents a comparison of results from the two models. The U-Net achieved a dice score of 0.9716 and a mean IoU of 0.9448, while the DeepLabV3+ model had a dice score of 0.9582 and a mean IoU of 0.9206. The U-Net also achieved a pixel-wise accuracy of 95.95%, compared to DeepLabV3+'s 94.17%.

The U-Net model outperformed DeepLabV3+ in pixel-wise accuracy and mean IoU, making it a more suitable choice for our road segmentation tasks due to its superior accuracy and IoU metrics.

B. Quantitative and Qualitative Analysis

In Table I, Mean Average Precision (mAP) scores for the object detection model are provided for various labels. The model achieves impressive mAP scores of 0.93 for people, 0.95 for motorcycles, 0.87 for riders, and 0.92 for trucks, showcasing its proficiency in these tasks. The model also performs well with mAP scores of 0.72 for animals, 0.85 for cars, and 0.80 for carts. However, it shows less consistency with a mAP of 0.50 for garbage detection and 0.63 for bicycles. These variations highlight areas where the model's performance can be further improved. Fig. 9 shows a few samples of the qualitative results for the detection models.

TABLE I
QUANTITATIVE RESULTS: MAP OF INDIVIDUAL LABELS FOR YOLO MODELS

YOLO Model	**Labels**	**mAP**
YOLOv5l (Frequent objects)	Animal	0.723
	Car	0.858
	Bicycle	0.635
	Person	0.934
	Motorcycle	0.941
	Rider	0.877
YOLOv5l (Less frequent objects)	Garbage	0.703
	Carts	0.802

Both YOLOv5 models achieved high accuracies: the model for frequent objects is 0.92 and the one for less frequent objects is 0.89 Table.II. The U-Net road segmentation model achieved an impressive mean IOU of 0.9448, outperforming other models Table.II. The final AdaBoost classifier achieved a binary classification accuracy of 97.88%, crucial for reliable detection and differentiation of road elements Table.II.

Followed by the quantitative results are the qualitative results of each of the models.Fig. 9 demonstrates sample

TABLE II
QUANTITATIVE RESULTS OF OBJECT DETECTION AND SEGMENTATION MODELS

Model	Task	Metric	Score
YOLOv5l	Detection for Frequent objects	mAP	0.92
YOLOv5l	Detection for Less frequent objects	mAP	0.89
U-Net	Road Segmentation	mIoU	0.9448
AdaBoost Classifier	Binary Classification	Accuracy	97.88%

results of all the 9 road friction elements labels also specifying the input image scenarios. The bounding boxes associated with each class in the output images are colour-coded to allow for easy differentiation and identification.

Fig. 9. Qualitative results for the detection models: The first column shows the input image, the middle column displays the output of the detection model, and the last column presents the ground truth. The detected labels in each row, from top to bottom are truck, car, garbage, motorcycle-rider, person, animal, carts, and bicycle in various scenarios. Each label is accompanied by its corresponding confidence score.

The qualitative outcomes of U-Net road segmentation using a sample input image and its corresponding mask delineating the road segment can be seen in Fig. 10.

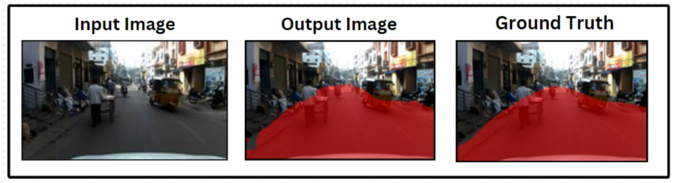

Fig. 10. Qualitative results for road segmentation: The first column shows the input image, the middle column displays the output of the segmentation model, and the last column presents the ground truth.

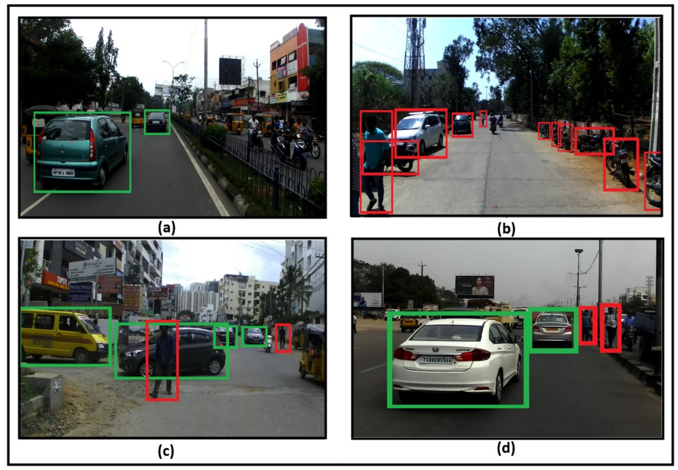

Fig. 11. Qualitative Results of AdaBoost classifier:(a)A scenario where there are no friction elements (green)(b)A scenario where there are all friction elements(red). (c) and (d) show road scenes which have both friction and nonfriction elements present.

Lastly, as observed in Fig. 11 there are various scenarios representing the presence and absence of road friction elements, classified by the AdaBoost classifier. In the depicted images, the objects identified as friction elements are highlighted in red, while those elements which are not road friction elements are marked in green.

C. Misclassification

As observed in Fig. 13(a) the model demonstrates strong performance across the five classes: person, bicycle, car, truck, and animal. The majority of values are concentrated along the diagonal, indicating successful classification for most objects.

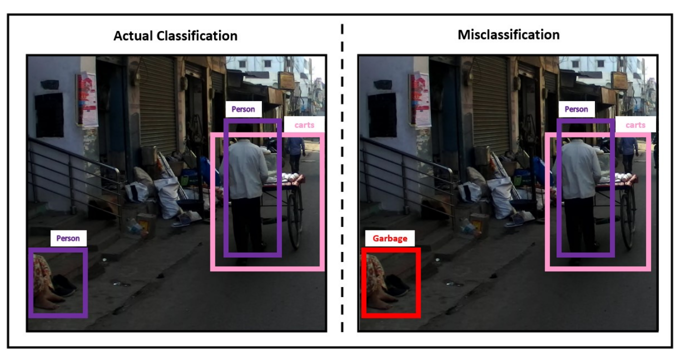

Fig. 12. The object person is misclassified as garbage.

However some misclassifications are present between person and rider arises due to their similar appearances. Misclassifications between persons and animals occur, due to limited examples of animals during model training, leading to struggles in differentiation, especially in unusual or ambiguous contexts or poses. Confusion also arises between bicycles, motorcycles, and cars, likely due to their similar structural elements such as wheels. Fig. 12 shows a special case where the object person gets confused as garbage. This is likely due to similar contextual cues and lesser training data for garbage.

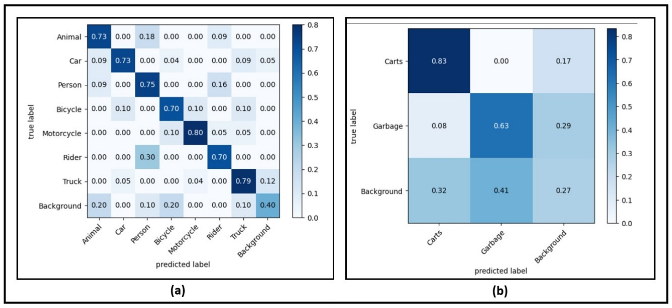

Fig. 13. (a) Confusion Matrix of YOLO Model 1 for Frequently Occurring Labels.(b) Confusion Matrix of YOLO Model 2 for Less Frequently Occurring Labels.

As depicted in Fig.13(b), the confusion matrix for YOLO Model 2 provides insights into the model's performance. Misclassifications between "Garbage" and "Cart" classes are due to limited training data for these classes.

VI. Conclusion

We have proposed a novel solution for detecting and classifying road friction elements on Indian roads, including nonmotorized vehicles, parked vehicles, garbage, pedestrians, and livestock, using road segmentation. Our object detection model, particularly the YOLOv5 for frequent labels, performed strongly, with cars detected most frequently at 85.8% confidence and bicycles at 63.5%. This model achieved an impressive mAP of 0.92, while the second YOLOv5 model for detecting garbage and carts achieved an mAP of 0.89. Additionally, our U-Net road segmentation model scored 94.48%, highlighting its precision in distinguishing roads. The AdaBoost classifier reached 94.88% accuracy in binary classification of friction elements. These metrics demonstrate the effectiveness of our solution for real-world applications. The proposed solution can be integrated into ADAS systems, providing alerts or automated actions for navigating through traffic. It can also enhance smart traffic management by identifying road friction elements, adjusting signals, or rerouting traffic to reduce congestion.However, challenges such as limited data coverage, difficulty in detecting objects like carts in low light, and the demands of real-time processing can be addressed by updating models with city-specific data and utilizing hardware acceleration, such as GPUs, to enhance real-time performance.

In the future, we aim to investigate the adaptability of our model to diverse road scenarios across regions and further integrate it into traffic management systems, addressing broader road safety and urban planning challenges.

References

[1] M. L. M. Chiguma, "Analysis of side friction impacts on urban roads: Case study Dar-es-Salaam," Ph.D. dissertation, KTH, 2007.

[2] K. Srivastava and A. Kumar, "Critical Analysis of Road Side Friction on an Urban Arterial Road," Engineering, Technology Applied Science Research, vol. 13, no. 2, pp. 10261–10269, 2023.

[3] "Indian Driving Dataset (IDD)," https://inai.iiit.ac.in/domains/idd.html, accessed March 5, 2024.

[4] "YOLOv5 Documentation," https://docs.ultralytics.com/models/yolov5, last accessed 2024.

[5] "YOLOv8 Documentation," https://docs.ultralytics.com/models/yolov8, last accessed 2024.

[6] S. Srivastava, A. V. Divekar, C. Anilkumar, I. Naik, V. Kulkarni, and V. Pattabiraman, "Comparative analysis of deep learning image detection algorithms," Journal of Big Data, vol. 8, no. 1, pp. 1–27, 2021.

[7] V. S. Arya, A. B. Binu, J. Pradeep, G. Thomas, A. Suminamol, and V. Aparna, "Impact of side friction on urban roads," Journal of Transportation Engineering and Traffic Management, vol. 1, 2020.

[8] H. N. Ağba and A. Tahir, "Street vendor detection" in 2021 29th Signal Processing and Communications Applications Conference (SIU), pp. 1–4, 2021.

[9] A. Jana, A. Sarkar, J. V. Kallakurchi, and S. Kumar, "Autonomous vehicle as a future mode of transport in India" in Proceedings of the Eastern Asia Society for Transportation Studies, vol. 12, 2019.

[10] S. U. Sharma and D. J. Shah, "A practical animal detection and collision avoidance system using computer vision technique," IEEE Access, vol. 5, pp. 347–358, 2016.

[11] P. Gulivindala and A. Mehar, "Analysis of side friction on urban arterials," Transport and Telecommunication Journal, vol. 19, no. 1, pp. 21–30, 2018.

[12] X. Xie, C. Wang, S. Chen, G. Shi, and Z. Zhao, "Real-time illegal parking detection system based on deep learning," in Proceedings of the 2017 International Conference on Deep Learning Technologies, pp. 23–27, 2017.

[13] A. Munteanu, T. Selea, and M. Neagul, "Deep learning techniques applied for road segmentation," in 2019 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), pp. 297–303, 2019.

[14] "U-Net," https://paperswithcode.com/method/u-net, 2022.

[15] "IDD Dataset," IIIT-Hyderabad, https://idd.insaan.iiit.ac.in/dataset/details/, accessed March 6, 2024.

[16] "CVAT.ai," https://www.cvat.ai/, accessed March 6, 2024.

[17] M. K. S. Mukund, "DeepLabV3Plus-Pytorch: A PyTorch implementation of DeepLabV3Plus," https://github.com/VainF/DeepLabV3Plus-Pytorch, accessed March 21, 2024.