Digital Object Identifier

LoLTV: A Low Light Two-wheeler Violation Dataset with Anomaly Detection Technique

SAMPRIT BOSE[1], MAHESHKUMAR H. KOLEKAR[1],(Senior Member, IEEE), SAHIL NAWALE[2] and DHRUV KHUT[2]
[1]Department of Electrical Engineering, Indian Institute of Technology, Patna, Bihar 801106 India
[2]Department of Information Technology, Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology, Mumbai, Maharashtra 400058 India
Corresponding author: Samprit Bose (e-mail: samprit_2221ee13@iitp.ac.in).

ABSTRACT The detection of traffic violations is important for enhancing road safety, promoting compliance with traffic regulations, optimizing traffic flow, and ensuring accountability. By identifying and addressing violations we can reduce the risk of accidents, deter potential violators, maintain smoother traffic flow, and hold accountable those who break traffic laws. Additionally, the data collected from violation detection helps in making data-driven decisions for improving road infrastructure and implementing targeted interventions. o address these challenges, there is a growing need to incorporate Artificial Intelligence (AI)-based techniques that can automate the detection and apprehension of violators, thereby reducing the reliance on manual intervention. Low light also affects traffic rule violation detection process as violations are difficult to detect in low lighting conditions. In this paper we built a first low light two wheeler violation detection dataset which contains a total of 1032 images and 1475 violations. We propose a deep learning-based system for real-time detection of two-wheeler violations using YOLO v8. Our system addresses the challenge of low-light conditions by incorporating a preprocessing step for low-light enhancement, enabling accurate detection even in challenging lighting scenarios. Evaluations were done in terms of precision, recall, accuracy and F1 score. We have obtained an accuracy of 97.05% on our low light dataset. Our approach also outperformed all models when tested on other publicly available dataset.

INDEX TERMS Deep learning, Low light enhancement, Road safety, Two wheeler violation detection

I. INTRODUCTION

A. MOTIVATION

TRAFFIC violation detection and control play a crucial role in promoting safety, efficient traffic flow, deterrence of risky driving behaviors, protection of vulnerable road users, and ensuring equal treatment for all motorists. By enforcing violations, authorities can reduce accidents, injuries, and fatalities, maintain orderly traffic flow, discourage reckless driving, safeguard pedestrians and cyclists, and uphold fairness on the roads [1]. According to the Ministry of Road Transport and Highways (MoRTH) data, the number of road accidents and fatalities in India is alarmingly high. Each year, around 1.5 lakh people lose their lives in road accidents, which averages to approximately 422 deaths per day [2]. Additionally, there are about 1,130 accidents reported daily, indicating a significant impact on road safety. In 2020, two-wheelers were the most common type of vehicles involved in road accidents, accounting for 36.7% of all accidents. MoRTH, in 2019, also reported around 40% of two-wheeler riders involved in fatal accidents were not wearing helmets [3]. This highlights the vulnerability of two-wheeler riders on Indian roads. The designated seating capacity of a two-wheeler is typically designed for the rider and one passenger. It is widely recognized that triple riding on two-wheelers is a dangerous practice and can significantly increase the risk of accidents and injuries [4]. Triple riding disrupts the balance and stability of the vehicle, compromises the rider's control, and can lead to reduced visibility [5]. Engaging in wheelies and stoppies on public roads is also considered a violation of traffic rules. Both maneuvers involve lifting either the front or rear wheel of a motorcycle off the ground, which can be dangerous and increase the risk of accidents. Performing wheelies and stoppies is illegal in many countries, including India, due to safety concerns. Authorities enforce penalties and fines for individuals who violate these rules to discourage such risky behavior and ensure the safety of all road users. Violating traffic rules on two-wheelers has a detrimental effect on road culture [6]. It creates an unsafe environment, encourages others to break the rules, shows a lack of respect for the law, increases tensions and deteriorates road discipline.

Intelligent Transportation Systems (ITS) play a crucial role in detecting classifying and addressing traffic violations and effectively [7] [8]. ITS enables accurate and real-time monitoring of traffic conditions, ensuring prompt identification and response to violations and anomaly detection [9]. It utilizes advanced technologies like surveillance cameras and data analytics, enhancing the accuracy of anomalous event detection and reducing false alarm rates [10]. Automated enforcement capabilities and targeted resource allocation optimize violation detection and deter potential violators. ITS also generates valuable data for insights into prevalent violations and informs evidence-based decision-making for road safety improvements [11]. Ultimately, ITS promotes compliance with traffic rules, creating a safer and more efficient transportation system [12] [13]. Moreover low lighting conditions can affect the violation detection process as violators cannot be detected due to low light hence enhancement of low light is a important preprocessing step [14]. This has motivated us to develop a two-wheeler traffic rule violation detection system that would not only detect violations at normal lighting conditions but also detect the violation by enhancing low light thereby helping the concerned authorities to monitor and take necessary actions against the violators. In this paper we have emphasized on mainly four types of two-wheeler violation that includes no helmet, triple riding, wheelie and stoppie. We have constructed a dataset of two-wheeler rule violation which contains more than 1000 images at low lighting conditions and their corresponding labels. The dataset mainly emphasizes on the aforementioned classes of two-wheeler rule violation. Furthermore we have also developed a real-time two wheeler traffic rule violation system that enhances low-light videos and can be accessed through any device. We have also incorporated an automatic violation alert system that captures images in case it detects any violation and mails it to the respective authorities.

B. RELATED WORKS

In 2022 Charran et al. [15] proposed the development of an automated system to detect and address two-wheeler violations in Indian road scenarios. The system focused on detecting violations such as not wearing a helmet, using a phone while riding, triple riding, wheeling, and illegal parking. To accomplish this, they suggested a custom trained Yolo-v4 + DeepSORT algorithm for violation detection and tracking. Additionally, they propose employing Yolo-v4 + Tesseract for number plate detection and extraction to capture the vehicle numbers. By integrating these technologies, the system automated the process of issuing tickets by storing violation data and corresponding vehicle numbers in a database. The main drawback of the suggested framework is that when the system is tested on live feed from real-world scenario the accuracy is pretty low. They have not provided their detection accuracy at low lighting conditions. In 2021 Mallela et al. [16] proposed the development of a deep learning framework based on a Convolutional Neural Network (CNN) and YOLO (You Only Look Once) algorithm for detecting triple riding and speed violations on two-wheelers. The framework utilized the power of CNNs to analyze images and identify instances of triple riding and high-speed riding. The paper only focused on one type of violation and the accuracy obtained is also pretty low. Raj et al. [17] in 2018 developed a system that utilizes image processing techniques and deep convolutional neural networks (CNNs) to identify motorcyclists who are violating helmet laws. The system consisted of three main components: motorcycle detection, helmet vs. no-helmet classification, and motorcycle license plate recognition. Through image processing, the system could accurately detect motorcycles in the captured images or video frames. The deep CNNs were trained to classify whether the detected motorcyclists were wearing helmets or not. Additionally, the system incorporated license plate recognition algorithms to extract and identify the license plates of the motorcycles. The major drawback of the system is their detection accuracy which is very low. The paper also restricts into only helmet detection and no other two Wheeler violation was detected. In their work, Dahiya et al. [18] introduced a comparable method for helmet detection. They began by employing background subtraction and object segmentation techniques to detect bike riders. Subsequently, visual features and a binary classifier were utilized to ascertain whether the bike rider was wearing a helmet or not. By combining these techniques, the authors aimed to accurately identify instances of helmet usage among bike riders but were unable to showcase their performance in low light. Vishnu et al. [19] in 2017 employed adaptive background subtraction on video frames to identify moving objects, specifically focusing on motorcyclists. Subsequently, a convolutional neural network (CNN) was utilized to select motorcyclists from the detected moving objects. To further distinguish motorcyclists driving without a helmet, another CNN was applied specifically to the upper one fourth part of the image. This multi-step process allowed for the recognition of motorcyclists and the identification of those not wearing helmets. The approach only restricts itself to helmet detection and computational complexity is high. In 2022, Lavingia et al. [20] designed an automated system model for monitoring the violation of traffic rules regarding the number of people sitting on a two-wheeler. A deep learning-based solution was proposed for monitoring the violation of traffic rules related to the number of people sitting on a two-wheeler. The solution utilized the YOLOv3 (You Only Look Once) model for object recognition, specifically identifying individuals sitting on a particular vehicle. To determine if the distance between individuals on the two-wheeler exceeds a minimum threshold, a depth estimation algorithm is implemented. This algorithm calculates the 3D distance between objects based on the 2D image, aiding in accurate violation detection and monitoring. The system fails in low light conditions and accuracy is low. In 2020, Soni et al. [21] created a computer vision system based on Tensorflow Keras. Even in real time, the system is capable of detecting whether or not a biker is wearing a helmet. The system will vigilantly monitor the scene and mark those without helmets as violating the rules. The drawback of the system is it works on images and only restricted to one type of violation. In their work, Reddy et al. [22] proposed the utilization of YOLO (You Only Look Once) for detecting persons and motorcycles in images. Additionally, they employed Histogram of Oriented Gradients (HOG) features and a Support Vector Machine (SVM) classifier to specifically detect helmets in the images. The SVM classifier was trained on color histogram features of helmets, allowing for accurate identification of helmet presence. In 2016, Doungmala et al, [23] proposed a new helmet detection technique. The technique combines two methods for helmet detection to improve the detection rates. The first method involves using a Haar-like feature-based face detection approach to distinguish between individuals wearing a full helmet and those without a helmet. The second method utilizes the circle Hough transform to differentiate between individuals without a helmet and those wearing a half helmet. In their work, Chiverton [24] proposed a technique that leverages the geometric shape of the helmet and the variation of lights across different parts of the helmet. This approach utilizes a method for detecting circular arcs based on the Hough transform. The technique achieved high accuracy in helmet detection; however, its reliability was limited due to the small number of test images used in the evaluation. While the approach showed promise, further validation and testing are necessary to ensure its effectiveness in diverse real-world scenarios. Most of the aforementioned techniques are limited to one or two violations and does not perform well on low lighting conditions. In this paper we have addressed the drawbacks of the previously discussed papers and considered four different types of two wheeler violations and have also implemented a preprocessing step that would enhance low lighting conditions.

C. CONTRIBUTION AND ORGANIZATION OF PAPER

Two-wheeler violation detection is important for enhancing road safety, preventing accidents, promoting compliance with traffic rules, improving traffic flow, and enforcing legal accountability. By addressing violations promptly and consistently, it helps change behavior, reduce the likelihood of accidents, and foster a safer and more efficient transportation system. Hence we have developed a real-time two wheeler violation detection system that would also ensure road safety and proper actions can e taken against the violators. The following summarises this work's key contributions.

1) We have built a comprehensive dataset of two-wheeler rule violations comprising over 1000 images along with their corresponding labels. The dataset is specifically focused on the classes of violations including not wearing a helmet, triple riding, wheeling, and stoppie. Each image in the dataset is labeled to indicate the specific type of violation present. This dataset serves as a valuable resource for training and evaluating derp learning models aimed at detecting and classifying two-wheeler rule violations accurately. By utilizing this dataset, researchers and developers can enhance the development of intelligent systems for automated violation detection and contribute to improving road safety.
2) In addition to the dataset, we have developed a real-time system for detecting four types of two-wheeler traffic rule violations. One of the key features of our system is its ability to enhance video quality in low-light conditions, ensuring accurate detection even in challenging lighting environments. The low-light enhancement has been added as a preprocessing step.
3) The system is designed to be accessible through various devices, allowing users to monitor and receive alerts about violations remotely.
4) Furthermore, our system includes an automatic violation alert mechanism. Whenever a violation is detected, such as not wearing a helmet or triple riding, the system automatically captures images or video footage as evidence. These images are then sent via email to the concerning authorities responsible for enforcing traffic regulations. This feature ensures prompt notification and enables authorities to take appropriate action against the violators.

The proposed approach's methodology is described in Section 2, which includes a preprocessing step for low-light enancement and YOLO v8 arcitecture for trained on our proposed dataset for two wheeler rule violation. Section 3 displays the results and analyses drawn from the implemented system. Finally, Section 4 concludes the research and outlines potential future work.

II. METHODOLOGY

A. OVERVIEW

YOLO (You Only Look Once) is a leading object detection algorithm in computer vision [25]. It revolutionized real-time object detection by dividing an image into a grid and directly predicting bounding boxes and class probabilities for objects within each grid cell. With its efficient single-pass architecture, YOLO achieves high detection speeds, making it suitable for applications like autonomous driving and surveillance [26] [27]. It's success lies in its simplicity, accuracy, and ability to handle complex scenes with multiple objects. In this paper, we have used YOLOv8 architecture for detection two wheeler traffic rule violation. The most recent state-of-the-art YOLO model, known as YOLOv8, may be utilised for tasks including object identification, image categorization and instance segmentation. This version of YOLO as outperformed all other preexisting versions in terms of accuracy.

B. NETWORK ARCHITECTURE

In this section, we will describe the proposed system designed to automate the process of traffic violation detection. The system comprises four main components, which are as follows: Creating our own low-light dataset, low-light video enhancement, detection of anomaly from the enhanced footages and storing the frame when anomaly is detected.

1) Gather low light footage at real time from sources like CCTV's, mobile cameras, dashcams, or any device with a camera, extract the frames and process them one by one.

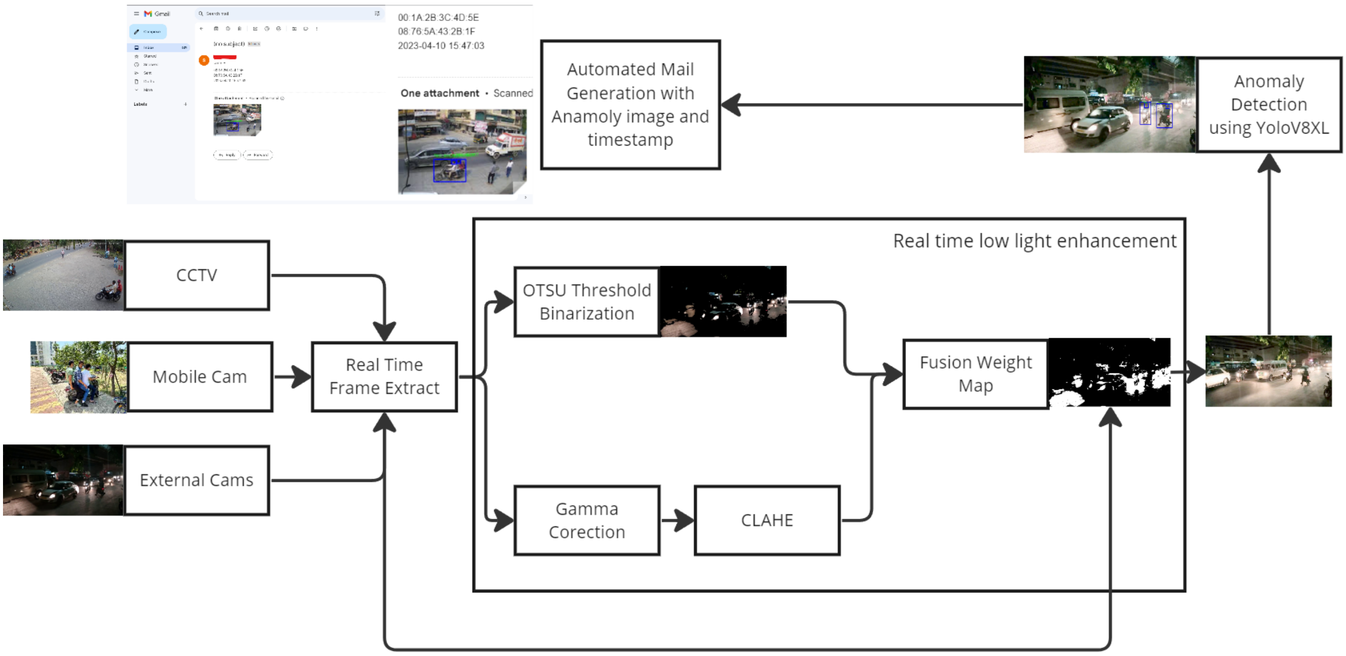

FIGURE 1. **Block diagram of two-wheeler violation detection at low lighting conditions with automated mail generation**

2) Real-time low-light video enhancement is implemented to improve the visibility and quality of low-light and night footages, enabling more effective anomaly detection. By enhancing the images in real-time, the system enhances details and reduces noise, enabling better detection and identification of anomalies or irregularities in challenging lighting conditions.
3) Detection of anomaly in the footages like no helmet, triple riding, wheelie and stoppie. By analyzing the video frames, the system identifies instances where individuals are not wearing helmets, multiple riders are on a two-wheeler, or dangerous maneuvers like wheelies and stoppies are performed. This automated detection enables timely intervention and enforcement actions to promote road safety and adherence to traffic regulations.
4) When an anomaly is detected, the system stores the corresponding frame along with the subsequent 'n' seconds of frames in a database. It then automatically generates an email containing the screenshot, timestamp, and relevant details, which is sent to the appropriate authority. This process ensures that the evidence of the detected anomaly is securely stored and promptly shared with the concerned authorities for further action.

The system for Two-Wheeler traffic violation detection at low lighting conditions and automatic mail generation is summarized in Figure 1. Further explanation of the system is provided below.

1) Pre-processing

Low light enhancement is an important pre-processing step for detecting two-wheeler rule violations because many traffic violations occur in low-light conditions, such as during the night or in poorly lit areas [28]. By improving the clarity of the footage, low light enhancement enables the system to effectively analyze and identify rule violations, reduces the chances for false count and ensures a higher level of accuracy in detecting and enforcing traffic regulations for two-wheelers. In this paper, we have used OTSU's thresholding to identify the low-light areas within each frame of a video. It is a popular image processing technique used to automatically determine the optimal threshold value for segmenting an image into foreground and background regions [29]. Once the frame is segmented into foreground and background using Otsu's thresholding, the low-light part of the image can be isolated. This allows specific enhancement techniques to be applied selectively to the low-light regions, improving their visibility and overall image quality. Let I_{xy}^{RGB} be the value of a pixel in a 24-bit RGB colour low-light image as input. First, the values of the pixel of the lightness image I_{xy} of an input image I^{RGB} are determined as follows:

$$I_{xy} = \frac{I_{xy}^{R} + I_{xy}^{G} + I_{xy}^{B}}{3} \tag{1}$$

This equation represents the calculation of the lightness value I_{xy} at each pixel location (x, y) by taking the average of the corresponding red, green, and blue component values I_{xy}^{R}, I_{xy}^{G}, and I_{xy}^{B}. Thresholding enables the method to be applied on backlit images and videos as well, along with low light enhancement. The initial step of the process involves generating a threshold from the input image in the first branch. This threshold determines the boundary between the foreground and background regions. By applying this threshold, a binary mask is created, which effectively highlights the areas of interest in the input image.

$$\mathrm{I'}_{xy} = \begin{cases} 1, & \text{if } \mathrm{I}_{xy} > \phi \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

I'_{xy} represents the intensity of the pixel at position (x, y) in the mask image. ϕ represents the threshold value. The input image is parallely enhanced in second branch using gamma correction followed by CLAHE (Contrast Limited Adaptive Histogram Equalization). The CLAHE algorithm consists of three steps: tile generation, histogram equalization, and bilinear interpolation. The input image is broken down into tiles, and histogram equalization is applied to each tile using a clip limit. Excess pixels are redistributed, and a cumulative distribution function (CDF) is calculated. The CDF values are then scaled and mapped to the input image pixels. Finally, the tiles are stitched together using bilinear interpolation to produce an output image with improved contrast. To enhance the image's visibility, the gamma conversion is applied to the original image, resulting in the light intensity conversion image, denoted as I^{γ}_{xy}. The equation is given by:

$$I^{\gamma}_{xy} = 255 \cdot \left(\frac{I_{xy}}{255}\right)^{\gamma} \tag{3}$$

For the purpose of calculating the output light-enhanced image's (O) pixel values, the α-blending of the two lightness conversion images I^{γ}_{xy} and I^{HE}_{xy} can be expressed as follows:

$$O^{xy} = (1 - \alpha) \cdot I^{\gamma}_{ij} + \alpha \cdot I^{\mathrm{HE}}_{xy} \tag{4}$$

The values of $\alpha \in [0, 1]$. To prevent degradation of the image quality the input image and the improved image are combined using a weight map. This fusion process yields the final output image, which maintains the image quality while incorporating enhancements. The weighted sum of the pixel value at (x, y) of the input image I and of the low-light improved one O as the pixel value of the resultant improved image $\tilde{O}$ as follows:

$$\tilde{O}_{xy} = \tilde{W}_{xy} \cdot O_{xy} + (1 - \tilde{W}_{xy}) \cdot I_{xy} \tag{5}$$

Where $\tilde{W}_{xy}$ represents the pixel value of the smoothed weight map. After determining the ultimate result, light enhanced image $\tilde{O}$, the pixel value at position (i, j) of the output color image $\tilde{O}^{\mathrm{RGB}}_{xy}$ is calculated using the subsequent equation:

$$\tilde{O}^{\mathrm{RGB}}_{xy} = I^{\mathrm{RGB}}_{xy} \cdot \frac{\tilde{O}_{xy}}{I_{xy}} \tag{6}$$

Using this technique we could enhance the low-light and backlight affected videos and to avoid the problem of false counting while detecting two-wheeler rule violations.

2) Violation Detection Module

In this paper, two wheeler rule violations were detected using YOLOv8. The key principle of YOLOv8 is that it performs object detection in a single pass of the neural network, making it highly efficient for real-time applications [30]. It generates bounding boxes and class probabilities for items inside each grid cell after dividing the input image into a grid. This approach allows for simultaneous detection of multiple objects in the image. YOLOv8 incorporates a deep neural network architecture, typically based on the DarkNet framework. It utilizes convolutional layers, downsampling, and upsampling operations to extract and process features at different scales, enabling it to capture both small and large objects effectively. The YOLOv8 architecture consists of three main components: the backbone, neck, and head [31]. The backbone extracts high-level features from the input image, enabling the model to understand the content and context. It typically comprises a deep convolutional neural network (CNN), such as DarkNet-53 or DarkNet-19. The backbone network consists of multiple stacked convolutional layers, pooling layers, and non-linear activation functions [32]. Its purpose is to capture and encode relevant spatial information from the input image at different scales and levels of abstraction. The neck fuses features from different scales, improving the model's ability to detect objects of varying sizes. This fusion process typically involves techniques like feature concatenation or feature pyramid construction. The head generates object predictions by transforming the fused features into bounding box coordinates, class probabilities, and other relevant attributes. It typically consists of convolutional layers, followed by a set of detection-specific layers. The head performs spatial transformation and dimensionality reduction operations to predict bounding box coordinates, object class probabilities, and any additional relevant attributes. The output is a set of bounding box predictions along with their associated class labels and confidence scores. This multi-stage architecture allows YOLOv8 to achieve efficient and accurate object detection in real-time applications. We have trained the YOLOv8 model with our own created dataset and have detected anomalies like no helmet, triple riding, wheelie and stoppie from the footages.

3) Automatic Mail Alert

We have also incorporated automatic mail generation system in this paper. Whenever an anomaly is detected, the system saves the corresponding frame along with the timestamp. Additionally, it saves all the frames for the next 'n' seconds, where 'n' is a threshold value determined through experimentation. This approach aids in the tracking of the detected violator. This method offers several advantages. First, it eliminates the need to record footage continuously. Instead, recording starts only when an anomaly is detected, which significantly saves memory that would otherwise be consumed by long, monotonous footage with no purpose. The CCTV cameras in a network are interconnected through a DVR system, which is connected to a central server where real-time detection is performed. The system is configured with the MAC addresses of all the cameras. When an anomaly is detected, the central server retrieves the MAC address of the camera in which the anomaly occurred and begins storing its footage. Furthermore, the central server maintains a table that contains the MAC addresses of the nearest cameras to each camera.

Consequently, when an anomaly is detected, the central server also starts recording the footage from the cameras that are closest to the camera where the anomaly was detected.

Once the footages from all the relevant cameras for a particular anomaly have been saved, an automated email is generated. This email includes the timestamp, MAC addresses of the cameras involved, and a screenshot of the anomaly as an attachment. The email is automatically sent to the corresponding authority's email address using an SMTP server provided by any mail service provider. This comprehensive approach ensures that the necessary footages are captured selectively, saving storage space and facilitating efficient tracking and reporting of anomalies.

III. EXPERIMENT

A. DATASET

In this paper, we have proposed our own two wheeler traffic rule violation dataset. To the best of our knowledge it is the first two-wheeler violation dataset on low-light. The dataset consists of total 1032 low-light images on two-wheeler traffic rule violations which include no helmet, triple riding, wheeling and stoppie. The low-light rule violation dataset was prepared from live CCTV footages installed at entry gate of IIT Patna and also from online sources. The distribution of the aforementioned two-wheeler rule violation dataset is shown in Table 1. From the table we can observe that there are 4

TABLE 1. Two Wheeler Violation Distribution

Classes	No. of images exhibiting violations	No of annotations of violation across all images
No Helmet	805	1320
Triple Riding	220	220
Wheeling	235	237
Stoppie	215	218

classes of rule violation. Each class has a certain number of images and their respective annotations. From Fig. 2. it

FIGURE 2. Sample annotation from low light two wheeler violation dataset

can be observed that there are multiple two wheeler violation annotations in a single image. For example in the figure 'a', it can be observed that there is a single violation made by a single violator. Similarly in figure 'c' we can see that in a single image there are cases of both triple riding and no helmet and in 'd' we can see the cases of wheelie and no helmet in a single picture which falls under the category of multiple violation and multiple violators. Also there are cases of single violation and multiple violators as shown in the images 'b', 'e' and 'f'.

B. EVALUATION TASKS

We evaluate the system across following tasks:

1) Evaluate our pipeline on real-world low-light videos initially to ensure proper low-light enhancement is taking place. This is an important preprocessing step since the present models fails to detect violations in low-lighting conditions.
2) To demonstrate real-time detection of two-wheeler traffic rule violation on live footages captured on CCTV, mobile camers or any other streaming devices with highest accuracy.
3) Compare our pipeline with other YOLO architectures on our proposed dataset in terms of true positive, false positive and precision.
4) Compare our pipeline with other benchmark datasets on two-wheeler rule violations in terms of precision, recall and time per detection.
5) To evaluate its capability and robustness, the entire pipeline is tested on 10 video feeds. These feeds encompass various scenarios, including both light and dense traffic conditions at low lighting environment. By subjecting the system to these diverse testing scenarios, its performance can be thoroughly assessed.

C. IMPLEMENTATION DETAILS

The experimental results were obtained using a laptop equipped with 16GB of RAM and a Ryzen 5 4600H processor running at 3.2GHz. The laptop also had an NVIDIA GeForce GTX 1650 graphics card with 4GB of dedicated memory. The operating system was 64-bit Windows 10, and the storage consisted of a 512GB SSD. The experimental implementation utilized Python 3.10 as the programming language, incorporating various libraries such as ultralytics, OpenCV, PyTorch, NumPy, and Pillow. The YOLO model used for the experiments was the extra-large version, comprising approximately 68 million trainable parameters. For the YOLOV8x model, the default settings were utilized except for a batch size of 5. The confidence threshold for object prediction was set to 0.5. Real-time detection was achieved at an average rate of 20 frames per second (fps). To optimize performance, every alternate frame was considered for detection, resulting in a net fps of approximately 40. The dataset used for training and testing the model was collected from various sources. It included footage from CP-PLUS CP/IP CCTV cameras placed at different heights in different parts of the city. Additionally, footage from mobile cameras such as the S21 FE and iPhone 14, as well as a 720p laptop webcam, was included in the dataset. This diverse collection of sources provided a comprehensive dataset for evaluating the model's performance.

IV. RESULTS AND DISCUSSION

The results section is divided into five parts:

1) **Section A** examines the effect of low-light enhancement on traffic videos, serving as a crucial preprocessing step.
2) **Section B** demonstrates the effectiveness of our methods in detecting two-wheeler violations even under challenging low-light conditions, surpassing the capabilities of other existing methods. Additionally, our approach successfully identifies violations that were previously undetectable.
3) In **Section C**, we evaluate our model using precision, recall, true positive, and false positive metrics, providing graphical representations of its performance and comparing it to state-of-the-art methods.
4) **Section D** showcases the accuracy of our model through testing on 10 live video footages, reaffirming its robustness and reliability. Lastly,
5) **Section E** outlines the implementation of an automated mailing system that promptly sends detailed reports, including timestamps, camera MAC addresses, and anomaly screenshots, to the appropriate authorities, thereby enhancing real-time response and reporting capabilities.

A. LOW LIGHT VIDEO ENHANCEMENT

Low light video enhancement is a crucial preprocessing step in computer vision applications, especially in low-light conditions. This step enables algorithms to better distinguish objects, capture important features, and accurately detect and track objects. By enhancing the input data quality, it improves the performance of computer vision models, leading to increased precision and recall. Moreover, it enhances the usability of vision-based applications in low-light environments, contributing to better situational awareness and decision-making. Figure 3 shows the low light enhanced output.

FIGURE 3. Low light images and their enhancements at different traffic conditions

Fig 3(a) and 3(d) shows low light image and its enhanced output respectively in less traffic conditions, Fig 3(b) and 3(e) shows low light image and its enhanced output respectively in moderate traffic conditions and Fig 3(c) and 3(f) shows low light image and its enhanced output respectively in heavy traffic conditions. The yellow boxes 3(a), 3(b) and 3(c) depicts low light areas and yellow boxes in 3(d), 3(d) and 3(f) depicts parts of enhanced portions.

B. TWO WHEELER RULE VIOLATION DETECTION

In this section we have demonstrated remarkable effectiveness in detecting two-wheeler violations, even in challenging low-light conditions, outperforming existing methods. These conditions typically pose significant obstacles to accurate detection due to reduced visibility and potential loss of crucial details. However, our approach has successfully overcome these challenges and proven its ability to identify violations that were previously undetectable. Fig 4 shows two wheeler rule violation detection in normal daylight scenario. At first the two wheelers are detected by blue bounding box. Fig 4(a) shows no helmet violation that has been detected by the green bounding box. In Fig 4(b), 4(c) and 4(e) we can observe multiple violations have been detected. The triple riding violation is denoted by red bounding box no helmet violation by green bounding box. Similarly in fig 4(e) and 4(f) at first the two wheelers are detected by blue bounding box and violations like wheelie and stoppie are detected by black bounding box. In fig 5, we have highlighted the benefit of low light enhancement in detecting two wheeler rule violation. From Fig. 5(a), 5(b) and 5(c) we can see that although there are violations that are being made by riders who are not wearing helmet, those violations cannot be detected due to low light. But in Fig. 5(d), 5(e) and 5(f) after low light is being enhanced we can observe that the violations that were previously not getting detected were getting detected after low light enhancement. It indicates that our methods have surpassed the limitations of existing approaches and are capable of unveiling violations that would have otherwise gone unnoticed. This has substantial implications for improving the overall safety and security in traffic monitoring and enforcement scenarios.

C. QUANTITATIVE AND COMPARATIVE ANALYSIS

Two wheeler violation detection was performed using Yolo-v8 on our proposed dataset on low lighting conditions. The model was tested on a total of 874 violation annotations which includes instances of no helmet, triple riding, wheelie and stoppie. Table 2 shows the result of violation detection on our low light test dataset. During the evaluation of our model

TABLE 2. Violation Detection on Low Light Test Dataset

Violation	No. of Actual violation	Precision	Recall	Accuracy (%)
No Helmet	482	0.959	0.951	94.80
Triple Riding	155	0.981	0.972	96.42
Wheelie	121	0.995	0.989	98.57
Stoppie	116	0.995	0.989	98.44
Average		**0.982**	**0.975**	**97.05**

FIGURE 4. Two Wheeler violation detection at normal daylight

FIGURE 5. Two Wheeler violation detection at night

on the test dataset, we employed various metrics including Precision, Recall and Accuracy. Precision measures the ratio of correctly predicted positive pixels to all predicted positives, while Recall measures the ratio of correctly predicted positive pixels to all actual positives. We also obtained accuracy by summing the number of true positives and true negatives and dividing it by the total number of instances. These evaluation metrics allow us to comprehensively analyze and validate the effectiveness of our model. The evaluation matrices can be formulated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

Where, TP represents the number of correctly identified positive instances (violations in this context), TN represents the number of correctly identified negative instances (non-violations), FP represents the number of false positive instances (incorrectly identified as violations), and FN represents the number of false negative instances (violations missed or not detected). From table 2 we can obseve that our test data has obtained an average precision of 98.2%, average recall of 97.54% and an average accuracy of 97.05% in low lighting conditions.

We have also plotted precision-confidence, recall-confidence, precision-recall and F1 curve to analyze the performance of our model in low light as shown in fig. 6. The precision-confidence curve shows the trade-off between

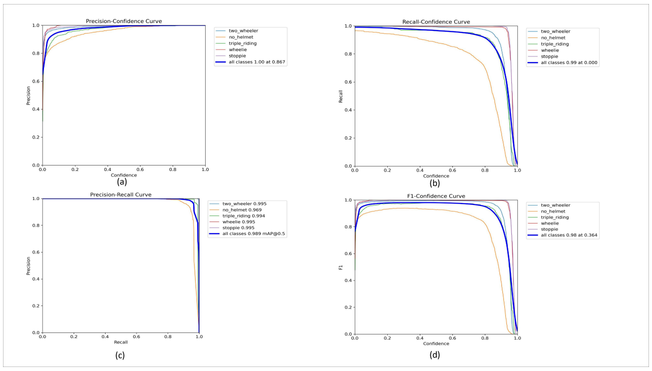

FIGURE 6. **Graphical representation of (a) Precision-Confidence curve (b) Recall-Confidence curve (c) Precision-Recall curve (d) F1 curve**

precision and confidence thresholds. Higher confidence leads to higher precision, while lower confidence results in lower precision. The recall-confidence curve depicts the trade-off between recall and confidence thresholds. Higher confidence leads to lower recall, while lower confidence increases recall. The precision-recall curve illustrates the relationship between precision and recall for different confidence thresholds, helping to determine the optimal threshold for a desired balance. The F1 score combines precision and recall into a single metric, providing an overall assessment of the model's performance. A higher F1 score indicates a better balance between precision and recall.

We have also compared our low light violation detection dataset with other existing Yolo models in terms of precision, recall and F1 score as shown in table 3 below. From the table 3 it is evident that the version of Yolo we have used in this paper (Yolo-v8) outperforms all other existing versions in terms of precision, recall and F1 score when tested on our low light rule violation dataset.

To evaluate the performance of our system, we conducted comprehensive tests comparing it to state-of-the-art models using similar deep learning frameworks. The dataset used in our evaluation focuses on two common types of violations: no helmet and triple riding. This dataset consists of 910 video clips captured at 12 observation sites in Myanmar during 2016. Each videoclip has a duration of 10 seconds, recorded at a framerate of 10fps with a resolution of 1920x1080. Notably, this dataset surpasses existing ones in terms of the number of individual motorcycles, totaling 10,006 instances. Each motorcycle in the dataset's 91,000 annotated frames is meticulously labeled with a bounding box, along with rider number per motorcycle and specific helmet usage data. Through these rigorous evaluations, we aimed to assess the efficacy and

TABLE 3. **Comparison with other Yolo models**

Sl no.	Violations	Yolo versions	Precision	Recall	F1 score
1.	**No Helmet**	Yolo-v3	0.803	0.809	0.805
		Yolo-v4	0.846	0.854	0.849
		Yolo-v5	0.902	0.897	0.899
		Yolo-v7	0.941	0.933	0.936
		Yolo-v8 (ours)	**0.959**	**0.951**	**0.954**
2.	**Triple Riding**	Yolo-v3	0.811	0.817	0.813
		Yolo-v4	0.854	0.861	0.857
		Yolo-v5	0.923	0.914	0.918
		Yolo-v7	0.962	0.954	0.957
		Yolo-v8 (ours)	**0.981**	**0.972**	**0.976**
3.	**Wheelie**	Yolo-v3	0.856	0.831	0.843
		Yolo-v4	0.889	0.874	0.881
		Yolo-v5	0.936	0.929	0.932
		Yolo-v7	0.974	0.970	0.971
		Yolo-v8 (ours)	**0.995**	**0.989**	**0.991**
4.	**Stoppie**	Yolo-v3	0.856	0.831	0.843
		Yolo-v4	0.889	0.874	0.881
		Yolo-v5	0.936	0.929	0.932
		Yolo-v7	0.972	0.968	0.974
		Yolo-v8 (ours)	**0.993**	**0.987**	**0.995**

superiority of our system compared to established models as shown in table 4. From the table we can observe that our

TABLE 4. **Quantitative Analysis on Benchmark Dataset. The Best Performance is Highlighted in Bold**

Approach	Precision	Recall	Accuracy (%)	Time per detection
Mallela et al. (2021) [16]	0.914	0.921	91.7	0.24 sec
Arshad and Kumar (2022) [33]	0.921	0.918	92.6	0.15 sec
Charran and Dubey (2022) [15]	0.963	0.940	95.2	**0.09 sec**
Proposed Approach	**0.971**	**0.965**	**96.9**	0.10 sec

method has achieved a precision of 0.971, recall of 0.965 and accuracy of 96.9% on the benchmark dataset discussed above which is the highest as compared to other methods in the comparison table. Time per detection is obtained to be 0.1 seconds in our case.

D. LIVE VIDEO

The evaluation of our methods also involved testing them on ten high-resolution low-light videos, each approximately one minute long and captured at different frames-per-second rates. These videos represented a range of traffic conditions, providing a diverse and realistic dataset for assessing the performance of our approach. The results obtained from the video feed are presented in Table 5. Out of a total of 188 violations present in the videos, our methods accurately captured 172 violations. This translates to an accuracy rate of 91.48%, indicating the effectiveness and reliability of our approach in detecting and capturing two-wheeler violations in low-light scenarios. The high accuracy achieved demonstrates the robustness and efficiency of our methods in identifying violations, even in challenging low-light conditions. This in-

TABLE 5. Two Wheeler Violation Distribution

Video	Frames per second	Actual Violations	Detected Violations	Error in Violation Detection (%)
Video 1	33	7	7	0.00
Video 2	24	14	13	7.14
Video 3	31	4	4	0.00
Video 4	25	27	25	8.00
Video 5	27	19	18	5.26
Video 6	31	9	9	0.00
Video 7	32	11	11	0.00
Video 8	22	37	33	10.81
Video 9	60	16	15	6.25
Video 10	21	44	37	15.9
Total		**188**	**172**	

dicates that our approach is capable of accurately recognizing and flagging instances of non-compliance, contributing to improved traffic monitoring and enforcement. Furthermore, an important aspect of the proposed system is that it achieved a zero false positive rate, meaning that no incorrect violations were mistakenly identified. This highlights the precision and accuracy of our approach in distinguishing genuine violations from other objects or movements even at low light. Upon analyzing the instances where violations were not detected (false negatives), it was observed that a significant portion of these cases involved individuals wearing caps or unconventional headgear. The model tended to classify such headgear as helmets, resulting in the violations not being captured.

E. AUTOMATED MAIL GENERATION

Automated mail generation plays a pivotal role in efficiently notifying relevant authorities or stakeholders when violations or anomalies are detected. This process involves the automated creation and sending of emails without the need for manual intervention. When a violation or anomaly is identified, the system triggers an automated procedure to generate an email containing pertinent information such as the event's timestamp, the nature of the violation, and supporting data or images. The email is sent through a mail server or SMTP service, either to pre-configured recipients or dynamically determined based on predefined rules as shown in figure 7. By

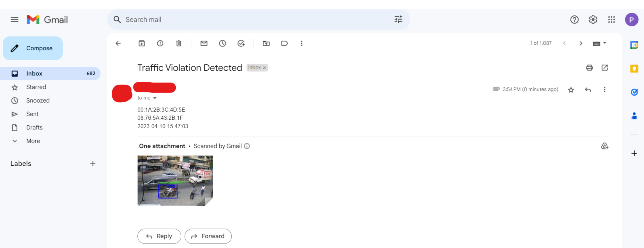

FIGURE 7. Automated mail generated output

automating the mail generation, the system ensures real-time notification and response, eliminating the need for manual effort and providing a documented record of events. This streamlines the monitoring process, enhances efficiency, and facilitates further analysis, tracking, and reporting.

V. CONCLUSION AND FUTURE SCOPE

Our proposed system for two-wheeler violation detection using YOLO v8 has demonstrated its effectiveness in detecting violations under challenging low-light conditions. The system successfully captured the majority of violations while minimizing false positives. The evaluation metrics, precision, recall, and F1 score, showcased the robust performance of our model. Additionally, the automated mailing system efficiently alerted relevant authorities, enabling prompt actions to be taken. Overall, our system showcases the potential of deep learning-based approaches for real-time violation detection and contributes to enhancing road safety.

Future work includes incorporating advanced techniques such as multi-object tracking could enable better analysis and tracking of violators across multiple frames and also incorporating additional data sources such as traffic flow patterns, weather conditions, and road infrastructure could further enhance the accuracy and contextual understanding of violations.

VI. AUTHOR CONTRIBUTIONS

All authors contributed equally to the manuscript

VII. ACKNOWLEDGEMENT

This research work is supported by TIH IIT Patna project titled ”Real-time anomaly detection in traffic video streams”. (project no: TIH/EE/RTDT/017)

REFERENCES

[1] P. D. Norton, *Fighting traffic: the dawn of the motor age in the American city*. Mit Press, 2011.

[2] S. K. Sivasankaran, H. Rangam, and V. Balasubramanian, “Investigation of factors contributing to injury severity in single vehicle motorcycle crashes in india,” *International journal of injury control and safety promotion*, vol. 28, no. 2, pp. 243–254, 2021.

[3] M. Ijaz, L. Liu, Y. Almarhabi, A. Jamal, S. M. Usman, and M. Zahid, “Temporal instability of factors affecting injury severity in helmet-wearing and non-helmet-wearing motorcycle crashes: a random parameter approach with heterogeneity in means and variances,” *International journal of environmental research and public health*, vol. 19, no. 17, p. 10526, 2022.

[4] V. Beanland, M. G. Lenné, E. Fuessl, M. Oberlader, S. Joshi, T. Bellet, A. Banet, L. Rößger, L. Leden, I. Spyropoulou *et al.*, “Acceptability of rider assistive systems for powered two-wheelers,” *Transportation research part F: traffic psychology and behaviour*, vol. 19, pp. 63–76, 2013.

[5] M. A. Regan, G. Lintern, R. Hutchinson, and C. Turetschek, “Use of cognitive work analysis for exploration of safety management in the operation of motorcycles and scooters,” *Accident Analysis & Prevention*, vol. 74, pp. 279–289, 2015.

[6] J. Damani and P. Vedagiri, “Safety of motorised two wheelers in mixed traffic conditions: Literature review of risk factors,” *Journal of traffic and transportation engineering (English edition)*, vol. 8, no. 1, pp. 35–56, 2021.

[7] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, “Data-driven intelligent transportation systems: A survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, 2011.

[8] S. Bose, C. D. Ramesh, and M. H. Kolekar, “Vehicle classification and counting for traffic video monitoring using yolo-v3,” in *2022 International Conference on Connected Systems & Intelligence (CSI)*. IEEE, 2022, pp. 1–8.

[9] N. Aslam, P. K. Rai, and M. H. Kolekar, "A3n: Attention-based adversarial autoencoder network for detecting anomalies in video sequence," *Journal of Visual Communication and Image Representation*, vol. 87, p. 103598, 2022.

[10] N. Aslam and M. H. Kolekar, "Unsupervised anomalous event detection in videos using spatio-temporal inter-fused autoencoder," *Multimedia Tools and Applications*, vol. 81, no. 29, pp. 42 457–42 482, 2022.

[11] A. Neilson, B. Daniel, S. Tjandra *et al.*, "Systematic review of the literature on big data in the transportation domain: Concepts and applications," *Big Data Research*, vol. 17, pp. 35–44, 2019.

[12] B. Chen and H. H. Cheng, "A review of the applications of agent technology in traffic and transportation systems," *IEEE Transactions on intelligent transportation systems*, vol. 11, no. 2, pp. 485–497, 2010.

[13] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2018.

[14] A. Aboah, B. Wang, U. Bagci, and Y. Adu-Gyamfi, "Real-time multi-class helmet violation detection using few-shot data sampling technique and yolov8," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5349–5357.

[15] R. S. Charran and R. K. Dubey, "Two-wheeler vehicle traffic violations detection and automated ticketing for indian road scenario," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 22 002–22 007, 2022.

[16] N. C. Mallela, R. Volety, and N. RK, "Detection of the triple riding and speed violation on two-wheelers using deep learning algorithms," *Multimedia Tools and Applications*, vol. 80, pp. 8175–8187, 2021.

[17] K. D. Raj, A. Chairat, V. Timtong, M. N. Dailey, and M. Ekpanyapong, "Helmet violation processing using deep learning," in *2018 International Workshop on Advanced Image Technology (IWAIT)*. IEEE, 2018, pp. 1–4.

[18] K. Dahiya, D. Singh, and C. K. Mohan, "Automatic detection of bike-riders without helmet using surveillance videos in real-time," in *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 3046–3051.

[19] C. Vishnu, D. Singh, C. K. Mohan, and S. Babu, "Detection of motorcyclists without helmet in videos using convolutional neural network," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 3036–3041.

[20] K. Lavingia, M. Vaja, P. Chaturvedi, and A. Lavingia, "Machine learning based approach for traffic rule violation detection," in *2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, vol. 7. IEEE, 2022, pp. 244–249.

[21] A. Soni and A. P. Singh, "Automatic motorcyclist helmet rule violation detection using tensorflow & keras in opencv," in *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. IEEE, 2020, pp. 1–5.

[22] B. Y. Reddy and M. Budka, "Detection of motor bicyclist violating traffic rules using computational neural networks," Tech. Rep, Tech. Rep.

[23] P. Doungmala and K. Klubsuwan, "Helmet wearing detection in thailand using haar like feature and circle hough transform on image processing," in *2016 IEEE International Conference on Computer and Information Technology (CIT)*. IEEE, 2016, pp. 611–614.

[24] J. Chiverton, "Helmet presence classification with motorcycle detection and tracking," *IET Intelligent Transport Systems*, vol. 6, no. 3, pp. 259–269, 2012.

[25] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[26] S. Liang, H. Wu, L. Zhen, Q. Hua, S. Garg, G. Kaddoum, M. M. Hassan, and K. Yu, "Edge yolo: Real-time intelligent object detection system based on edge-cloud cooperation in autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 25 345–25 360, 2022.

[27] Y. Cao, C. Li, Y. Peng, and H. Ru, "Mcs-yolo: A multiscale object detection method for autonomous driving road environment recognition," *IEEE Access*, vol. 11, pp. 22 342–22 354, 2023.

[28] H. Wang, Y. Chen, Y. Cai, L. Chen, Y. Li, M. A. Sotelo, and Z. Li, "Sfnet-n: An improved sfnet algorithm for semantic segmentation of low-light autonomous driving road scenes," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 21 405–21 417, 2022.

[29] C. Sha, J. Hou, and H. Cui, "A robust 2d otsu's thresholding method in image segmentation," *Journal of Visual Communication and Image Representation*, vol. 41, pp. 339–351, 2016.

[30] J. Kaur and W. Singh, "A systematic review of object detection from images using deep learning," *Multimedia Tools and Applications*, pp. 1–86, 2023.

[31] J. Terven and D. Cordova-Esparza, "A comprehensive review of yolo: From yolov1 to yolov8 and beyond," *arXiv preprint arXiv:2304.00501*, 2023.

[32] A. B. Amjoud and M. Amrouch, "Object detection using deep learning, cnns and vision transformers: A review," *IEEE Access*, 2023.

[33] M. Arshad and P. Kumar, "Detection of two-wheeler traffic rule violation using deep learning," in *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)*. IEEE, 2022, pp. 109–114.

SAMPRIT BOSE completed his master's degree (M.Tech) in Biomedical Engineering at the National Institute of Technology, Rourkela in 2021. He is currently pursuing a Ph.D. degree in the Electrical Engineering Department at the Indian Institute of Technology, Patna. His current research focuses on developing a real-time traffic anomaly detection system. His research interests encompass artificial intelligence, video surveillance, computer vision, and deep learning.

MAHESHKUMAR H. KOLEKAR (Senior Member, IEEE), was born in Satara, India. He received the Ph.D. degree in electronics and electrical communication engineering from the Indian Institute of Technology Kharagpur in 2007. From 2008 to 2009, he was a Post-Doctoral Research Fellow with the Department of Computer Science, University of Missouri, Columbia, USA. Since 2010, he has been working as an Assistant Professor of Electrical Engineering with the Indian Institute of Technology Patna (IIT Patna), where he served as a Coordinator with the Department of Electrical Engineering, in 2013, for one year and has been a Coordinator with the Center for Advanced Systems Engineering since 2014. He is currently an Associate Professor at IIT Patna and also holds the position of Institute Ph.D. Coordinator. His research interests are in digital image and video processing, video surveillance, and medical image processing. He was a recipient of the Best Paper Award from the Computer Society of India for his research paper. In 2012, he was an Organizing Chair of the IEEE sponsored International Conference on Signal, Image and Video Processing at the IIT Patna, a Guest Editor of the Journal of Medical Imaging and Health Informatics (American Scientific) for Special Issue on Biomedical Signal and Image Processing, a Tutorial Chair of the 4th IEEE International Conference on Intelligent Human Computer Interaction, a Keynote Speaker of the International conference on Biomedical Systems, Signals and Images at the IIT Madras, an Editorial Board Member of the American Journal of Signal Processing (Scientific and Academic). In 2014, he successfully organized Special Session on Intelligent Video Surveillance Systems at the International Conference on Systems Engineering, Las Vegas, NV, USA.

SAHIL NAWALE 10+2 qualified in science from Matoshri College Of Engineering in 2020, currently pursuing senior year in B.tech Information Technology from college Sardar Patel Institute Of Technology (S.P.I.T). He is also currently a research intern at Indian Institute Of Technology, Patna (IITP) and is working in research fields of image enhancement and video surveillance. His research interests include computer vision, deep learning and artificial intelligence.

DHRUV KHUT completed his 10+2 education in science from Sardar Vallabhbhai Patel Vidyalaya and Jr. College in 2020. Currently, he is a senior year student pursuing a B.Tech in Information Technology from Sardar Patel Institute of Technology (S.P.I.T). Additionally, he is working as a research intern at the Indian Institute of Technology, Patna (IITP), where he is actively involved in various projects related to image enhancement and video surveillance. His research interests revolve around image processing, artificial intelligence, deep learning and computer vision.