YOLO based approach for Helmet and Seatbelt Detection

PREETHI B V[1], PREETHI T[2], SAMEEKSHA B[3], SHARVANI S[4], SUDEEP MANOHAR[5]
[1, 2, 3, 4, 5] *Jawaharlal Nehru New College of Engineering, Shivamogga, Karnataka, India*

Abstract—Increase in the number of transportation vehicles has led to heavy traffic loads and thus many transportation routes are created. Due to this people might lose patience which may lead to many accidents and traffic injuries. Hence rules are enforced upon people to wear helmets while riding bikes and seat belts while driving cars to minimize the chances on casualty in case of accidents. Many people do not follow the traffic rules and get involved in traffic rule violation. Since the monitoring of those rules are done manually by traffic police things might get difficult for the police to monitor each and every vehicle. Hence, a model has been proposed to detect whether helmet or seat belt is worn using the machine learning technique. In this paper, a survey has been done and found that helmet detection model could be built with the help of Convolutional Neural Network (CNN) and You Only Look Once (YOLO)v3. YOLOv2 can also be used which has to be trained with COCO datasets. For seat belt detection YOLOv5 method is used along with the combination of CNN and Support Vector Machine (SVM) models. The proposed model decreases the burden on traffic police and also helps in achieving efficiency.

Indexed Terms—Helmet, Seat belt, YOLOv2, YOLOv3, YOLOv5, Convolutional Neural Network, SVM

I. INTRODUCTION

The report of Mumbai Environmental Social Network in December 2018 has shown that the percentage has been reduced to 2.2% from 6.2% for the criteria of road acquired by buses, where in past twenty years the percentage has been increased to 77% from 5% for the road space acquired by private vehicles in India. According to the World Bank, India is having 1% of automobiles and shares the death rate by 15% of typical overall traffic deaths. To increase the use of helmets and seat belts, Government of India has proposed Motor Vehicles Bill with penalties. A rider who does not wear a helmet or a driver who doesn't wear a seatbelt will be fined 1000 rupees and have his driver's license would be revoked for 3 months. Motorcycle has been a main mode of transportation in a country of big cities with highest population density. Nowadays the heavy load of traffic is due to the increased number of personally owned cars. Due to violation of traffic rules such as not wearing helmet and seat belts, the death rate from accidents is significantly more in metro cities in the past few years. To avoid this situation, it is not possible for traffic police alone to monitor each and every vehicle that passes by. So, there is a need for an automated system that monitors motorcyclists and drivers to detect whether the helmet and seat belt are worn by the rider and driver. Due to the lack of experience and violation of helmet use by motorcyclist and car seat belt use, there is an increase in the accidents and injuries related to road accidents. Government is focusing on the law-abiding behavior and enforcing safety measures in traffics to encounter the problem with the traffic safety [1]. The technique that is used for implementation is Deep Learning which is a subset of machine learning. To perform the process of machine learning, hierarchical level artificial neural networks are used. Machine learning is the field that gives the ability to learn without being explicitly programmed. It is defined as a machine that has the ability to imitate intelligent human behavior. To overcome these problems, it is necessary to design and train a machine learning model, using the YOLOv3 algorithm and CNN algorithm for detection of helmet and seat belt. Sample data, known as "training data", are built by Machine learning algorithms with a mathematical model are used in order to make the predictions of data or decisions that are necessary and also in object detection applications. Therefore, a Helmet and Seat belt detection model can be implemented by training the model with a specific dataset. There is a necessity for controlling and monitoring the use of helmet and seat belt in order to reduce the traffic accidents and its number of injuries and deaths in the country. Much attention has been received by intelligent vehicle monitoring systems based on deep learning [2].

II. OBJECT DETECTION METHODS

YOLO is an algorithm that uses neural networks for real-time object recognition. This algorithm is very popular because of its accuracy and speed. It has been used in various applications for recognizing traffic signals, parking meters, animals and humans, and is capable of recognizing objects in a single algorithm pass. The YOLO algorithm consists of several variants. The most common variants include tiny YOLO and YOLOv3

A. YOLOv2

YOLO version 2 is a network that performs object detection in a single stage. YOLOv2 is much faster than other two-stage deep learning object detectors such as CNN regions.

B. YOLOv3

YOLO version 3 algorithm detects objects in videos, images, or live feeds in real time.

C. YOLOv5

In this algorithm a single neural network is used to process an entire image, then decompose it into parts and predict the bounding boxes and probabilities for each component. This algorithm is trained with the dataset COCO.

D. CNN

The Convolutional Neural Network or ConvNet is a subtype of neural networks used primarily for image and speech recognition applications. Its built-in convolutional layer reduces the high dimensionality of images without losing their information.

E. SVM

The Support Vector Machine is used for both regression problems and classification and is one of the most popular supervised learning algorithms.

III. HELMET DETECTION

Several challenges arise while detecting helmets as there is no standard helmets in the country. To detect the helmet three algorithms are used based on CNN, YOLOv3 and YOLOv2.

A. Helmet detection using CNN : To recognize objects, earlier researchers used conventional image processing methods for feature extraction. Nowadays, CNN is the most advanced method for object recognition. To achieve higher accuracy, CNN automatic feature extraction is the optimal choice. In this paper, a model is proposed for detecting motorcyclists riding a motorcycle with or without helmet [3]. For the analysis of the image, a convolution tool is used to segregate and identify different features, which is called feature extraction. The feature extraction network contains many pairs of pooling or convolutional layers. Pooling is one of the feature commonly used in convolutional neural network architectures. The main aim behind this layer is to collect features from the maps generated by convolution of a filter over the particular image. As shown in Fig. 1 the function of the pooling layer is to gradually reduce its geometric size of representation to deduct the total amount of computations and parameters in the network. Max pooling is the most common form of pooling. To obtain an abstracted form of the representation, Max pooling is performed to avoid over-fitting and to give the internal representation a basic translation invariance and to reduce the number of parameters to be learned, so that the computational cost is also reduced. The results of the convolution process are used by a fully connected layer that predicts the class of an image based on features extracted in the previous stages as shown in Fig. 2. To reduce the number of features which are present in the dataset, the CNN model of feature extraction is used. It generates new features from the existing features contained in the original feature set. Here are several Convolutional Neural Network layers as shown in Fig. 3. of CNN architecture.

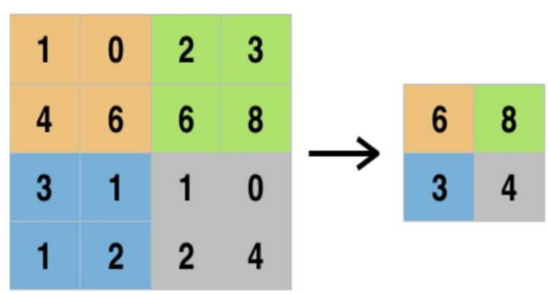

Fig. 1 Geometric size reduction in Max pooling

Layer	Filters	Size	Stride
C1	16	3 × 3	1
Max-pooling		2 × 2	2
C2	32	3 × 3	1
Max-pooling		2 × 2	2
C3	64	3 × 3	1
Max-pooling		2 × 2	2
C4	128	3 × 3	1
Max-pooling		2 × 2	2
C5	256	3 × 3	1
Max-pooling		2 × 2	2

Fig. 2. Layer wise filters in Max-pooling

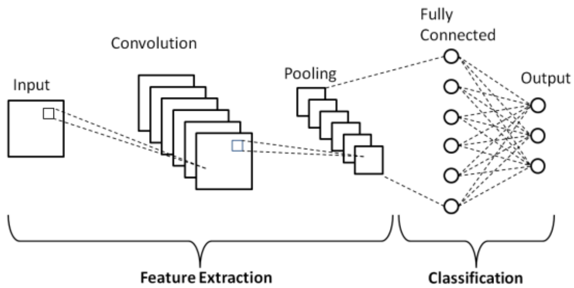

Fig. 3. CNN architecture

In this proposed model, moving objects are detected using adaptive background subtraction method. Adaptive Gaussian mixture model is strongly built to face some of the challenges like occlusion, shadows and other changes and variations. During variable and complex situations solo Gaussian is not sufficient for the pixel so more than one Gaussian model can be used for each pixel. It can be achieved by bounding the objects into a box. CNN model is built, and inputs are passed to separate motorcyclist from other moving objects. CNN learns about all the texture and structure in the training set while the training is done. Next, to detect whether the motorcycle rider is wearing helmet or not the obtained input images are cropped at the top $1/4^{th}$ portion where the head region of motorcyclist is usually present. Later by using binary classification method separation of images with and without helmet are done [4].

B. Helmet detection using YOLOv3: The name YOLO comes from the fact that 1×1 folds are used by prediction, which means that the size of prediction map in the folds is exactly the same as the size of feature map before it. Since YOLOv3 detects the small objects much better compared to YOLO, YOLOv3 is used for detection of multiple riders. YOLOv3 is an improved version of YOLO and YOLOv2. To identify specific objects in images, videos and live feeds, YOLOv3 is used as it detects the object in real time. First, the YOLOv3 algorithm is used to divide an image into grid. Each cell of the grid predicts a certain number of the bounding boxes around objects that have a high score in the predefined classes mentioned earlier.

Each bounding box contains its respective confidence score to assume how accurate this prediction should be, and only one object per bounding box is detected. To detect the most repeated shapes and sizes from the original dataset, boundary boxes are generated by grouping the dimensions of ground truth boxes [5]. Basically, YOLOv3 is used for object recognition and is capable of recognising a large number of classes, including motorcycles and people. To locate the objects, bounding boxes are drawn and the network predicts 4 coordinates. The Euclidean distance between the center of the two bounding boxes of the person and the motorcycle is calculated. The same method can be used to detect multiple riders [6].

C. Helmet detection using YOLOv2: The goal of YOLOv2 is to significantly improve accuracy while becoming faster. For training the classifier, YOLOv2 starts with 224 × 224 images and then trains the classifier again with 448 × 448 images, using much fewer epochs. This makes training the detector easier and map increases by 4 percent.

The proposed approach is based on the YOLOv2 model, which is trained with the dataset COCO and detects the different classes in the input image. Fig. 4 shows the process of helmet detection using YOLO v2. Helmet detection is done in three steps as explained below:

a) Person detection
In the first step, the person in an input image is recognized. This is easily accomplished because the input image is passed through the YOLOv2 model trained on the COCO dataset.

b) Intermediate processing
In this step, all classes except the recognized person are discarded. And then the bounding box of the

person is automatically pruned to be used for the further steps.

c) Helmet detection

Once the person is detected, it will detect the person with helmet and without helmet with the help of trained YOLOv2 model. The person wearing helmet is detected in the cropped image.

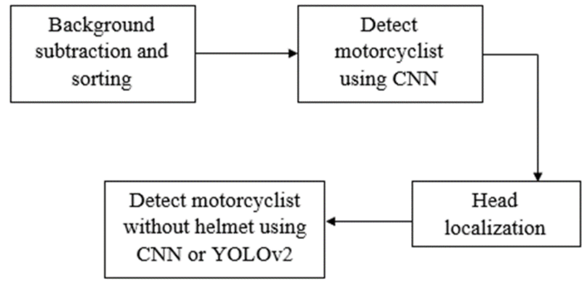

Fig. 4. Process for helmet detection using YOLO v2

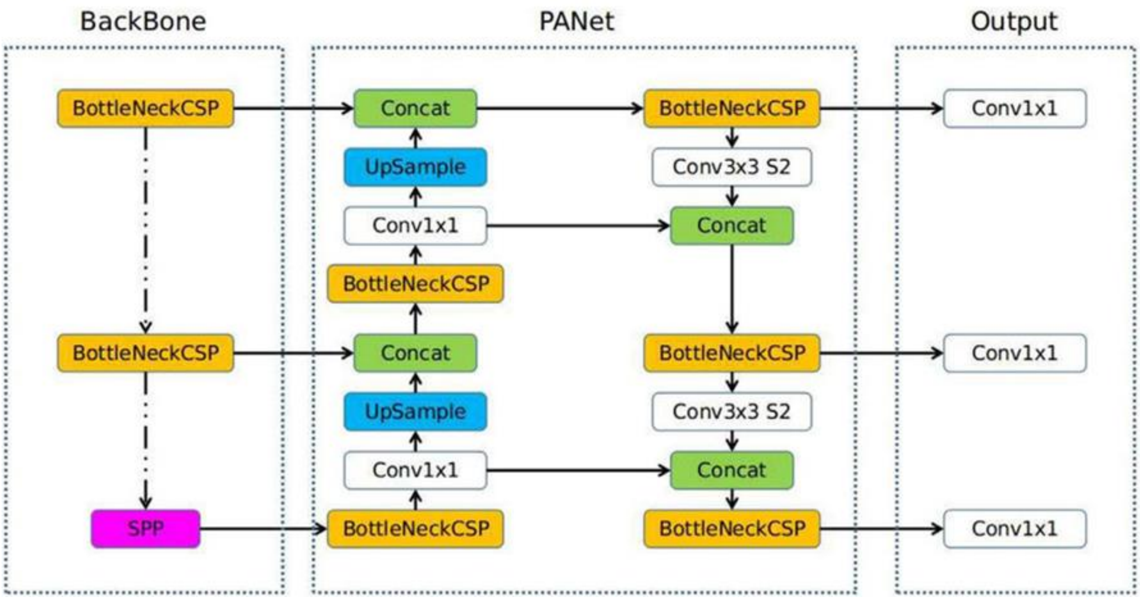

Fig. 5. The network architecture of YOLOv5

IV. SEAT BELT DETECTION

Detecting seat belt is also a challenging task but using the advanced machine learning techniques has made it easier to build applications that can detect them with acceptable accuracy. Seat belt detection is done using two approaches, one using and the other using CNN and SVM models.

A. Seat belt detection using YOLOv5: YOLOv5 is used here to detect seatbelts in cars. According to size of model, the YOLOv5 is mainly divided into four versions which are noted as YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. As shown in Fig. 5, the YOLOv5 network architecture mainly consists of three parts which namely Backbone which comprises of CSPDarknet, Neck which is PANet, and the Head which is YOLO layer. CSPDarknet takes the input first for feature extraction, and then its output is fed to PANet for feature fusion. The results of the detection which are class, score, location, size are finally the outputs of the YOLO layer. In this proposed method the detection of the occupant seats and seat belt of driver is detected using YOLOv5 model. Though the time taken for image recognition is longer the accuracy of the model is higher. The detection process has two steps, namely windshield detection and seatbelt rule violation detection.

a) Windshield detection: In the first step when the vehicle image is given as input, the windshield area of the car is detected by the YOLOv5 model. The windshield region Detection and segmentation is done by the YOLOv5 model that is pre-trained by the dataset of COCO dataset.

b) Seat belt rule violation detection: Here the windshield regions is segmented into two portions vertically where the detection of the violation of seat belt rule will be used from the right part and the detecting of passenger will be the windshield's left part. Here for the seat belt rule violation detection only the right part is used whereas to detect the passenger seat occupancy is done with the windshield left part which is a future work. The process of seat belt detection is shown in Fig. 6.

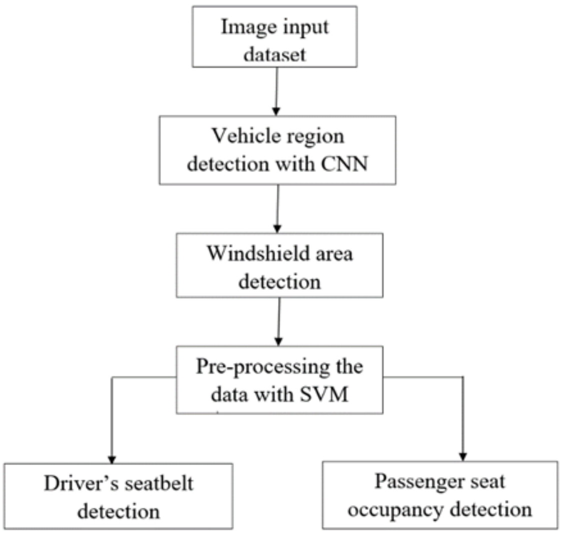

Fig. 6. Process for seat belt detection

B. Seat belt detection using CNN and SVM : SVM is used for regression and classification. To train the models used for recognition using CNN, the multi-level features are extracted from the labeled vehicle, windshield, and seat belt regions. Then rough candidates are recognized. The support vector machine is trained by the relative positions along with the detection scores of those vehicle components, where a post-processing will be deployed to get detection results which are more accurate. The seat belt region is finally detected using this classification model which will be passed under the fine mapping process along with the process of identification.

V. RESULTS

Fig. 7 shows the result of helmet detection module. Motorcylists are segmented and the head region is checked for the presence of helmet. And Fig. 8 shows the result of seatbelt detection module. The left region of the windshield is considered and checked whether the seat belt is worn or not.

Fig. 7. Detecting riders for helmet

Fig. 8. Detecting seat belt

CONCLUSION

The proposed framework for automatic traffic violation detection uses adaptive background subtraction that is insensitive to various challenges such as lighting, poor video quality, etc. Using Deep Learning to automatically learn discriminative representations for classification tasks improves recognition rates and reduces false positives, resulting in a more reliable and dependable system. The YOLOv3 method is the fastest and most accurate method for helmet detection. And YOLOv5 is used to detect seat belt with high accuracy.

REFERENCES

[1] Fahad A Khan, Nitin Nagori, Dr. Ameya Naik, Helmet and Number Plate Detection of Motorcyclists using Deep Learning and Advanced Machine Vision Techniques, ICIRCA, IEEE, 2020.

[2] S Sanjana, V R Shriya, Gururaj, Vaishnavi, K Ashwini, A review on various methodologies used for vehicle classification, helmet detection and number plate recognition, Evolutionary Intelligence, Springer, 2020.

[3] Meghal Darji, Jaivik Dave, Nadim Asif, Chirag Godawat, Vishal Chudasama, Kishor Upla, License plate Identification and Recognition for non-Helmeted Motorcyclists using Light-Weight Convolutional Neural Network, International Conference for Emerging Technology (INCET), IEEE, 2020.

[4] C. Vishnu, Dinesh Singh, C. Krishna Mohan, Sohan Babu, Detection of Motorcyclists without Helmet in Videos using Convolutional Neural Network, Visual Intelligence and Learning Group (VIGIL), Indian Institute of Technology Hyderabad, IEEE, 2017.

[5] Madhuchhanda Dasgupta, Oishila Bandyopadhyay, Sanjay Chatterji, Automated Helmet Detection for Multiple Motorcycle Riders using CNN, Conference on Information and Communication Technology (CICIT), IEEE, 2019.

[6] Shashidhar R, A S Manjunath, Santosh Kumar R, Roopa M, Puneeth S B, Vehicle Number Plate Detection and Recognition using YOLOv3 and OCR Method, IEEE, 2021.