Grenze International Journal of Engineering and Technology, June Issue

Towards the Safer Roads: Driver Helmet and Number Plate Detection System using YOLOv8 and Google Vision API

Kapil Tajane[1], Rahul Pitale[2], Atharva Khairmode[3], Atharva Kore[4], Prathmesh Khutale[5] and Sachin Girawale[6]
[1-6]Pimpri Chinchwad College of Engineering, Computer Engineering, Pune, India
Email: kapil.tajane@pccoepune.org, {rahul.pitale, atharva.khairmode21, atharva.kore21, prathmesh.khutale21, sachin.girawale21}@pccoepune.org

***Abstract*— Now a days road safety is a growing concern, due to numerous accidents caused by non-compliance with road safety laws. This research represents an automated system for identification of driver helmet and number plate using YOLOv8 deep learning computer vision model combined with Google Vision API for text recognition from number plate leveraging latest cloud technology for enhanced accuracy and faster processing. The system is designed to efficiently detect whether a rider is wearing a helmet or not and simultaneously extract license plate information for identification purposes.**

Our method utilizes YOLOv8's real time object detection features to reliably identify riders and their helmets in different lightning and environmental situations. After detecting a vehicle, we extract the number plate and process it with Google Vision API, which transforms the text from images into a format that can be processed. This two-tiered system allows for effective enforcement of traffic regulations and can be incorporated into smart surveillance systems to improve compliance with road safety. Experimental results show a strong ability to detect helmets and extract number plate information, indicating that this method could be a valuable tool for automated traffic monitoring. Future improvements might include facial recognition and vehicle tracking to create a more effective enforcement system. This research plays a significant role in advancing intelligent transportation systems by offering an efficient, AI-powered approach to enhancing road safety and law enforcement.

***Index Terms*—Object Detection, YOLOv8, Google Vision API, Helmet Detection, Number Plate Recognition, Computer Vision, Road Safety, Traffic Surveillance, Real Time Detection styling.**

I. INTRODUCTION

Road safety is becoming an increasingly important issue worldwide, especially for those who ride two-wheelers. In response, governments and traffic authorities have put strict laws in place requiring the use of helmets to reduce the chances of serious injuries [8]. Despite this, a large number of motorcyclists still ignore these helmet laws, resulting in dire outcomes during accidents. Additionally, vehicle identification through number plates is essential for enforcing traffic laws, preventing crime, and enhancing transportation systems [9]. Traditional approaches to monitoring helmet use and identifying vehicles depend on manual checks by traffic officers or fixed surveillance systems, which can be inefficient, resource-heavy, and susceptible to human mistakes.

Grenze ID: 01.GIJET.11.2.535

With the rise of deep learning and computer vision, automated solutions have become a practical way to tackle these issues. Recently, object detection models have shown impressive advancements in both speed and accuracy [3]. Notably, the YOLO (You Only Look Once) series has attracted considerable interest for its capability to detect objects quickly and accurately [3]. The newest version, YOLOv8, brings better feature extraction, more precise bounding boxes, and lower computational demands, making it an excellent option for real-time tasks like helmet detection and vehicle identification [6].

In addition to object detection, extracting text from license plates necessitates a strong Optical Character Recognition (OCR) system [5]. Conventional OCR methods frequently face difficulties with inadequate lighting, motion blur, and different font styles. Cloud-based solutions like Google Vision API have shown great effectiveness in overcoming these obstacles by utilizing deep learning for precise text extraction [13]. By combining YOLOv8 with Google Vision API, we can develop a robust system that not only detects riders without helmets but also identifies vehicles and extracts license plate information for additional processing.

While there have been advancements in deep learning for traffic surveillance, current methods still encounter various challenges. Many real-time systems either do not operate quickly enough for practical use or sacrifice accuracy because of hardware limitations [9]. License plate recognition continues to be difficult in situations where plates are blocked, angled, or photographed in different lighting conditions [5]. Furthermore, most existing models demand substantial hardware resources, which makes large-scale deployment expensive and inefficient [11]. The main goals of this study are to create a reliable deep learning model for detecting helmets and number plates, to extract license plate information using OCR, and to assess the system's effectiveness in real-world situations. By combining object detection with text recognition, the system seeks to automate the enforcement of traffic rules and improve road safety monitoring.

A. Motivation and article overview

To tackle these challenges, this research introduces an intelligent system that integrates deep learning-based object detection with cloud-based OCR for detecting helmets and identifying vehicles. The system employs YOLOv8 to spot motorcyclists and determine if they are wearing helmets. At the same time, the model identifies number plates and extracts text using the Google Vision API. This proposed system is crafted to operate efficiently in real-time while ensuring high accuracy and resilience across various environmental conditions.

The rest of this paper is organized as follows. The next section examines the current literature on deep learning-based traffic surveillance systems, emphasizing significant advancements and challenges. After that, the methodology section describes the dataset preparation, model selection, training process, and system architecture. The experimental results section provides an analysis of the proposed system's performance, covering accuracy, processing time, and resilience to environmental changes. Lastly, the conclusion winds up the findings, addresses limitations, and suggests possible avenues for future research.

B. Glance at previos research

This research depends on the previous studies in real-time object detection for helmet detection dataset. In our last work, we compared various object detection deep learning models for helmet detection [2]. The results represented that YOLOv8 surpassed other models in speed, and computational efficiency, establishing it as the best option for this application. By utilizing these results from our previous research, this study expands the system by integrating real-time OCR capabilities and thoroughly assessing system performance across various conditions [2]. Our previous research represented and concluded YOLOv8 model as best suited model for real time helmet and number plate detection, with high accuracy and faster speed over traditional models like Faster-RCNN or previous versions of YOLO itself for real time conditions [2].

II. LITERATURE REVIEW

A. Evolution of object detection

Object detection has evolved from handcrafted feature-based methods to neural network-based models. The first approaches used hand-extracted features like Haar Cascades and Histogram of Oriented Gradients (HOG) combined with Support Vector Machines (SVM). These did well on simple detection tasks but failed for more complex scenarios.

Models like Deformable Part Model (DPM) and Region-based CNN (R-CNN) with better accuracy were developed later but they were very computationally expensive [1]. Faster R-CNN, Single Shot Detector (SSD) and You Only Look Once (YOLO) with faster detection speed and higher accuracy were developed which opened the door for real time applications [1].

YOLO in particular was able to perform object detection on entire images in a single forward pass thereby achieving high efficiency [3]. YOLOv8 takes this even further with optimized feature extraction, anchor-free detection, and improved accuracy – especially suitable for real time safety applications like helmet and number plate detection [4].

B. Object detection and road safety application

Another important application of object detection in road safety is the detection of motorcycle helmets [9]. There are many traffic accidents that cause severe head injuries because the victims do not use helmets as required by law. Automated motorcycle helmet detection systems that use deep learning models can effectively detect the absence of helmets on the heads of riders and help law enforcement enforce the laws. Studies show that CNN-based models are more accurate than conventional image processing techniques in detecting helmets in various lighting and environmental conditions [9].

Another road safety application that object detection improves is automatic number plate recognition (ANPR), which is important in traffic management, law enforcement, and toll collection. ANPR depends on the traditional methods for license plate text extraction [9], but with the emergence of deep learning-based approaches these have been improved. Models such as YOLO, R-CNN can be utilized with OCR frameworks such as Google Vision API to notably improve accuracy. This makes it suitable to detect vehicles quickly and effectively and to enforce traffic regulations.

Object detection also plays an important role in ADAS systems and autonomous vehicles [7]. ADAS and autonomous vehicles are dependent on real-time object detection to identify road hazards, pedestrians, lane markings, and other vehicles. This increases driving safety. Object detection on road objects is well-studied with models like YOLOv4 [7], YOLOv8 [7], and other deep learning models [7]. It is demonstrated that deep learning models outperform state of the art on detecting road objects, especially in challenging conditions such as low light or obstructions [7]. Object detection is also backbone of traffic surveillance systems where it helps to monitor blockages, detect accidents, and improve road infrastructure planning [7].

There are still many challenges in using object detection models in road safety applications [8]. The impact due to environmental conditions, obstructions, and real-time requirements on detection accuracy and response time have been analyzed [3]. Furthermore, scalability and optimization for embedded and edge computing devices is critical for practical applications in surveillance cameras and roadside monitoring systems [8]. Nonetheless, ongoing research in deep learning architectures and hardware acceleration can help make object detection systems more reliable and efficient for road safety and traffic management [8].

C. Previous research and existing approaches

Previous road safety object detection research focused on improving accuracy and robustness for real-time applications [10]. Feature-based approaches such as edge detection and more recently machine learning classifiers such as Support Vector Machines (SVM) and Decision Trees [10] have been the mainstay of previous research. While these approaches were robust in controlled environments, they often struggled with dynamic lighting conditions, background complexity and the demands for real-time processing [10].

Compared to the previous studies in helmet detection, they have used color-based segmentation and manual feature extraction methods to classify the riders who wear helmets and those who do not wear helmets [10]. But recently, deep learning has improved the accuracy of detection using CNN based models like Faster R-CNN and YOLO which automatically learn important features from images [11]. Many studies have used R-CNN based model performing high in accuracy and also research has shown that YOLOv3 and YOLOv5 are good in real-time helmet detection, so they are suitable for using in traffic surveillance systems [11].

Traditionally, number plate recognition relied on Optical Character Recognition (OCR) techniques and image preprocessing to extract text [5]. These methods have limitations such as different font styles, image resolutions and obstructions. Recent techniques use YOLO for license plate detection followed by CNN-based OCR systems such as the Google Vision API which have improved accuracy and adaptability to traffic situations [5].

Helmet and number plate detection Object detection is also an important task in traffic monitoring and ADAS [7]. Faster R-CNN, SSD, and YOLO architectures have been extensively studied for vehicle detection, pedestrian tracking, and traffic law enforcement. Comparisons of these architectures show that YOLO offers a trade-off between speed and accuracy, making it suitable for real-time applications [2].

D. Limitations in previous research and scope of advancement

While the state-of-the-art deep learning-based object detection models have improved the accuracy of helmet and number plate detection there are still challenges to be addressed. Most state-of-the-art methods use

Convolutional Neural Networks (CNNs) for feature extraction and classification with models such as Faster R-CNN, SSD and previous versions of YOLO commonly used in traffic surveillance [2]. However, these models have limitations in terms of computational cost, real-time performance and the ability to perform well in different environmental conditions.

The major drawback of most deep learning models is their slow inference time. Faster R-CNN achieves high accuracy (often over 90% in controlled environments) but has slower inference time due to its two-stage detection process [2]. This makes it less suitable for real-time applications where fast detection is critical, such as traffic monitoring and law enforcement. YOLO (You Only Look Once) models have a single-stage pipeline, greatly reducing processing time while still achieving competitive accuracy [2].

Another issue is accuracy in different conditions. Motion blur, low-light conditions, occlusions, and bad weather conditions can all affect how well the model performs. Studies show that traditional YOLO models (e. g., YOLOv3 and YOLOv5) can reach real-time speeds of 45–60 FPS, but have poor performance for small or partially obscured objects [2]. Moreover, number plate recognition accuracy can drop by 15-20% in low-light conditions with standard OCR methods, emphasizing the need for improvements in text extraction models (e. g., Google Vision API and deep learning-based OCR) [5]. YOLOv8, overcomes these limitations and achieves better accuracy and speed than previous YOLO variants and Faster R-CNN [4]. YOLOv8 achieves mean Average Precision (mAP) scores of over 50% on benchmark datasets with inference speeds of over 70 FPS on high-end GPUs [4], making it particularly suitable for real-time traffic monitoring applications. YOLOv8 also features improved anchor-free detection and transformer-based improvements that allow it to identify smaller objects such as helmets and license plates with greater accuracy than previous models [2].

III. METHODOLOGY

The proposed system is an automated system for real time helmet and number plate detection wherein the system captures whether the motorcyclist is wearing helmet or not and simultaneously fetching number plate to fetch information of respective vehicle and motorcyclist by using YOLOv8 deep learning model and cloud-based Google Vision API.

A. Implementation approach

- Data collection and preprocessing: Data collection and preprocessing First step is data collection and preprocessing, where motorcyclist dataset (with and without helmet) and vehicle number plate dataset is collected. For robustness several data augmentation techniques are used like rotated images, changed brightness, added noise [12]. Images are resized and normalized to keep them in same scale during training. All these dataset are segregated as 80: 20 split (80% training data and 20% testing and validation data) [12].
- Model training: The YOLOv8 model is then trained for Object Detection. Here we focus on Helmet and Number Plates Detection. The model is fine-tuned with optimized hyperparameters like learning rate, batch size and IoU threshold. For ensuring accurate detection, performance is evaluated using metrics like Precision, Recall and mAP [6]. The model is trained for 20 epochs to avoid uderfitting and overfitting. Graph against various parameters like Precision, Recall, mAP (mean average precision) are generated to validate model performance and accuracy.
- Integration with system GUI and cloud API: A Tkinter-based desktop application has been created to enable user interaction. This application allows users to upload images or videos, process them with just one click, and view the results of the detection. The system highlights the detected objects using bounding boxes and shows the extracted text. After training, the system incorporates Google Vision API for Optical Character Recognition (OCR), which captures alphanumeric text from the identified number plates. The extracted text undergoes processing and formatting to enhance accuracy, ensuring dependable recognition even in images that are motion-blurred or taken in low light.

B. System Workflow

The system begins by processing input video frames to identify motorcyclists and vehicle number plates using the YOLOv8 object detection model. It extracts the helmet status (whether the rider is wearing a helmet or not) and the area containing the number plate. After detecting the numberplate, we use the Google Vision API for OCR to get the alphanumeric text. The data and confidence scores for each detection is presented in a desktop application built with Tkinter which shows the results with bounding boxes around the detected objects. The final application is implemented for real time performance and packaged as a standalone executable so that it can be easily run without any extra dependencies.

IV. Results & Discussion

The Results and Discussion section evaluates the proposed system using performance metrics and visual inspections. The proposed system was evaluated on a dataset consisting of diverse motorcycle images taken under different conditions such as different lighting conditions and angles [12]. Key performance metrics such as precision, recall and mean Average Precision (mAP) [12] were used to analyse the performance of YOLOv8 on detecting helmets and number plates.

Further analysis is done using performance curves which includes Precision Recall curve, F1 score confidence threshold curve, Precision-Confidence and Recall-Confidence curves, emphasizes the balance between detection accuracy and threshold selection. The confusion matrix and mean Average Precision (mAP) give additional validation of the system's robustness. The outcome shows that YOLOv8 surpasses traditional CNN-based methods in terms of both detection speed and accuracy, marking its effectiveness for real-time object detection tasks in road safety applications.

A. Confusion Matrix and mAP analysis

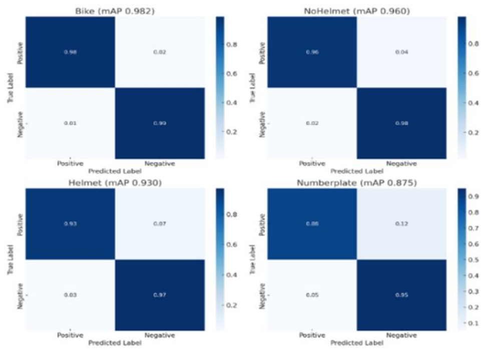

Fig 1: Confusion matrices for different classes

Each confusion matrix visually represents how well the model classifies data by comparing the true labels with the predicted ones. The mean Average Precision (mAP), a key evaluation metric, is also shown for each category. The confusion matrix for Bike detection shows an mAP of 0.982, indicating very accurate predictions with few misclassifications. The NoHelmet detection model has an mAP of 0.960, reflecting its strong reliability in identifying riders without helmets. The Helmet detection model records an mAP of 0.930, which is slightly lower but still effective in distinguishing between helmeted and non-helmeted riders. Lastly, the Number Plate detection model achieves an mAP of 0.875, which, while a bit lower than the others, still performs very well in diverse conditions. The model performs at combined mAP of 0.937@0.5 (combined for all at 0.5 confidence threshold).

Each confusion matrix exhibits a similar structure, where the diagonal elements represent correct classifications, and the off-diagonal values indicate misclassifications. The high diagonal values (close to 1) suggest strong detection capabilities, while the low misclassification rates affirm the model’s effectiveness. These results confirm the YOLOv8 model’s superiority in real-time object detection.

B. Model Benchmarking and Performance trends

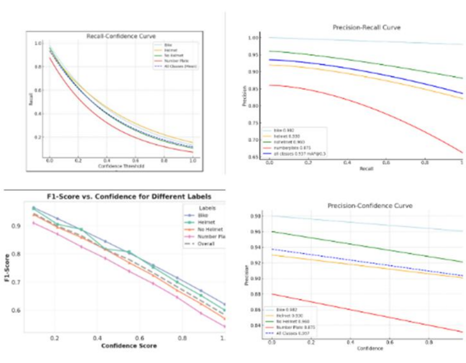

Fig 2: Result curves and performance trends

- Precision Recall curve: The given Precision recall curve shows trend between precision and recall function of the model at threshold confidence of 0.5. The curve demonstrates that ability of model to determine bike class is almost close to one representing strong detection ability as AUC-PR for bike class is much higher than others. AUC-PR for numberplate class is lowest among all the classes making model little less accurate compared to other classes. Helmet and no-helmet class has good AUC-PR curve representing strong ability of model to identify whether rider is wearing a helmet or not. The overall PR curve shows a good trend in accurately detecting classes with high precision.
- Precision Confidence curve and Recall Confidence curve: Both the curves show trend between precision and varying confidence threshold and recall and varying confidence threshold. As per the trend in both the curves it is clear that model becomes more selective as value of confidence threshold increases. Model declines less sure cases decreasing number of false positive but also rejects some correct predictions showing slight decline in precision. However the recall decreases with increase in confidence threshold as number of flase negative increases.
- F1-score vs confidence threshold curve: Based on the F1 score at varying threshold the best confidence threshold is chosen. The most balanced point in the curve is considered as best confidence threshold for the model to perform accurate predictions. The mid range of confidence threshold shows best trade offs declaring it as best range for real world scenarios. Further that model becomes more selective and before it model increses number of false positives.

V. CONCLUSION

This research shows an efficient way for real time helmet and number plate detection using YOLOv8 deep learning model and Google Vision API for text identification. This study shows advantages of using YOLOv8 and latest deep learning approaches over traditional models showing superior performance and accuracy. The evaluation metrics shows the robustness of proposed system in detecting helmets and number plates with high reliability.

By leveraging deep learning based object detection and OCR (Optical character recognition), the system gives practical solution for traffic monitoring. The results show that raising the confidence threshold enhances precision but lowers recall, highlighting the importance of finding the right balance between the two. Additionally, the findings support the effectiveness of YOLO-based architectures in real-time situations, where both speed and accuracy are essential.

Further enhancement in the system may include multi-angle detection, enhancing low light performance of the model, applying new deployment techniques such as edge computing deployment for real time and enhanced processing on edge devices. This research sets a foundation for future developments in automated traffic rule enforcement and intelligent surveillance systems aimed at improving road safety.

REFERENCES

[1] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. Proceedings of the IEEE, 111(3), 257-276.

[2] Tajane, K., Khutale, P., Kore, A., Khairmode, A., & Girawale, S. (2024, November). Object Detection: A comprehensive study of famous Deep Learning approaches with case study of Helmet Detection dataset. In 2024 2nd International Conference on Advancements and Key Challenges in Green Energy and Computing (AKGEC) (pp. 1-5). IEEE.

[3] Chandana, R. K., & Ramachandra, A. C. (2022). Real time object detection system with YOLO and CNN models: A review. arXiv Prepr. arXiv2208, 773.

[4] Sohan, M., Sai Ram, T., Reddy, R., & Venkata, C. (2024). A review on yolov8 and its advancements. In International Conference on Data Intelligence and Cognitive Informatics (pp. 529-545). Springer, Singapore.

[5] Karthick, K., Ravindrakumar, K. B., Francis, R., & Ilankannan, S. (2019). Steps involved in text recognition and recent research in OCR; a study. International Journal of Recent Technology and Engineering, 8(1), 2277-3878.

[6] Bao, J., Li, S., Wang, G., Xiong, J., & Li, S. (2023, November). Improved YOLOV8 network and application in safety helmet detection. In Journal of Physics: Conference Series (Vol. 2632, No. 1, p. 012012). IOP Publishing.

[7] Balasubramaniam, A., & Pasricha, S. (2022). Object detection in autonomous vehicles: Status and open challenges. arXiv preprint arXiv:2201.07706.

[8] Oza, S., & Rathod, S. (2020). Object Detection using IoT and Machine Learning to Avoid Accident and Improve Road Safety. development, 900, 2.

[9] Kulkarni, Y., Bodkhe, S., Kamthe, A., & Patil, A. (2018, March). Automatic number plate recognition for motorcyclists riding without helmet. In 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) (pp. 1-6). IEEE. (USed CNN to achieve 98.1 accuracy but speed is less).

[10] Manocha, D., Purkayastha, A., Chachra, Y., Rastogi, N., & Goel, V. (2019, March). Helmet detection using ML & IoT. In 2019 International Conference on Signal Processing and Communication (ICSC) (pp. 355-358). IEEE.
[11] Khan, F. A., Nagori, N., & Naik, A. (2020, July). Helmet and number plate detection of motorcyclists using deep learning and advanced machine vision techniques. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 714-717). IEEE.
[12] Vakani, A. M., Singh, A. K., Saksena, S., & Vanamala, H. R. (2020, December). Automatic license plate recognition of bikers with no helmets. In 2020 IEEE 17th India council international conference (INDICON) (pp. 1-5). IEEE.
[13] Vaithiyanathan, D., & Muniraj, M. (2019, December). Cloud based text extraction using google cloud vison for visually impaired applications. In 2019 11th international conference on advanced computing (ICoAC) (pp. 90-96). IEEE.