An efficient model for camera mounted helmet and number plate detection on custom dataset

Adithya Krishna R
Department of Computer Science and Application
Amrita School of Computing, Amrita Vishwa Vidyapeetham
Amritapuri, India
am.en.p2mca22051@am.students.amrita.edu

Sandeep P
Department of Computer Science and Application
Amrita School of Computing, Amrita Vishwa Vidyapeetham
Amritapuri, India
am.en.p2mca22051@am.students.amrita.edu

Sonu P
Department of Computer Science and Application
Amrita School of Computing, Amrita Vishwa Vidyapeetham
Amritapuri, India
sonu@am.amrita.edu

Abstract—The model is developed using YOLO v8 for detecting Motorbike riders who are violating road rules. Initially, the model differentiates motorbike riders who are wearing safety bike helmets from those who are not wearing them. Riders without helmets are marked and processed by OCR to detect number-plates. Images with helmets are checked if used with cameras attached to them. If detected, it is marked and processed by OCR to detect number plates. Detected number-plates are classified under offense. The whole model is run on a custom dataset that we prepared which contains real-life road images. Then the images that are found to violating rules will be loaded to a OCR for number-plate detection. Bounding boxes is used to represent the detected objects and has made a prediction of 98% accuracy. The numbers detected from the OCR will be exported into a csv file and any duplicates removed.

Keywords— helmet, cameras, motorcycles, license plates, YOLO, ALPI.

I. Introduction

This Project uses machine learning, advanced computer vision and image and data processing to detect violations in road safety. This project uses You Only Look Once (YOLO)[1] to detect the usage of cameras or recording devices attached to helmets worn by motor bike riders. This innovation adds precision and portability making a better environment for safe traveling roads. A Deep Neural Network (DNN)[13] is a structure with multiple layers between input and output layer. DNN consists of nodes that each perform simple tasks of computation. Passing of data through these nodes to next to create a final output is called Forward propagation. During training the initial predicted output is compared with that actual output and then changes are made to biases and weights of the nodes the remove the error created. This process is called Back propagation. Our existing systems lacked the automation in detecting certain rules that are passed under Motor Vehicle Act. Our aim was to build a system that detects these missing violations. After researching, we found some. They were: (1) Recording devices or Go-Pros attached to Helmet of Motor bike riders , (2) Parking on the road side inside the line, (3) Stopping vehicle in opposite lane while waiting at a railway cross, (4) Not stopping before the line at a traffic signal.

Recently many states in India has adapted to similar new road rules and are gradually implementing them. This will be a one stop solution for all those changes. On observing the fact that all these violations are happening around us on daily because it is impossible for the police to reach everywhere to check these violations. Our roads are already integrated with surveillance cameras. With some minor adjustments we can integrate this model into it and regularize detection of rule violation. Our aim was to build a model comprising of all these features, but unfortunately due to the lack of time, we could only solve some of them. There is a future scope to integrate these and other violations are the rule system is also evolving. The aim of this project is to automate the detection of road rule violations addressing human error and other challenges to enhance road safety and build a safe travelling environment.

In the past few years, there has been a noticeable increase in the interest and efforts put into developing object detection models that are adept at crucial tasks such as pinpointing locations, drawing out features, and classifying objects within both images and videos. The labeled image dataset was organized into training, testing, and validation subsets, with proportions of 60%, 30%, and 10% respectively, to facilitate the model's learning and prediction accuracy. Bounding boxes help visually identify the different classes in an image, making it easier to determine whether a traffic rule has been violated based on the classified images. This process is crucial for accurately detecting offenses in the images.

II. Literature review

A. *Kharade et al.[1]*

This paper represents deep learning-based systems for detecting road helmet violations by motorists using the past version of YOLO which is v4. It was a cutting-edge object detection model that aimed to differentiate between riders with and without helmets in traffic to enhance road safety conditions. The approach started with computing a pre-trained CNN model on the COCO dataset, which contained 695 images, augmented to 1595 images, and split into train, test, and validate sets. The model resulted in a mean average precision of 70% and a detection accuracy of about 97% on the test dataset. This framework presented a promise to the future developments of road safety enhancement, considering more complex scenarios.

B. *Adithya H et al. [3]*

The paper proposes a dynamic traffic rule violation monitoring system that utilizes Automatic Number Plate Recognition(ANPR) with an SMS feedback thus aiming at a very effective automated traffic law enforcement system by identifying vehicles that violate rules and informing vehicle owners through an SMS system. The model developed in MATLAB, collects vehicle images, then extracts vehicle number plates using image segmentation, and instructs an OCR(Optical Character Recognition). A GSM modem is attached and used for sending SMS responses to the vehicle owner and authorities. The model achieved a 95% success rate in number plate detection and extraction thus proving the efficiency in road safety system management.

C. Vardhan et al. [4]

The paper presents a framework for distinguishing head protector infringement and perusing license plates utilizing the YOLO machine-learning calculation, close by Automated License Plate Recognition (ALPI). It expects to distinguish motorcyclists without protective caps from live transfer video outlines, characterize them, and concentrate their license plate numbers whenever found disregarding head protector regulations. This approach incorporates improving picture quality for indistinct photographs because of climate or climatic circumstances. The framework utilizes a few strategies, including Convolution Neural Network (CNN), Cross Stage Partial (CSP), and Support Vector Machine (SVM) for protective cap identification and number plate extraction, with a Programming interface for license plate recognition. The review accomplished compelling identification and characterization by handling video into outlines, applying YOLO for object recognition, and ALPI for plate number extraction, and improving transit regulation implementation effectiveness.

D. Wang et al. [5]

The paper presents an upgraded head protector wearing calculation based on the better Yolov5 s model, expecting to address issues like low accuracy and speed in complex scenes. The calculation incorporates the CBAM consideration module into the spine organization to further develop imperfection region identification, integrates the BiFPN structure instead of PANet to improve highlight representation, and adopts the SIoU loss function over CIoU to refine recognition precision. These modifications result in the model's elevated capacity to focus on significant features and further develop generally speaking identification accuracy. The better Yolov5 s model resulted in a mean average precision (mAP @ 0.5) of 98.8%, which is a 2.3% improvement over the first model, demonstrating its suitability for certifiable applications in safety head protector identification.

E. Matheus HF, et al. [6]

The paper assesses the presentation of YOLOv5 and YOLOv8 models in distinguishing vehicles and tags for Intelligent Transport Systems (ITS). Employing a dataset curated through transfer learning, the study is aiming for enhancement in urban mobility safety and efficiency. The training utilized a dataset expanded to 4,075 quality images, ensuring diversity in locations and lighting conditions. A validation set of 814 images and a test set of 409 images were used for performance evaluation. The results, analysed through confusion matrices, indicated YOLOv8's slight superiority, achieving an accuracy of approximately 97.98% and precision of 97.19%. YOLOv8 also required less training time compared to YOLOv5. Future work includes integrating a license plate number reading model and conducting tests using the model on a Raspberry Pi for real application metrics, aiming to implement effective parking system monitoring and access control.

F. Jamtsho et al. [7]

The study presents a real-time method for detecting the license plates (LPs) of non-helmeted motorcyclists using the YOLO algorithm. This study shows to employs a single CNN to detect LPs from video streams, with a focus on ensuring motorcyclist safety by enforcing helmet use laws. A centroid tracking method with a horizontal reference line is utilised for reducing false positives from helmeted motorcyclists as they exit the video frame. The method achieved a high LP detection rate of 98.52%. This system, by automating the detection of non-helmeted motorcyclists and identifying their LPs, aims to decrease the workload of traffic police and reduce helmet law violations.

G. Armstrong et al. [8]

The review presents a continuous multi-class protective cap infringement identification framework that uses a few- shot data sampling technique and YOLOv8 calculation. This framework aims to enhance traffic well-being by distinguishing head protector utilization infringement under different circumstances. The few-shot data sampling technique takes into consideration fostering a powerful model with insignificant annotations by choosing a delegate set of pictures for train- ing. YOLOv8, known for its continuous article identification capacities, is used to distinguish protective cap infringement in video outlines. The framework exhibited its viability in the 2023 AI City Challenge, Track 5, securing the 7th place with a mAP score of 0.5861. This outcome features the framework's effectiveness and heartiness in certifiable situations. Further- more, the review thought about the performance of YOLOv5, YOLOv7, and YOLOv8, tracking down that YOLOv8, partic- ularly with test time augmentation (TTA), outflanks the others in precision and is appropriate for continuous expectations.

H. Maheswaran et al. [9]

The paper presents a continuous framework for distinguishing helmet usage and perceiving number plates on motor- cyclists utilizing YOLOv5, PyTorch, and XAMPP. Pointed toward improving street well-being by implementing helmet laws, the framework processes live video feeds to distinguish riders, characterize them in light of helmet usage, and perceive their number plates utilizing OCR procedures. A commented- on dataset with pictures of riders with and without helmets is utilized for preparing the model. The framework accomplishes certainty scores of 88% for riders, 85% for helmets, and 80% for number plates. It recommends future improvements in OCR motor turn of events and UI security. This approach features the reconciliation of profound learning and web innovations for traffic law enforcement, offering a versatile answer for decreasing street fatalities connected to rebelliousness with helmet usage.

I. Nigel Dale et al. [10]

This study assesses the viability of profound learning models YOLOv5, v6, and v7 in distinguishing individuals wearing safety-helmets, a basic part of personal defensive hardware on building destinations. Using the Google Collaboratory notebook with a Tesla T4 GPU, the models are prepared and tried on a dataset with explanations for ”helmet,” ”person,” and ”head,” among others. The performance of these models is assessed based on their mean Average Precision (mAP), with YOLOv7 outperforming the others by achieving the highest mAP of 89.6%. The findings suggest that the YOLOv7 model, despite its longer training time, offers superior accuracy in safety helmet detection, making it a promising tool for enhancing workplace safety. Future work includes refining the dataset and exploring model pruning’s impact on performance.

III. PROPOSED METHODOLOGY

A. Creating custom dataset

The dataset for detection of helmets with camera attached was not available. We wanted the model to very precisely close to real world scenarios. So we clicked photos in roads (as shown in Fig.1.) where we recreated a real world traffic where two wheeler riders are breaking these motor vehicle rules. We included the cases:

- Riders not wearing helmets.
- Riders wearing helmets, but no camera attached to it.
- Riders wearing helmets, with camera attached to it.

We collected a total 1200+ images which resulted in our custom dataset. We also maintained a strict ratio of the three scenarios given above to create a perfect learning spectrum for the model.

B. Image labelling

We used the LabelImg tool to label a collection of images by drawing bounding boxes around the objects of interest, specifically identifying classes such as No Helmet, Helmet and Camera. The annotations are stored in XML format, using the YOLO format, and were also made compatible with YOLO v8 and Create-ML formats. The labelled image dataset was organized into training, testing, and validation subsets, with proportions of 60%, 30%, and 10% respectively, to facilitate the model’s learning and prediction accuracy. Bounding boxes help visually identify the different classes in an image, making it easier to determine whether a traffic rule has been violated based on the classified images.

Their ability to capture and process visual information makes them essential in modern computer vision applications.

Fig. 1. **Riders that are 1.Wearing a Helmet, 2.Wearing a helmet with a camera attached, 3. Not wearing a helmet**

C. CNN

Convolutional Neural Networks (CNNs) assume a basic part in the undertaking of item identification, which includes distinguishing and finding objects in pictures or recordings. The consideration of pooling layers in CNNs is pointed toward down-sampling highlight guides and diminishing their spatial aspects. In object identification undertakings, CNNs frequently use anchor boxes to anticipate and refine bounding boxes. Generally speaking, CNNs act as incredible assets for object identification by autonomously learning various levelled highlights, overseeing spatial varieties, and considering start- to-finish preparation in both confinement and characterization undertakings. Their capacity to catch and deal with visual data makes them fundamental in present-day computer vision applications.

D. YOLO v8

YOLO is a series of evolving models widely utilized in computer vision for object detection. The latest version, YOLOv8, expands on the accomplishments of previous versions, improving both precision and processing speed for detecting objects. The excursion began with YOLOv1 and has since developed through progressive iterations, each addressing previous limitations and boosting performance. YOLOv8 addresses the cutting-edge advancements in object detection innovation. Like its ancestors, YOLOv8 succeeds in real- time object detection, making it ideal for applications that demand speedy and accurate analysis. It gauges the model’s overall precision in detecting and categorizing objects across various classes. For instance, a mean Average Precision (mAP) of 90% connotes that the model consistently distinguishes and categorizes objects with 90% accuracy. To register mAP, we initially calculate the Average Precision (AP) for each object class, which involves averaging the precision-recall bend across all Intersection over Union (IoU) thresholds from 0.5 to 0.95. YOLOv8 Loss Function: The total loss, L total, is often a sum of individual losses:

$$L_{total} = (\lambda_{coord} L_{coord}) + (\lambda_{obj} L_{obj}) + (\lambda_{class} L_{class}) \quad (1)$$

Where:

- L coord is the bounding box regression loss.
- L obj is the objectness confidence loss.
- L class is the class prediction loss.
- Lambda coord is the weight for the bounding box regression loss.
- Lambda obj is the weight for the objectness confidence loss.
- Lambda class is the weight for the class prediction loss.

Coefficients L coord, L obj, L class are typically used to balance the impact of each loss term.

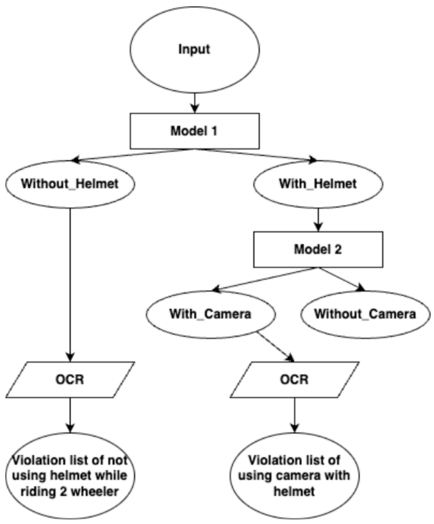

Fig. 2. **Architectural Diagram**

E. Training

As represented in Figure 1, all pictures have now been categorized and labelled in like manner.

In this manner, we partitioned these pictures into two unmistakable gatherings: one for training and the other for testing and validation of our model. This approach ensures that our model is trained on a diverse set of data, enhancing its ability to accurately classify new images in real-world scenarios. The dataset is split into a ratio of 60%(Training), 30%(Testing), 10%(Validation). The Train set of data is used to train the model for the purpose we require it. Test set of data is used to test the model's detecting and validation set of data is used to determine how precise the prediction has been done.

F. Process

The Input images is taken into YOLO Model-1 as shown in Fig 2, which detects if the image contains a biker that either wears a helmet or not. The prediction classifies images into two groups, With helmet and Without helmet.

Images classified as "Without helmet" is considered as offence and these images are saved into a separate directory named the same. Further the offenders are identified by detecting their number plates using a OCR. The list of offenders is saved.

YOLO Model-2 is made to run images that are classified and saved under With helmet, so the model can classify its prediction under two classes. With Camera and With- out Camera.

Images under the classification of "With Camera" is identified as the second offence and images are saved into a separate directory named the same. Then these images are loaded into the OCR for vehicle identification.

Both the results of OCR model is exported in csv file format and then any duplicates in vehicles are removed for each file.

From here we check the accuracy of the model. And later we use the images stored in the folder to assign penalties.

IV. Result Analysis

The model provided a result shown in [Fig. 3]. This image contains a helmet(with and without camera attached). Here each image is processed and detected accurately. Images that detected will be stored to respective directories after classification. Which is further used for detecting number plates using a OCR and assigning penalties.

[Fig. 8] displays the performance of a classification system in a normalized format, where each cell represents the pro- portion of predictions within each actual class. The matrix is divided into four categories:

Predictions for "Camera" are perfectly accurate (1.00), indicating no misclassifications. Predictions for "Helmet" show some confusion with "NoHelmet" and "Background", indicating some instances where helmets were misclassified as not being helmets or as background. No Helmet with only 33% correctly identified, and significant confusion with "Background".

[Fig. 6] graph shows the relationship between precision and confidence for the detection of various classes (Camera, Helmet, NoHelmet) over different confidence thresholds:

As confidence increases, precision also generally increases, which is expected as higher confidence usually corresponds to more accurate predictions. The precision for "Camera" and

”Helmet” starts high and remains relatively stable across confidence levels, indicating robust detection capabilities. In contrast, ”NoHelmet” shows lower precision, suggesting difficulties in accurately identifying instances without helmets,

possibly due to fewer features distinguishing this class from the background.

[Fig. 7] plots illustrates the trade-off between precision and recall for different classes at varying thresholds:

The model performs exceptionally well for ”Camera” and ”Helmet”, maintaining high precision across most recall levels. ”NoHelmet” has a lower precision, particularly as recall in- creases, which may indicate a higher number of false positives for this class. The curve for ”all classes” indicates a balanced performance across the model, with a mean Average Precision (mAP) of 0.941 at an Intersection over Union (IoU) threshold of 0.5.

Fig. 3. Result: Detects helmet with camera attached

Fig. 4. Result: Detects multiple riders not wearing helmets

Fig. 5. Result: Detects multiple riders wearing helmet

TABLE I. PRECISION

CLASSES	PRECISION
CAMERA	0.971
HELMET	0.979
NO HELMET	0.874

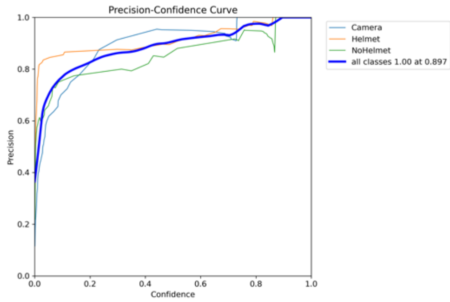

Fig. 6. Precision-Confidence curve

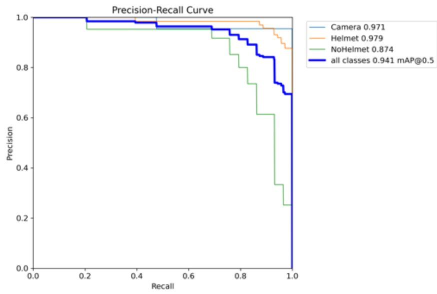

Fig. 7. Precision-Recall curve

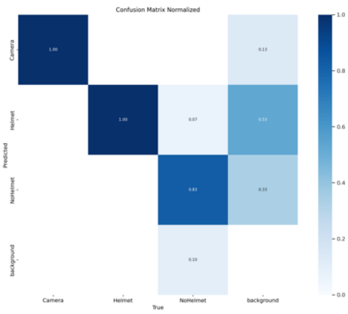

Fig. 8. Confusion Matrix

V. CONCLUSION

Our model showed an accuracy of 98% and precision of 99%. These statistics proves the practicality of the model and how well the model perform under a real-world scenario. Furthermore, this system can be deployed on a larger scale and help our law and order system to enhance road transport environment.

Our system has only solved some of the many missing entities of our existing system. Our options are limitless and can be improved for automating detection of more violations in our world. There are more efficient add-on systems like heat-image sensor, weather detection system that can be used to teach the model of the changing environment and how it affects road transport. Adapting to all scenarios provides better results.

REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] N. Kharade, S. Mane, J. Raghav, N. Alle, A. Khatavkar and G. Navale, ”Deep-learning based helmet vio- lation detection system” 2021 International Confer- ence on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021, pp. 1-4, doi: 10.1109/AIMV53313.2021.9670937.

[2] Ghahremannezhad, Hadi, Hang Shi, and Chengjun Liu. ”Object Detection in Traffic Videos: A Survey.” IEEE Transactions on Intelligent Transportation Systems (2023).

[3] ShreyasRKumar,B.V.Pradeep,AdithyaH.B,Padmaja B and Sunil M. P. “Dynamic traffic rule violation mon- itoring system using auto- matic number plate recogni- tion with SMS feedback”. 2nd International Conference on Telecommunication and Networks August 2017.

[4] Vardhan, Rongali Lalith, et al. ”Helmet Violation and Number Plate Detection System.” (2023).

[5] Wang, Zizhen, et al. ”Higher Accuracy Yolov5 Based Safety Helmet Detection.” International Conference on 5G for Future Wireless Networks. Cham: Springer Na- ture Switzerland.

[6] Afonso, Matheus HF, et al. ”Vehicle and Plate Detection for Intelligent Transport Systems: Performance Evalua- tion of Models YOLOv5 and YOLOv8.”

[7] Yonten Jamtsho, Panomkhawn Riyamongkol, Rat- tapoom Waranusast. ”Real-time license plate detection for non-helmeted motorcyclist using YOLO”, ICT Ex- press, 2020.

[8] Armstrong Aboah, Bin Wang, Ulas Bagci, Yaw Adu- Gyamfi. ”Real-time Multi-Class Helmet Violation De- tection Using Few-Shot Data Sampling Technique and YOLOv8”, 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2023

[9] MaheswaranS,GomathiRD,SatheshS,DeepanKumar S, Murugesan G, Prakash Duraisamy. ”Real-time Imple- mentation of YOLO V5 Based Helmet with Number Plate Recognition and Logging of Rider Data using PyTorch and XAMPP”, 2023 14th International Con- ference on Computing Communication and Networking Technologies (ICCCNT), 2023

[10] Nigel Dale Then Yung, W. K. Wong, Filbert H. Juwono, Zee Ang Sim. ”Safety Helmet Detection Using Deep Learning: Implementation and Comparative Study Using YOLOv5, YOLOv6, and YOLOv7”, 2022 International Conference on Green Energy, Computing and Sustain-able Technology (GECOST), 2022

[11] Nife, Nadia Ibrahim, and Mohammed Chtourou. ”Im- proved Detection and Tracking of Objects Based on a Modified Deep Learning Model (YOLOv5).” Interna- tional Journal of Interactive Mobile Technologies 17.21 (2023)

[12] M. Mohith K. Reddy, D. Rohith, V. Krishna, C. Reddy, A., and Mamatha, I., “Smart Helmet using Advanced Technology”, Information and Communication Technol- ogy for Intelligent Systems. Springer Singapore, Singa- pore, pp. 479–488, 2021

[13] Sunil, S., Sonu, P. (2022). An Effective Approach for Detecting Acute Lymphoblastic Leukemia Using Deep Convolutional Neural Networks. In: Mandal, J.K., Hsiung, PA., Sankar Dhar, R. (eds) Topical Drifts in Intelligent Computing. ICCTA 2021. Lecture Notes in Networks and Systems, vol 426. Springer, Singapore.

[14] Karthika, R. and Parameswaran, L., 2022. A Novel Convolutional Neural Network Based Architecture for Object Detection and Recognition with an Application to traffic sign recognition from Road scenes. Pattern recognition and image analysis, 32(2), pp.351-362

[15] S. Prabhakaran, R. K. Menon and R. V. U, ”Auto- mated Non-Helmet Rider Detection using YOLO v7 and OCR for Enhanced Traffic Monitoring,” 2023 4th IEEE Global Conference for Advancement in Tech- nology (GCAT), Bangalore, India, 2023, pp. 1-6, doi:10.1109/GCAT59970.2023.10353406.

[16] U. Subbiah, Kumar, D. K., Dr. Senthil Kumar T., and Parameswaran, L., “An Extensive Study and Comparison of the Various Approaches to Object Detection using Deep Learning”, 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT). 2020.

[17] B. N. Krishna Sai and Sasikala T, “Object Detection and Count of Objects in Image using Tensor Flow Object Detection API”, in 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), 2019.