

“HELMET DETECTION AND LICENSE PLATE RECOGNITION USING YOLO MODEL”

Prof. MUNESHWAR R. N.
Miss. POTE PRANAVI VIJAY
Mr. BHAWAR SHIVAM ASHOK
Miss. JADHAV PRATIKSHA SITARAM
Miss. GITE NIKITA RAJENDRA

DEPARTMENT OF INFORMATION TECHNOLOGY
AMRUTVAHINI COLLEGE OF ENGINEERING, SANGAMNER

Abstract

In today's world, the increasing use of Motorcycles has prompted increment in road accidents and injuries. Helmet not used by the motorcycle rider is one of the major causes. Currently, one procedure is to physically check use of helmet at the pavement junction or through the CCTV footage video, which requires human energy to detect motorcyclists without helmet.

The object detection and tracking is the important steps of computer vision algorithm. The robust object detection is the challenge due to variations in the scenes. Another biggest challenge is to track the object in the occlusion conditions. Hence in this approach, the moving objects detection using OpenCV object detection API. Further the location of the detected object is pass to the object tracking algorithm. A YOLO object tracking algorithm is used for robust object detection.

The proposed approach is able to detect the object in different illumination and occlusion. The system extracts objects class based on feature extracted. The system uses You Only Look Once (YOLO)-Darknet deep learning framework which consists of Convolutional Neural Networks trained on Common Objects in Context (COCO) and combined with computer vision.

Keywords: COCO, YOLO, CCTV, OpenCV, Neural Network, video detection ,tracking object, etc

Introduction

Background and Basics

In countries like India, Brazil, Thailand, majority of population uses motorcycles for daily commute. In India, in most of these countries, wearing helmet for motorcyclists is mandatory by law. Also, considering safety of people using motorcycles, wearing helmet is paramount. Currently, in practice, Traffic Police are entrusted with the task of ensuring that motorcycle riders wear helmet. But, this method of monitoring motorcyclists is inefficient due to insufficient police force and limitations of human senses. Also, all major cities use CCTV surveillance-based methods. But those require human assistance and are not automated.

Due to the increasing number of motorcycles and the concern for human safety, there has been a growing amount of research in the domain of road transport. The system proposed in this paper automates the task of monitoring motorcyclists. The system detects motorcyclists not wearing helmets and retrieves their motorcycle number plate in real time from videos captured by CCTV cameras at road junctions by making use of Machine Learning.

The object detection and tracking is the important steps of computer vision algorithm. The robust object detection is the challenge due to variations in the scenes. Another biggest challenge is to track the object in the occlusion conditions.

Hence in this approach, the moving objects detection using TensorFlow object detection API.

Further the location of the detected object is pass to the object tracking algorithm. A YOLO based object tracking algorithm is used for robust object detection. The proposed

approach is able to detect the object in different illumination and occlusion. Usually, frame contains background and foregrounds information. In this foreground object is represented by features points in the ROI and remaining features are considered as background. In general, surveillance system consists of two major steps such as moving object detection and motion estimation. The first step is the object detection and it is influenced by the background pixels information.

Background subtraction is the method which extracts the interested moving object from the video frames. The background subtraction is affected by mostly nonstationary background and illumination changes. To overcome this limitation, in this approach a novel and generalized Tensor flow-based object detection and CNN based object tracking algorithm has been presented. These approaches are robustly detected and track the object in complex scenes and complicated background conditions. Tracking objects is a fundamental problem in computer vision. Traditional methods based on feature, such as those based on color or blobs movement, follow-up maintaining a simple model of the objective and adapting this model over time. However, real situations in practice pose enormous challenges to these techniques.

Tkinter

Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit.

Tkinter provides various controls, such as buttons, labels and text boxes used in a GUI application. These controls are commonly called widgets.

All Tkinter widgets have access to specific geometry management methods, which have the purpose of organizing widgets throughout the parent widget area. Tkinter exposes the following geometry manager classes: pack, grid, and place.

Pillow

Python Imaging Library (expansion of PIL) is the de facto image processing package for Python language. It incorporates lightweight image processing tools that aids in editing, creating and saving images. Support for Python Imaging Library got discontinued in 2011, but a project named pillow forked the original PIL project and added Python3.x support to it.

Pillow was announced as a replacement for PIL for future usage. Pillow supports a large number of image file formats including BMP, PNG, JPEG, and TIFF. The library encourages adding support for newer formats in the library by creating new file decoders.

LabelImg

LabelImg is a free, open-source tool for graphically labeling images. It's written in Python and uses QT for its graphical interface. It's an easy, free way to label a few hundred images to try out your next object detection project. It can be used to create a custom dataset to train our object detection models.

OpenCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

TensorFlow

TensorFlow offers multiple levels of abstraction so you can choose the right one for your needs. Build and train models by using the high-level Keras API, which makes getting started with TensorFlow and machine learning easy.

If you need more flexibility, eager execution allows for immediate iteration and intuitive debugging. For large ML training tasks, use the Distribution Strategy API for distributed training on different hardware configurations without changing the model definition.

YOLO

YOLO is an abbreviation for the term 'You Only Look Once'. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.

This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously. The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO and YOLOv4.

YOLO algorithm is important because of the following reasons:

- **Speed:** This algorithm improves the speed of detection because it can predict objects in real-time.
- **High accuracy:** YOLO is a predictive technique that provides accurate results with minimal background errors.
- **Learning capabilities:** The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.
- **How the YOLO algorithm works?**

YOLO algorithm works using the following three techniques:

Residual Blocks:

- First, the image is divided into various grids. Each grid has a dimension of $S \times S$. The following image shows how an input image is divided into grids.

Fig: 1 Residual Blocks



- In the image above, there are many grid cells of equal dimension. Every grid cell will detect objects that appear within them. For example, if an object center appears within a certain grid cell, then this cell will be responsible for detecting it.

1. Bounding Box Regression:

- A bounding box is an outline that highlights an object in an image.

- Every bounding box in the image consists of the following attributes:
 1. Width (b_w)
 2. Height (b_h)
 3. Class (c)
 4. Bounding Box center (b_x, b_y)

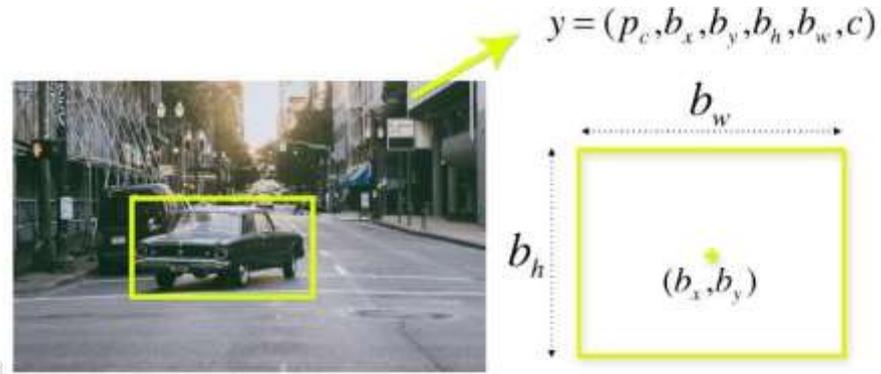


Fig 2: Bounding Box Regression

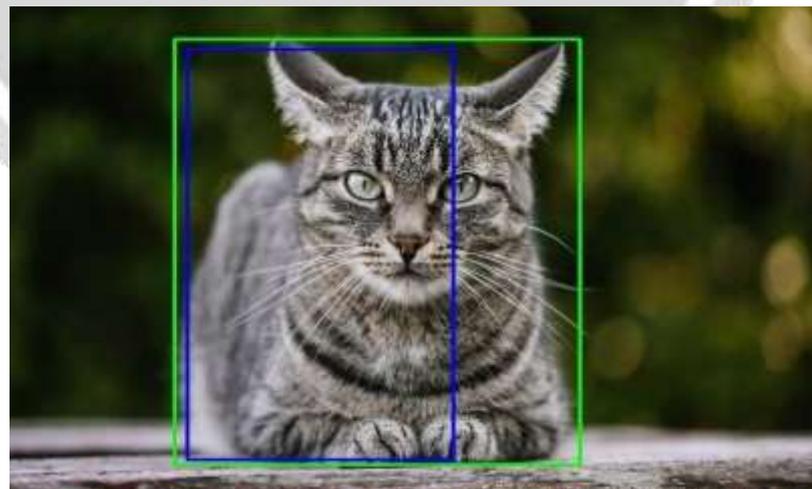
- YOLO uses a single bounding box regression to predict the height, width, center, and class of objects. In the image above, represents the probability of an object appearing in the bounding box.

2. Intersection Over Union (IOU) :

- Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.
- Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

In the image above, there are two bounding boxes, one in green and the other one in blue. The blue box is the predicted box while the green box is the real box. YOLO ensures that the two bounding boxes are equal.

Fig 3: Intersection Over Union



- **Combination of three techniques:**

- The following image shows how the three techniques are applied to produce the final detection results.

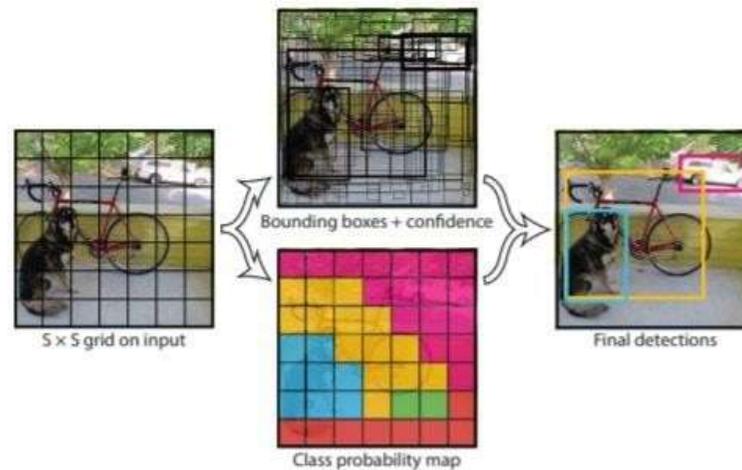


Fig 4: Combination of three techniques

- First, the image is divided into grid cells. Each grid cell forecasts B bounding boxes and provides their confidence scores. The cells predict the class probabilities to establish the class of each object.
- For example, we can notice at least three classes of objects: a car, a dog, and a bicycle. All the predictions are made simultaneously using a single convolutional neural network.
- Intersection over union ensures that the predicted bounding boxes are equal to the real boxes of the objects. This phenomenon eliminates unnecessary bounding boxes that do not meet the characteristics of the objects (like height and width). The final detection will consist of unique bounding boxes that fit the objects perfectly.
- For example, the car is surrounded by the pink bounding box while the bicycle is surrounded by the yellow bounding box. The dog has been highlighted using the blue bounding box.

OCR

Literally, OCR stands for Optical Character Recognition. It is a widespread technology to recognize text inside images, such as scanned documents and photos. OCR technology is used to convert virtually any kind of image containing written text (typed, handwritten, or printed) into machine-readable text data.

Methodology:

- To reduce the noise we need to blur the input image with Gaussian Blur then convert it to grayscale.



2. Find vertical edges in the image.
3. To reveal the plate we have to binarize the image. For this apply Otsu's Thresholding on the vertical edge image. In other thresholding methods we have to choose a threshold value to binarize the image but Otsu's Thresholding determines the value automatically.



4. Apply Closing Morphological Transformation on thresholded image. Closing is useful to fill small black regions between white regions in a thresholded image. It reveals the rectangular white box of license plate.
5. To detect the plate we need to find contours in the image. It is important to binarize and morph the image before finding contours so that it can find more relevant and less number of contours in the image. If you draw all the extracted contours on original image, it would look like this:



6. Now find the minimum area rectangle enclosed by each of the contour and validate their side ratios and area. We have defined the minimum and maximum area of the plate as 4500 and 30000 respectively.

7. Now find the contours in the validated region and validate the side ratios and area of the bounding rectangle of the largest contour in that region. After validating you will get a perfect contour of a license plate. Now extract that contour from the original image. You will get the image of plate:



Problem Statement

Helmet Detection and License Plate Recognition System using YOLO Model.

Scope Statement

Motorcycle accidents have been rapidly growing throughout the years in many countries. The helmet is the main safety equipment of motorcyclists; however, many drivers do not use it. The main goal of helmet is to protect the drivers head in case of accident.

In case of accident, if the motorcyclist does not use can be fatal .and it is not possible for traffic police force to watch every motorcycle and detects the persons who are wearing helmet or not, so there was need to make the automates system that's automatically monitor motorcycles and detects the persons wearing helmet or not and also detect number plate to penalize those persons without a helmet.

2.1 Introduction

Project (System) Analysis

Proposed System:

In this system, the video is feed to the system as an input. Frames are extracted for further processing. The two main algorithms object detection and license plate recognition is process through deep learning methods.

The object detection using computer vision algorithm is affected by different aspects like light variation, illumination, occlusion and system has difficulty to detect the multiple objects. Hence in this paper, Tensor flow-based object detection algorithm has been used.

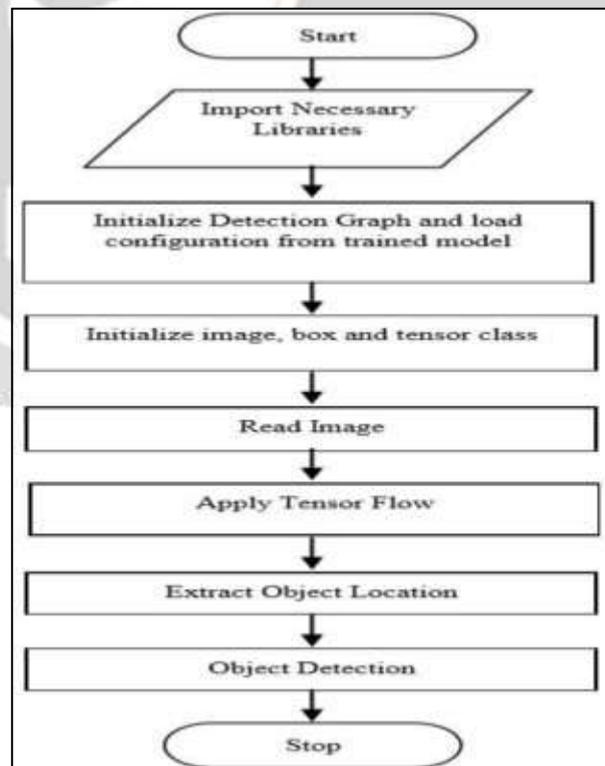


Fig. 5 General Control Flow of System.

YOLOv4 Introduction

The YOLO v4 algorithm uses a new CSPDarknet53 neural network as the main model of the detector. YOLO v4 contains 29 convolutional layers of $3 \times 3, 725 \times 725$ receptive fields and 27.6M parameters. The SPP block is added to CSP Darknet53 which significantly increases the receptive field and separates contextual features. At last, PANET is used instead of FPN in YOLOv3 as parameter aggregation methods for different detector levels. In addition, YOLOv4 also proposes some new improvement strategies: 1) Apply a new data augmentation method: mosaic and adversarial training-

2) Select the optimal hyperparameters through GA; Improve existing methods to be more suitable for efficient training and inference: improve SAM and PAN.

Terminology

- TP: moving object correctly identified moving object
- FP: Stationary object incorrectly identified as moving object
- TN: Stationary object correctly identified as Stationary object
- FN: moving object incorrectly identified as Stationary object

Sensitivity

Sensitivity It is the ratio of truly object present in the scene who are correctly identify as an object. This term presents the number of positive samples correctly identified. Higher the true positive element higher is the sensitivity
 $Sensitivity = TP / (TP + FN) (5.1)$

Specificity

It is the ratio of truly stationary object present in the scene that are correctly identify as a stationary object. This term presents the number of negative samples correctly identified. Higher the true positive element higher is the Specificity
 $Specificity = TN / (TN + FN) (5.2)$

Accuracy

Accuracy is the overall performance of the system including sensitivity and specificity.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) (5.3)$$

Architectural Design

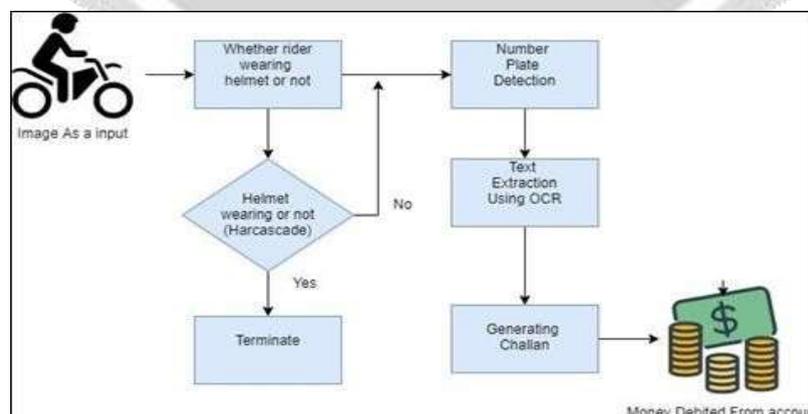


Fig. 6 System Architecture

Project Planning

The Project Planning concept is very important. After we selected our project topic project planning is important. In software engineering and development, project planning is intertwined into the methodology in a way that sometimes it is difficult to separate these subjects from each other.

To understand how to prepare a plan for our project, consider the situation that we are standing in. Lack of planning and being care- less about scheduling can create an irreversible catastrophic circumstance that puts our entire project in a hazardous situation. The dangerous outcomes can be very difficult.

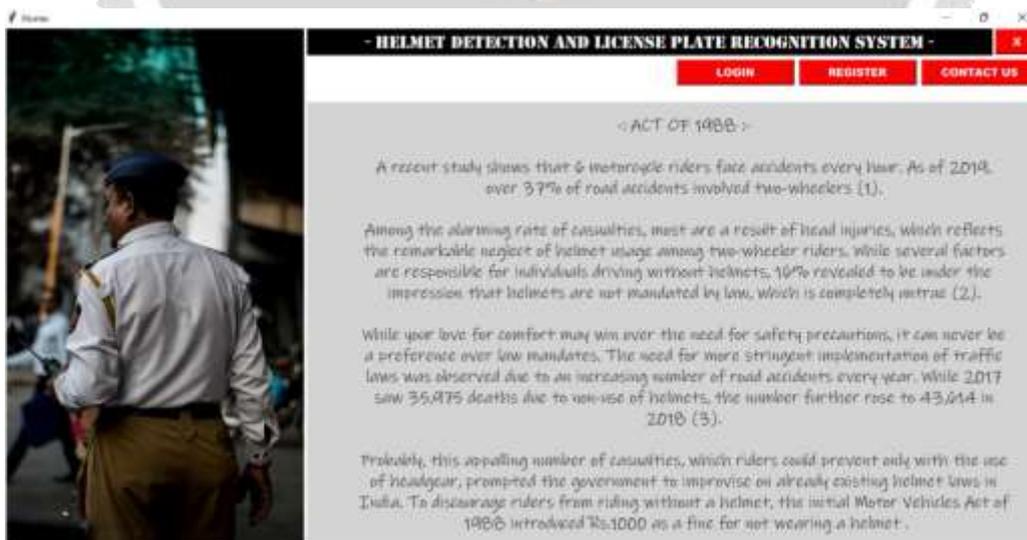
The extent can be from low quality project delivery at one extreme to total failure at the other. Now how we can prevent this important scenario? How we can guarantee that our project would provide the requirements to a pass as a maximum achievement? How we can turn it to a real success?

Indeed, taking simple steps can help to overcome with this situation properly. As a first action, if we are already understanding planning concept then consider the following steps:

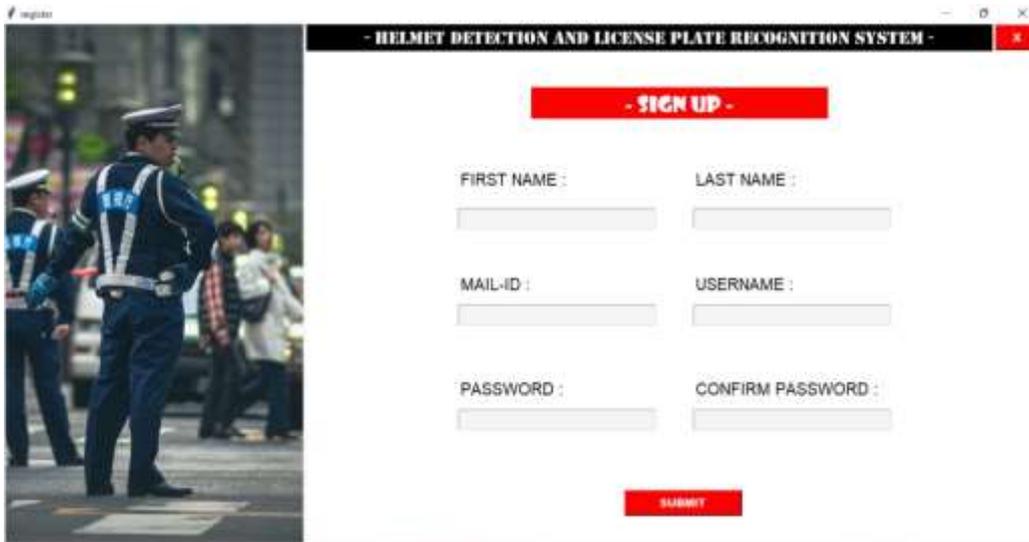
- Main phase of project plan
- Understanding the problem area
- Literature review
- Requirements Management
- Analysis
- design
- Implementation

Project Output Images (Screenshots)

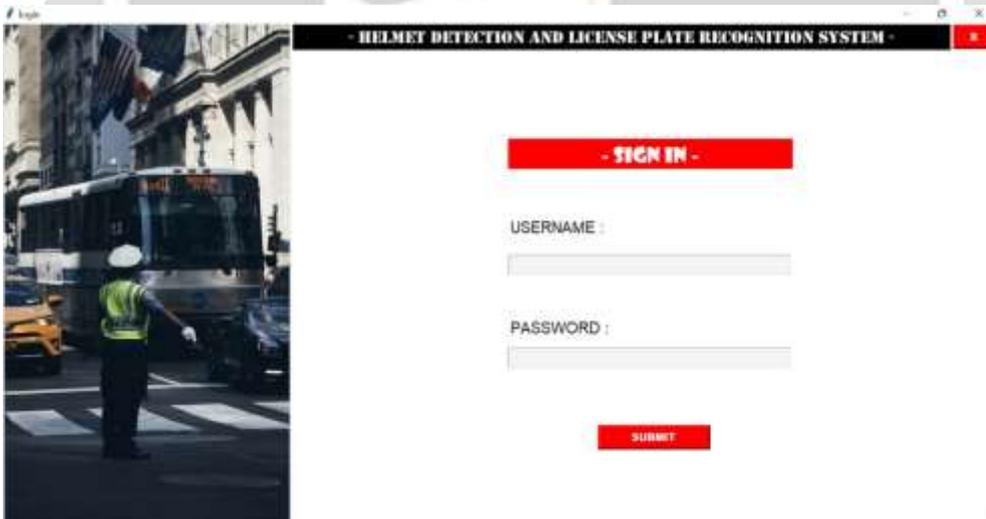
Home Frame:



SignUp:



SignIn:



Menu:



Contact Us:



Conclusion

In this paper we have described a framework for automatic detection of motorcycle riders without helmet from CCTV video and automatic retrieval of vehicle license number plate for such motorcyclists. The use of Convolutional Neural Networks (CNN) and transfer learning will help in achieving good accuracy for detection of motorcyclists not wearing helmets. But, only detection of such motorcyclists is not sufficient for taking action against them. So, the system will also recognize the number plates of their motorcycles and store them. The stored number plates can be then used by Transport Office to get information about the motorcyclists from their database of licensed vehicles. Concerned motorcyclists can then be penalized.

References

- [1] Cheng, R.; He, X.; Zheng, Z.; Wang, Z. Multi-Scale Safety Helmet Detection Based on SAS-YOLOv3-Tiny. Appl. Sci. 2021, 11, 3652. <https://doi.org/10.3390/app110836>
- [2] Zhang, Z.; Xia, S.; Cai, Y.; Yang, C.; Zeng, S. A Soft-YoloV4 for High-Performance Head Detection and Counting. Mathematics 2021, 9, 3096. <https://doi.org/10.3390/math9233096>

- [3] Dirir, A.; Ignatious, H.; Elsayed, H.; Khan, M.; Adib, M.; Mahmoud, A.; Al-Gunaid, M. An Efficient Multi-Object Tracking and Counting System Using Deep Learning in Urban Vehicular Environments. *Future Internet* 2021, 13, 306. <https://doi.org/10.3390/fi13120306>
- [4] Park, S.-S.; Tran, V.-T.; Lee, D.-E. Application of Various YOLO Models for Computer Vision-Based Real-Time Pothole Detection. *Appl. Sci.* 2021, 11, 11229. <https://doi.org/10.3390/app112311229>
- [5] Kim, M.; Jeong, J.; Kim, S. ECAP-YOLO: Efficient Channel Attention Pyramid YOLO for Small Object Detection in Aerial Image. *Remote Sens.* 2021, 13, 4851. <https://doi.org/10.3390/rs13234851>
- [6] Wang, Z.; Wu, Y.; Yang, L.; Thirunavukarasu, A.; Evison, C.; Zhao, Y. Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches. *Sensors* 2021, 21, 3478. <https://doi.org/10.3390/s211034>

