*International Journal of Research in Engineering and Science (IJRES)*
*ISSN (Online): 2320-9364, ISSN (Print): 2320-9356*
*www.ijres.org Volume 11 Issue 5 ǁ May 2023 ǁ PP. 80-87*

# Helmet detection system using Mask R-CNN

Dewashish Tarale
*Department of Computer Science and Engineering specialization in Software and Engineering*
*SRM Institute of Science and Technology Chennai, India*

Parth Chauhan
*Department of Computer Science and Engineering specialization in software and Engineering*
*SRM Institute of Science and Technology Chennai, India*

***Abstract***—
*Today in India as the industry and economy are growing. The increase of usage of two vehicular vehicles is increasing day by day. Many start-ups have emerged whose primary work is to deliver food items and other essential things and offer these two vehicular systems. Accidents are increasing daily due to not wearing helmets and family losing their loved ones due to negligence. This project presents a mechanized system to recognize motor bikers without a helmet.*
*In this process we are using mask RCNN method to detect helmets. Mask R-CNN is a state-of-the-art model for instance segmentation, developed on top of Faster R-CNN. Faster R- CNN is a region-based convolutional neural networks that returns bounding boxes for each object and its class label with a confidence score.*

---

Date of Submission: 19-04-2023 Date of acceptance: 03-05-2023

---

## I. INTRODUCTION

According to the Ministry of Road Transport & Highways [1], nearly 28 two-wheeler riders died every day in 2016 due to a lack of helmet use. In 2017, 31 people per 100 were killed in vehicle accidents, up from 21.6 deaths per 100 in 2005. Every year, 14 Lakh Indians suffer from traumatic brain injury (TBI). TBI takes the lives of over 50,000 people per year. Care for these accidents' costs almost 76.5 billion dollars, i.e., 5.6 Lakh crore rupees.

In developing countries like India, a two-wheeler is a popular means of transportation and is extensively used by the people. Many start-ups have emerged in this period. Motorcycles are used by many people around the city, such as Rapido, Swingy, Zomato and many other start-ups. India is a developing country. It has become crucial to wear helmets. Despite governments making efforts and making it illegal to ride a bike without a helmet and have implemented enforcement methods to apprehend violators. People don't follow the rules and get into accidents by not following directions. There is a reason why accidents have increased overall in the country.

We have seen this information, so we decided to contribute towards this project. In the interest of public safety, a system for automated helmet identification that can detect anyone not wearing a helmet on the road is required. This type of system will help the officials who are wearing the helmet or not and impose fines who commit the mistake by two-wheeler users.

In these project we are going to use mask RCNN but lets first understand how the faster R-CNN works. These works in two stages.

Stage1: It consists of two networks and region proposal network.We have to give one input at a time to get a set of region proposals and region proposals are regions in the feature map contain the object.

Stage 2:

In the second stage, the network predicts bounding boxes and object class for each of the proposed region obtained in stage1. Each proposed region can be of different size whereas fully connected layers in the networks always require fixed size vector to make predictions. Size of these proposed regions is fixed by using either Rol pool or RoIAlign method.

## II. Literature Review

For many years, researchers have been striving to solve the challenge of helmet detection from surveillance videos.

These are the papers from videos and images.

Kumar and his team used various techniques to detect objects in images and videos. Extraction and recognition of a particular thing were achieved through the process of Object Recognition with features like category recognition and its detection along with a few algorithms. Further pre-processing takes place. They are pre-processing enhanced the extracted features by implementing methods like resizing the image and normalizing the contrast and brightness effects. Edge Detection implies the feature extraction process that retains only the essential elements. This reduced part is termed a feature vector. Detection in images is accomplished through the concept of bounding boxes. The confidence score of each bounding box provides the certainty that the box encloses an object. CAL VOC 2007 dataset includes images of nearly 20 different classes of images.

Prajwal and his team proposed a computer vision model that would solve the problem of detection of motorcyclists violating helmet rules. It consists of a model including Convolutional Neural Networks (CNNs), image processing, and classifiers such as SVM and HOG to solve the problem. The R-CNN architecture has higher accuracy for the detection of objects but contains a few disadvantages, such as high training costs and memory consumption. Faster R-CNN was thus used. The input fed into one of the CNNs for feature extraction is the entire image in Faster CNNs. The SoftMax layer replaces the set of SVMs used for classifications. In Single shot multi-box detectors (SSD), the type and selection of region proposals are made simultaneously. The processing speed is increased when compared to R-CNN. SSD omits a filtering step, which creates many boxes that do not have objects. Thus, the probability of object presence and spatial shift of each box is performed, and bounding boxes that are true compared to the ones predicted and the false packages are deleted. The result of comparing the three neural networks states that Faster R-CNN's accuracy is the highest compared to the others but requires a longer time for image processing. SSD is used for real-time image processing, but it does not guarantee a result with higher accuracy.

Felix Wilhelm Siebert and his team proposed a deep learning- based method to automatically perform three elements of human observer motorcycle helmet use registration, i.e., detection and tracking of active motorcycles, as well as identification of rider number per motorcycle, rider position, and rider-specific helmet use. The results show high accuracy of their approach, i.e., around 95.3%. The results of our ablation study show that our approach achieves a comparatively high accuracy against ablation experiments. While this high accuracy comes at the expense of computational efficiency, our system can process more than 8 FPS on consumer hardware, which is close to real-time speed for 10 FPS video data. Overall, our work shows that all four essential elements of helmet use registration through human observers can be implemented in a CNN-based approach that is computationally efficient on consumer hardware. There are some limitations to their work; they had comprised when dealing with everyday traffic environments or street scenes with parked motorcycles. Hence, the current approach is partly constrained in real-world applicability, as observation site-specific elements could decrease detection accuracy. And while the HELMET dataset is a first step towards using more diverse datasets for developing automated helmet detection approaches, further data needs to be collected to make methods universally applicable.

Pu Yan and his team proposed an algorithm based on face localization which can detect helmets under different conditions. They used the Haar-AdaBoost algorithm to locate the face, which avoids the interference of background environment on helmet-wearing detection. Secondly, the potential relationship between the helmet and the face. Thirdly the colour the shape feature is used to detect the wearing status of the helmet. Finally, the shape feature is used to further eliminate the colour interference similar to colour of the helmet.

Vignesh and his team proposed their algorithm using single shot detector (SSD). Their paper described a method for solving the problem using the SSD model. SSD model is a mainstream algorithm for object detection. It does both image segmentation and image classification simultaneously by running at a single runtime. The SSD method will separate the images. This model will access the bounding box region of the motorcycle and the rider with just one unit. After the area has been picked, the proposed model will classify whether the biker is wearing or not wearing a helmet in real-time. The Convolutional Neural Network is used to recognize motorcycles without helmets and to distinguish motorcycles among moving targets. The results show that the integrated target detection network increases the reliability of identifying safety helmets, which can provide better comparative results for security surveillance systems, with an accuracy of around 94.65%. We'll use neural network models and frontend video capture plugins to create a more effective protective helmet recognition system.

In this paper, Sushant and his team used YOLO v4 in this proposed model. YOLO is a model trained on COCO dataset and capable of detecting 80 classes. The YOLOv4 is based on the convolution neural network to the entire input image. It divides it into SXS YOLO is a model trained on the COCO dataset capable of detecting 80 classes. YOLOv4 performs well for detection and gives decent precision. The number plate of the motorcyclist not wearing the helmet was successfully extracted. The model can be further improved by increasing the dataset. Different images in different environments should be added to the dataset to increase detection accuracy further.

Gaurav Patil and his team did helmet detection using Detectron2, and Efficient applied their knowledge of the state- of-the-art models Efficient and Detectron2 for solving this helmet detection problem. The proposed method can be used as a detection algorithm in a surveillance system on roads with significantly less prediction time. This paper does not cover a real time use case using these models.

Sanjay Chatterjee and his team proposed an approach to detect single or multiple riders on a motorcycle. The First YOLOv3 model is used for motorcyclist detection. Then, the proposed lightweight convolutional neural network sees the wearing of the helmet or no helmet for all motorcycle riders. The proposed model works quite well for helmet detection in different scenarios, with accuracy of 96.23%. The architecture used here performs comparably well with other CNN-based helmet detection and can be extended to detect more complicated cases of multiple riders, including child riders. Further, the work on this paper can be extended to even more complex scenarios of bad weather for the detection of helmetless motorcyclists. There were limitations in this paper.

Yaqui Guan and his team proposed this paper proposes a helmet-wearing detection method using an improved YOLOV5 model. This method mainly performs input size, initial candidate frame adjustment and loss function on the original YOLOV5 model. The improvement of YOLOV5 makes the YOLOV5 model more suitable for recognizing wearing helmets. It combines the trained model with the camera and other hardware to design a complete real-time detection system. During training, the size of the input image is changed to increase the model's adaptability, and the hyperparameters and optimizer are adjusted to be the best after improvement. The detection model has an accuracy rate of 90%, and the detection speed has reached 37.8fps, which meets the requirements of real-time detection of helmets. Through this model and hardware, such as cameras, real-time detection of whether a person wears a helmet is designed and implemented. The system realizes the three functions of picture detection, video detection and real-time monitoring.

The model that they've used is the Faster R-CNN with Feature Pyramid Network (FPN) which is considered to be the basic Bounding box detector. Here , Helmet Detection was done using Detectron 2 and EfficientDet. Basically depending on the number of epochs and the training time, each model got a different accuracy after training on the given dataset. The inferences of both the EfficientDet and the Detectron2 can be seen with the overall loss.Art models like Detectron2 and Efficient Det to solve the problem of detecting helmets on bike riders.The FPN backbone makes the model a multi-scale detector due to which it yields high accuracy for small as well as large objects which makes it very powerful.The proposed method can be used as a detection algorithm in a surveillance system on roads with very less prediction time. Feature of P6 was not able to passed using non-max suppression but all the other features passed to the box head which outputs 100 boxes.

In this paper Lokesh and his team propesed Helmet Detection using Machine Learning and Automatic License Plate Recognition where they were able to deliver not only obtained and persistant images but also used mobiles. In their proposed methodology Mobiles can be used as an webcam for input, OCR model was able to detect helmet in image with accuracy up to 85%. All the libraries and software used in this project is open source and hence is very flexible and cost efficient.Their method was not applicable to real world issue that live readings as it is only applicable for images. Also , it has Bad performance when there are groups of small objects, because each grid can only detect one object.

Xitian Long and his team from the Zhe Zheng State Grid Key Laboratory of Power Industrial Chip Design and Analysis Technology Beijing Smart-Chip Microelectronics Technology proposed safety Helmet Wearing Detection Based On Deep Learning where they were able to determine Data Acquisition helps to prepare a relatively larger dataset (>5000 images) . The system also achieves real-time speed of 21fps , which is better than most of the present protocols. It was observed SSD trains faster and has swiffer inference than fast RCNN but it doesn't involve cases when the person is holding the helmet, therefore not able to comprend the problem statement which we have deliverd in our proposed methodology.

In this paper detection of helmets on motorcyclists was done using Image Descriptors and Classifiers. They proposed a methodology which uses computational vision for the detcetion of helmets use by motorcyclists on public roads later divided into two different modules which are: vehicle segmentation and classification, and the detection of helmet use.First detection of object was done by moving pixels then the segmentation of moving objects was done on the scene the evaluation of the objects of interest in the image. The method of background subtraction receives all points that were altered in a scene. In the proposed system, the task of classifying the vehicles consists of differentiating the segmented objects into two classes: motorcycles and non-motorcycles.

Therefore,HOG algorithm was used which helped the feature extraction that can be done without an edge position foreknowledge The ML classifier using the HOG achieved the very best results with accuracy rate of 0.9137 but the proposed system do not detect more than one helmet in an image. Since a passenger without helmet is a traffic violation also, Quality of Image was compromised.

In this paper Rao Cheng , Zhonglong Zheng and Zhentao Wang carried out various Ablation Experiments ,result analysis and Training settings to prove that the improved algorithm is more suitable for natural complex scenes in terms of accuracy. So For small-scale objects, occluded objects and dense objects, SAS-YOLOv3-tiny is superior to the YOLOv3-tiny algorithm. As it can be seen from the first and second set of images, SAS- YOLOv3-tiny and the latest YOLOv4-tiny can detect all objects, but YOLOv3-tiny may leave out an ordinary object which isn't as big. As can be seen from the remaining set of images, SAS-YOLOv3-tiny can detect all objects while YOLOv4-tiny neglects a helmet object. Also in this paper they were able to solve the problem that the original lightweight algorithm has low accuracy than Mask RCNN. Even though with faster speed and some more few parameters ut can be seen that the accuracy of detection of objects espically minor can be improved.

The main objective of this paper was to construct an automatic detection of the motorcyclist withuout a helmet from video using a YOLOV4 trained model.The following proposed system intends to help the traffic police department throught the city to detect multiple bike riders who are not wearing helmet from traffic surveillance videosand to assist the traffic authorities. This is done by passing the surveillance video through multiple Machine Learning Models (i.e Rider etection Model, Helmet detection Model face cover.Since covid19 regulations people started wearing masks.Their proposed model was able to deliver Number Plate detection Images Average accuracy :100% The model detected the number plates in the images with 100% accuracy.There were no false negatives.Around 60 -70 images were tested but the model detected numberplate accurately OCR conversion Images Average accuracy ~ 98%. The accuracy was affected as some riders didn’t had numberplates and some were having numberplates written in design in such a way that human also cannot tell the number in one go.

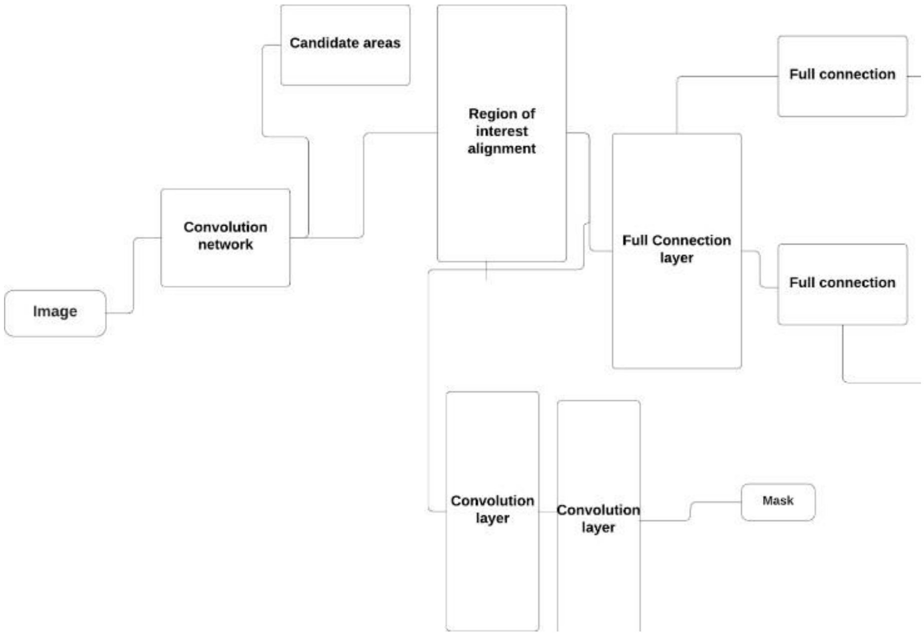
Miaomiao Chen and his team from Pioneer Colledge Inner Mongolia University Inner Mongolia University Inner Mongolia Hohhot, Chin proposed a hardware environment of Mask RCNN target decision experiment which uses CPU in this paper abd GPU uses Nvidia GTX970, which has a processing speed upto 8GB and the software environment is Windows10For Mask RCNN target detection. In each of these sample images which were tested for an average of 30 seconds. Here, it was observed the Mask RCNN target detection,when the highest accuracy is 100%, the recall rate is 20%. AP value reached 98%. Hence it is proved that Mask RCNN target detection algorithm is more accurate in target positioning and detection accuracy, and detection speed is also quite fast.

In this Paper Songbo Chen and his team from Zhejiang University College of Electrical Engineering The Improved Faster R-CNN algorithm for feature extraction is based on the ResNet-101 where they used Feature extraction network giving right after a convulation layer,one batch normalisation layer and one ReLU layer, and one max-pooling layer respectively. On the side note instead of traditional sliding window and Selective Search method,RPN will be used so it can generate region proposals directly, as it improves the generation speed of region proosals.Since no public data set has been established in the research on helmet wearing detection. The data used in this experiment was collected from the Internet , from a substation/subway which is located in Qingyuan City on-site, with a total of 1065 images, including images with and without helmets of various background and various other cases which may include small and big objects , day and night,different construction sites, substations and other places.0

**Comparison table:**

| Paper Title | Accuracy | Technique |
|---|---|---|
| Method - I A Review on Helmet Detection by using Image Processing and Convolutional Neural Networks | 92% | Image Processing, Convolutional Neural Networks and Support Vector Machine |
| Method - II Detection of Motorcyclists without Helmet in Videos using Convolutional Neural Network | 92.87% | Gaussian mixture model and Convolutional Neural Networks |
| Method - III Automatic Helmet Detection System on Motorcyclists Using YOLOv3 | 80% | YOLOv3 and Region of Interest |

**System architecture diagram:**



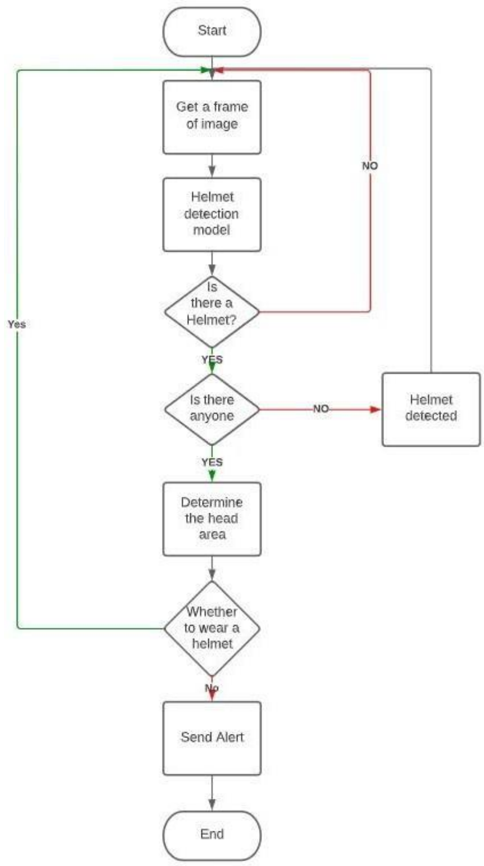

In this diagram we import the image from the dataset of any size the overall features of the image are extracted through the combination of Residual Neural Network (ResNet) and Feature Pyramid Network (FPN). The feature map of the image is then summarized by the feature pyramid network.

Then the next step is to recommend regions of interest through the region generation network. Each picture can have N regions of interest. ROI align method is used to correct the pixels corresponding to the feature map of input image to obtain the candidate frame feature map of each image. Finally, classify and regress the candidate frame targets according to the features on the feature map, perform a fully connected network operation in each candidate area and finally generate a mask for detection.

It is evident from this architectural diagram that we are taking data from the data set and passing it through text preparation. This could consist of bag-of-words, vectorizer, TFIDF, etc. And then separating it into training and testing datasets.

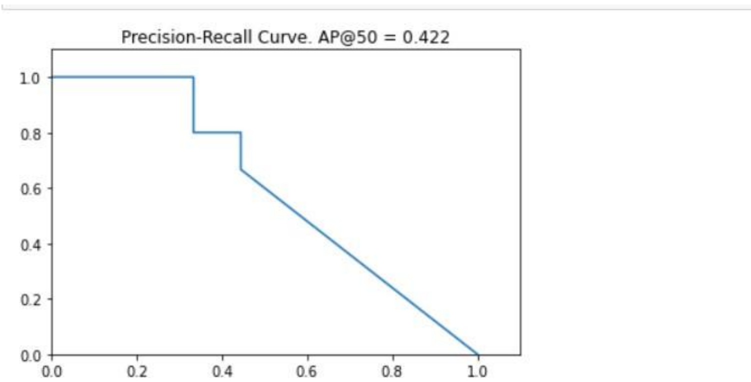
Now, the training datasets are processed by our ML algorithms. Additionally, we adjust the parameters to produce better outcomes, ultimately selecting the optimal one. After completing these steps, we compare our results to the test datasets and attain the needed accuracy.

## III. Proposed methodology

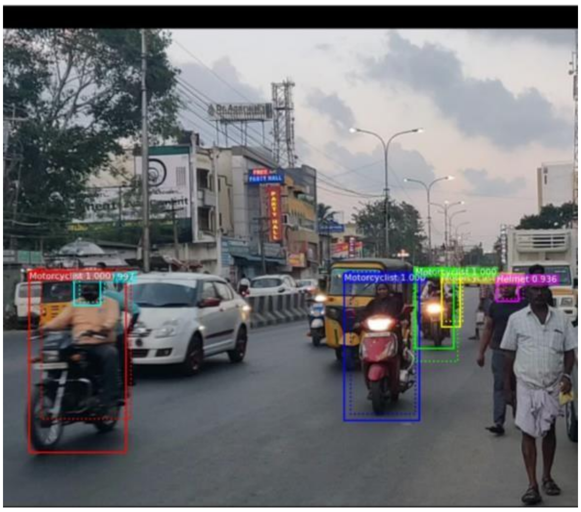


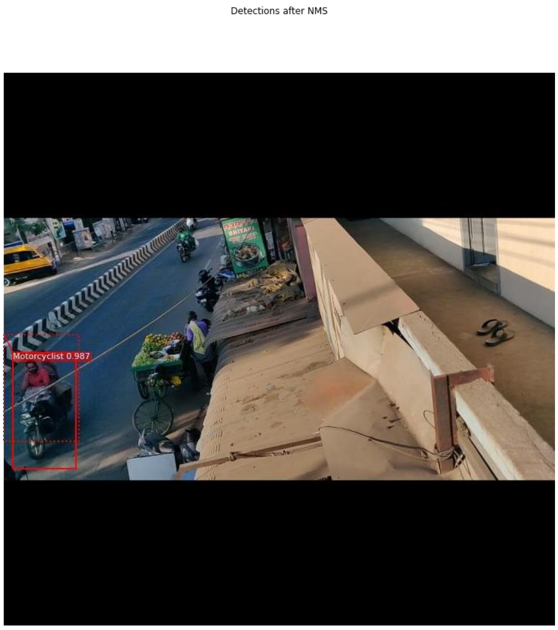

Flowchart diagram we import the image and run through helmet detection model and classify the vehicle first then we detect if there is any helmet or not then we detect is there anyone or not. After that we detect the head area and check whether he is wearing the helmet or not after that we bound the box and give confidence score of the head area**.**

## IV. Results and discussion:

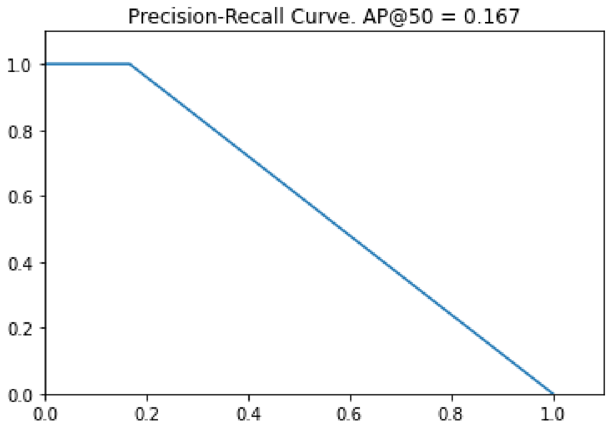



**Precision** recall curve with map value is 0.422.
We were able to detect the Images and helmet through our machine learning algorithm.

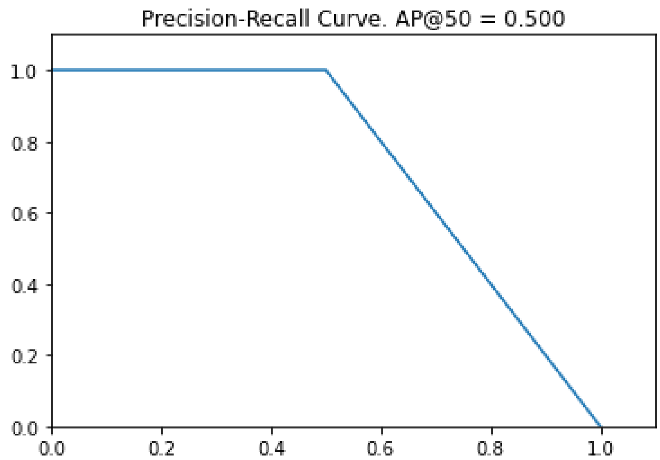





**Graph for this image**




**Precision recall curve for this image:**




From these 3 images were input for the system we were able to predict the helmets on the road.

## V. Conclusion:

From this project we were able to detect the helmet with their confidence score. Through this project we can detect many other objects people who are wearing seat belts or not. This project can be further extended to detect the number plate of the cars or bikes. From this project were able to detect the multiple persons on bike these can help us to detect children who are hidden away by their parents when they sit on the bikes.

## References


[1]. **Object Detection and Recognition in Images by** 1) Sandeep Kumar, 2) Aman Balyan, 3) Manvi Chawla Computer Science &Engineering Department, Maharaja Surajmal Institute of Technology, New Delhi, India.

[2]. **A Review on Helmet Detection by using Image Processing and Convolutional Neural Networks** by Prajwal M., Tejas K. B , Varshad V , Shashidhar R..

[3]. **Helmet use detection of tracked motorcycles using CNN-based multi-task learning. p**roposed by HANHE LIN 1, JEREMIAH D. DENG2, (Member, IEEE), DEIKE ALBERS3, AND FELIX WILHELM SIEBERT 4.

[4]. **Safety Helmet Wearing Detection Based On Deep Learning** by Xitian Long, Wenpeng Cui, and Zhe Zheng.

[5]. **A Helmet wrong detection algorithm based on Face Localization** by PuYan and Chong Chen

[6]. **Automated Helmet detection for Multiple Motorcycles Riders using CNN** by Madhuchhanda Dasgupta

[7]. Oishila Bandyopadhyay Sanjay Chatterji. (Computer Science & Engineering IIIT Kalyani West Bengal, India)

[8]. **helmet Detection using Single Shot Detector (SSD)** proposed by Vignesh Raj A,Manohar and Dhyanjith from Department of Computer science Amrita School of Arts and Sciences, Mysuru

[9]. **Automatic Detection of Bikers with No Helmet and Number Plate Detection** proposed by Sushant Kadam, Rushikesh Hirve, Nikhil Kawle, Payal Shah from the Department of Electronics Engineering, Sardar Patel Institute of Technology Mumbai, India.

[10]. **Design and Implementation of Safety Helmet Detection system based on YOLO v5** by Aqi Guan , Wenqiang Li ,Tianyu Hu, Qun Hou Jianghan from University Artificial Intelligence Institute Wuhan, China houqun@jhun.edu.c

[11]. **Helmet Detection using Machine Learning and Automatic License Plate Recognition** by Lokesh,Allamki1, Manjunath Panchakshari , Ashish Sateesha, K S Pratheek.

[12]. **Safety Helmet Wearing Detection Based On Deep Learning** by Xitian Long, Wenpeng Cui, Zhe Zheng State Grid Key Laboratory of Power Industrial Chip Design and Analysis Technology Beijing Smart-Chip Microelectronics Technology Co., Ltd

[13]. **Helmet Detection on Motorcyclists Using Image Descriptors and Classifiers** by Romuere Silva, Kelson Aires and Rodrigo Veras (Departamento de Computac¸ao˜Universidade Federal do Piau´ı - UFPI Teresina, Brasi.

[14]. **Multi-Scale Safety Helmet Detection Based on SAS- YOLOv3-Tiny proposed** by Rao Cheng, Xiaowei He *, Zhonglong Zheng and Zhentao Wang