Integration of Edge Detection and Machine Learning Techniques in Smart Traffic Management

Prajwal V[a], Samarth R Bogar[a], Shambhavi Jagadeesh[a], Vishwas R[a], D V Ashoka[a$]
[a]Department of Information Science & Engineering
JSS Academy of Technical Education, Bengaluru, India
$: dr.dvashoka@gmail.com Orcid.org | 0000-0003-1326-2387

Abstract - Traffic congestion problems are nowadays escalating owing to the rising usage and growing vehicle count, this brings about the need for the usage of advanced technology in managing traffic congestion. This system manages traffic with the help of processing of images and machine learning techniques instead of the simpler ways of having a timer for each traffic light or embedding a sensor to detect vehicles. It captures images using a camera installed on the traffic light, along with an empty road image for reference. After processing the image, the two images are matched for density calculation of traffic according to which the duration of traffic lights is controlled. Amidst the congestion the ambulances are detected in order to ensure free vehicular movement of the lane. Traffic violations such as triple riding and travelling without a helmet are detected using the licence plate detection.

Keywords– traffic congestion, machine learning techniques, image processing, traffic violations, detection

1. INTRODUCTION

In today's time, one of the most pressing issues that are directly affecting the lives of people is the problem of traffic congestion. Mainly in metropolitan cities of India, the traffic congestion is so high that 3 cities - Mumbai, Bengaluru, and Delhi, are rated the top 10 worst traffic cities in the globe. Experts have come up with different kinds of solutions to overcome this problem. Each solution works best in certain conditions and regions.

In our project, we have proposed a smart traffic control system that has greater advantages over traditional traffic control systems. Time-controlled traffic lights or sensor detection do not work effectively. Time Controlled traffic lights cannot analyse the real-time traffic congestion and the sensors have a limit to the number of vehicles they can detect. Further, the sensors may become outdated in the future and constant maintenance and installation are not cost-efficient. Digital image-based processing of traffic parameters is a better alternative. Hence we are proposing Image Processing. Image processing solves both the problems posed by timer-controlled and sensor-controlled traffic systems. A camera captures the live broadcast of the road along with the density of the vehicles is calculated to determine the length of the traffic signal cycles. Image processing is a technique in which digital images are taken as input and one or more algorithms are run on it to get the necessary results. It is one of the few technologies growing rapidly as it contains many advantages. A camera can cover a wider area than sensors. Live images give better and more accurate information on traffic situations. Using efficient algorithms and good-quality cameras, we can achieve higher percentages of accuracy.Bikes and mopeds are an exceptionally famous method of transportation in pretty much every country. Due to less protection, high risk is involved. For the risk to be decreased it is exceptionally alluring for two-wheel riders to utilise protective helmets. The most elevated the number of street mishaps is by bikes. Although reckless and rash driving is the primary cause of these mishaps, head wounds form the single biggest justification behind street mishaps. More than 33% of street mishaps have brought about death that might have been avoided by wearing a helmet. Despite the fact that the use of helmets is required in numerous nations, a few motorcyclists don't use them or use them inaccurately. To develop a helpful capable framework for recognizable proof of helmet usage of bike riders, a promising technique for achieving this automated recognition of motorbike protective helmets is to use Machine Learning. Utilising a particular dataset, a Helmet detection model can be developed and carried out. Riders without helmets can be effectively distinguished utilising this model. From the caught picture of the helmetless rider, the tag of the vehicle is trimmed out and saved. The saved picture is given to an AI model to distinguish the characters in the licence plate. This system aims to change unsafe behaviours and consequently reduce the number of accidents and their severity.

In many regions of the world today, road transit is one of the most rudimentary ways of transportation. The number of vehicles on the street is increasing dramatically everyday. Because of which, in urban areas traffic congestions cannot be avoided. When the traffic is not managed in an efficient way, it can lead to an increase in pollution, wastage of travellers time, increases the cost of transportation and driver stress. But the most important factor that we need to emphasise is that the emergency vehicles movement is restricted due to the congestion. Especially in countries like India, the speed at which the vehicles are increasing is inversely proportional to the pace at which the road expansion is being done. Creatingspecial lanes for ambulances given the length and width of Indian roads is hard to achieve. Our model reduces the impact caused by vehicular congestion on ambulances. Traffic controlling based on timers, sensors and fuzzy logic network being used in the existing systems are simply ineffective. Using video surveillance from cameras present at traffic congestion junctions helps to detect ambulances thus controlling the light signals until it passes through the junction.

Street security is the main part of this vehicle-driven mechanical world. If the number of people taking road transport as the means to reach their destination are considered, the number of people who are met with accidents instead of reaching their homes safely is increasing day-to-day. The lack of responsibility while driving the bikes or the weighty speeding of the four-wheelers is the significant justification for the occurring accidents. These irresponsible drivers are making it hard for the drivers who actually follow the traffic rules. The current increase in the fine/challan system might control these irresponsible drivers to a degree, yet this is certainly not a long-lasting arrangement that we can depend on. Monitoring the roads continuously for these sorts of unreliable drivers is human exertion-consuming. The triple riders cause the majority of two-wheeler related accidents.. Our system proposes a way to detect these riders through CCTV surveillance systems. These surveillance systems are automated to reduce the human effort in monitoring the triple riding violators. By capturing the images of these riders, the vehicle number or the licence plates can be detected. An automated model will then provide an output after recognition of the characters in the licence plates.

2. LITERATURE SURVEY

The Indian traffic system is rigid and inflexible as the vehicle count keeps growing rapidly, and does not consider the change in vehicular density throughout the day. Which in turn results in wastage of fuel and time. So, this paper aims to analyse and improve congestions using operation management based on traffic density. Depending on the vehicular density, traffic is splitted into various packets. When vehicles need to pass through the intersection, at any moment maximum packet having high density will pass through the intersection resulting in higher throughput of packets. Shadows from buildings, bridges cast on traffic lanes are resolved using otsu's multiple thresholding. To provide an optimal solution and to detect traffic patterns, machine learning techniques are used by the proposed system [1]. Due to the usage of the IoT (Internet of Things) and image and video processing techniques to achieve the number of passing cars, traffic management has become one of the most critical concerns in today's world. Furthermore, the Raspberry Pi device and the OpenCV tool are used to implement it. The proposed models' efficiency in the direction of the crossroad is demonstrated by analytical and experimental data. The image is set as a reference image for each direction in the Raspberry direction, else the ready images of traffic control cameras are prepared for each side and then provided to the original image. White pixels are counted after the edge image is displayed as a white colour and a reference image is obtained. The information is then sent to the scheduling algorithm, which decides on any other place to monitor on the fly, and traffic is controlled to reach path traffic statistics using the Raspberry Pi and the on site network platform [2]. Proposed a system that recognises licence plates of automobiles using vertical Sobel edge detection and extracts them using morphological operations with 100% accuracy. Scanning techniques are used to align, segment, and isolate the characters on licence plates. Finally, the Prewitt detection of edge method will recogniselicence plates. Initial step of the procedure is to photograph the automobiles. The licence plates must then be detected and extracted from the collected photos. These images are adjusted to required formats by converting RGB into the grey image, applying vertical Sobel edge detection, binarization of the image, edge horizontal histogram for threshold selection, and the morphological operation are used to locate candidate licence plate locations in a binary vertical gradient image Because vehicle photos may be obtained from different perspectives, the licence plate must be aligned in the next stage. The segmentation of characters, numerals, and words from the licence plate image is the following step. All unwanted licence plate regions will be deleted, photos will be adjusted as needed, and the licence plate will be binarized. The final stage of the proposed system is recognizing each segmented character, number, and word that are extracted from licence plate regions. The Prewitt edge detection algorithm is used in this stage, where the percent value will be determined for each character, number, and word. The resulting value compares with 11 the values stored in the prior determined database [3]. The suggested technique includes a framework that may continually transmit the vehicle count to the controlling station and produce an alarm in the event of a big vehicle gathering. From this, their vehicle count is found using contour properties. Furthermore, the monitoring data is sent through the internet to a remote controlling centre located anywhere in the city. The footage in real-time of the movement of traffic on the road was captured using a camera system. After obtaining the footage from the camera, the data was processed using a portable Raspberry Pi processing device. Finally, after determining the average vehicle count during a certain time frame, the same processing mechanism is employed to send the same data to the system of central control. Vehicle counting here is achieved by first extracting the individual frames from the input video stream. Then using OpenCV in python, the foreground image is developed using background subtraction. The quality of the image is improved and the noise is removed usingmorphological operations. Then the otsu thresholding method is applied for obtaining the binary threshold image. Finally, the vehicle count is obtained from the binary image using contour detection and using contour properties. This data is then transmitted for traffic management using the internet through the following steps. Fetching the per sec vehicle count data, interfacing of firebase google server Raspberry Pi, establishing connection with firebase, and sending data to the real-time cloud [4]. This research presents an image processing-based method for minimising traffic congestion by recognising and tracing blobs. It also provides a solution to clear the traffic route for emergency vehicles by using Bluetooth. The camera is firstly mounted to cover the full lane for vehicle recognition and counting. Then minimization of image is performed, in which the frames from the video are compared to identify the contrasts between them. The contrast then traces the blobs. The changes detected are the blobs or the moving vehicles. To ignore pedestrians, blobs of very small size are ignored. To track the blob, a line of thresholding is done to compute the number of automobiles. Next to this traffic signal, drawing of a line is done to an optimal grade. These vehicles are only counted when they pass those lines. This automobile count can determine high or less density, as well as adjust the traffic signal time. A Bluetooth module (BM) and a phone with Bluetooth enabled are used to detect an ambulance. When the emergency vehicle gets close to indication, the individual in charge of the ambulance will send out an order to the BM, which instructs the signal to adjust accordingly. The processed information will be sent onto an A-Microcontroller and terms are decided then [5]. Utilising IoT situations, clearing the traffic by conveying messages to the message board is conceivable, therefore an ambulance can reach the hospital without time delay and without any waste of time by clearing the traffic load. A model is developed using Embedded and IoT for clearing the traffic when ambulances are coming in the way of lanes. This approach involves a GPS device inside the ambulance which traces the location of the ambulance andsends it to the GPS associated with the ARM processor and the GSM module, which then sends it to the Traffic Control Management Unit. This is done to ensure that the message sent to the Traffic Control Management and any acknowledgment received from the receiver side is done with information being encrypted of the highest security. Communication is done with the help of the internet of things as the IoT data can be sent immediately. The IoT information is sent by the GSM module to the signal logic gate and all the data is passed to the PC using the Optocoupler circuit.The PC controls the Traffic Signal LED in like manner [6]. Have proposed a system that uses a mobile application through which an ambulance driver has to activate the emergency mode. As soon as emergency mode gets activated, the app will keep tracking the nearby traffic signals using GPS automatically. The destination will be pinned by the driver to help the app understand the route which can be followed by the ambulance. Whenever it finds a nearby signal within a radius of 500 metres, it will promptly ask the control room to check for the current status of the signal in which the emergency vehicle has to cross and it will open the signal automatically for giving the way for an ambulance, and when it crosses the signal, a request will be sent again for setting back the light signal to its normal timings. If the signal is already green, it will get the remaining time, and based on the remaining time, the application will naturally demand the control room to freeze the light signals until the emergency vehicle has crossed utilising AI strategies [7]. This paper proposes a framework comprising four modules: Acquiring image, Pre-processing, 16 Motorcycles, and Helmet Detection. A video stream, that is the camera film is captured and is examined frame by frame. Background subtraction is done to identify moving objects and isolate them from static objects within the scene, where a static camera is used. Basic thresholding is then applied to form a binary image, expelling shadows within the process. Next, morphological operations like opening, closing, erosion, and dilation are utilised to decrease noise. The following step is to draw contours and bounding boxes overthe moving objects. Two reference lines are used to detect the relative movement of a vehicle over the screen, and when a moving object crosses them, they are classified. The aspect ratios of bounding boxes are considered and used to determine if the moving object may be a motorcycle, car, or a different heavy vehicle. Execution of the helmet detection framework used the YOLOv3 network. Thousands of images are utilised to prepare this YOLOv3 network, which uses Darknet-53, a CNN which analyzes the ROI (Region of interest) to decide on the off chance that the rider of the cruiser is wearing a helmet or not. The tests on a large set of helmet-positive images gave up to 80%accuracy [8]. This paper proposes an image processing technique on detection of helmets. The technique is classified into four parts; primarily is image acquisition where the video of the traffic is captured by a camera and is reduced to frames. Second, comes the pre-processing techniques where noise removal and binarization take place. Pre-processing techniques include Centre Line, Median Blurring, Thresholding, and Morphological Operations. The third part is the Vehicle Classification. By fixing some criteria on two parameters (Aspect ratio and area) of the contour of a particular vehicle, motorcycles are classified and processed further. The final part is Helmet Detection. After extracting the Region of Interest [ROI], which is the main part, its features are given to the classifier which is being trained by a certain number of pictures of helmets. By matching and using trained features and ROI, we determine whether or not the motorcyclist is wearing a helmet [9]. Emphasises the development of a Non-Helmet Rider Identification System, which aims to automate the detection of traffic violations and the licence plate number extraction of the vehicle. Object detection using Deep Learning at 3 layers is the basic principle followed. A human, a motorcycle/moped at the first level using YOLOv2, a helmet at the second level using YOLOv3, and a licence plate at the third level using YOLOv2 are the objects observed. After that, OCR is used to retrieve the licence plate registration number (Optical Character Recognition). Initially, frames are collected at regular intervals from videofiles. The frame chosen is loaded to the YOLOv2 object detection model, where the classes to be detected are MotorBike or Person. Using the Image AI library, only the detected objects are extracted. The details of these extracted images are stored in a dictionary which can be 17 later used for further processing. The bounding box around the helmet along with the detection probability is displayed. If the helmet is detected no further processing is done, else the licence plate detection will take place if no helmet is detected. OCR can be used to extract licence plate numbers if the helmet is not worn by the rider [10]. Have proposed a machine learning model for predicting air quality index using various machine learning algorithms. All the algorithms have been hyperparameter tuned resulting in the XGBoost algorithm with highest accuracy [11].

3. METHODOLOGY

The first step is to acquire the data in the form of images from various sources, by capturing the real-time live video of the Vehicular congestion. From the live video, video frames will be extracted every second. The obtained video frames will be in the form of RGB colors, it needs to be converted into grayscale. On the grayscale image, shadows of buildings and bridges will be casted on the road, which is treated as noise in the image. During thresholding, important details from the image can be lost due to the shadows. Every image will have a threshold value i.e., above the threshold value the image will be brighter compared to the value below. Grayscale images contain pixel values in the range from 0 to 1, a threshold value needs to be selected in the closed range [0.0, 1.0]. An image can have multiple threshold values. The value of the threshold is determined using the grayscale image histogram. For multiple thresholding, Otsu's algorithm is used over various pixel densities of the image.

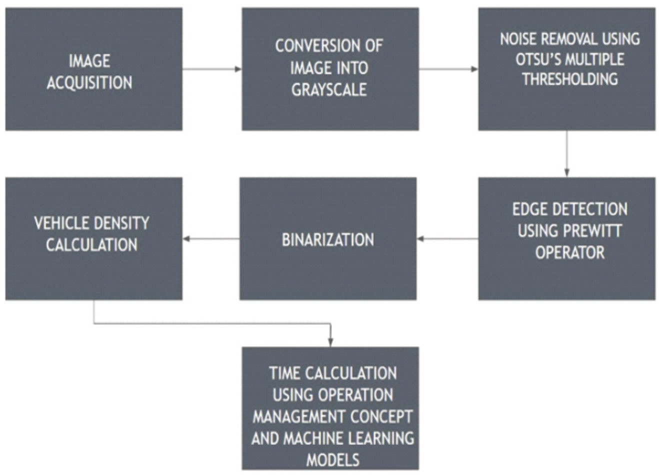

Fig. 1 : Block Diagram of Traffic Congestion Management System

The input image containing different pixels is separated to form several classes, the classes are formed based on the gray intensity level present in the image. Further different thresholds are calculated based on the number of obtained classes. Algorithm returns two threshold values for three classes, which is the default number of classes. This results in better output, thus removing the noise. In a digital image, edges are large local changes in intensity. A set of connected pixels that creates a border between two discontinuous regions is known as an edge. Horizontal edges, Vertical edges, and Diagonal edges are the three types of edges.

Edge Detection highlights the structural features and reduces the data present in the image. This helps in identifying various image sections, and marks the end and start of each section. There are two types of edge detection operators: Gradient and Gaussian. Prewitt operator is based on edges being calculated on the difference between image pixel intensities. Image is a2D signal. Mask being a signal is represented with a 2D matrix. In order to make images sharper, masks are used in the edge detection process.

Prewitt operator will output a couple of masks for edge detection, first is in the horizontal direction and the other in the vertical direction. On applying the mask horizontally, vertical edges are made prominent. Similar to first order derivative based on the difference of various denisties of pixel images of an edge. This results in increasing the intensity of the edge Resulting in increased edge intensity and enhancing the original image. On applying the mask vertically, horizontal edges are made prominent. Its working is based on the principle mask and it's calculated on the difference between image pixel intensities about that edge. The mask center row will have zeros thus not containing any original values of edges of the image, rather calculates the difference of pixel densities about the edges. This causes the image to become more sudden and prominent.

The next step is to convert the image into a binary image, which consists of the values 0 or 255 by eliminating gray levels in the image. Edge based matching is performed by comparing the image with all the other edges with the prewitt operator. Detected edge images will be matched and vehicle area density is calculated. Time Calculation for controlling traffic signal lights is done using machine learning models. The vehicular area density is the independent parameter and the dependent parameter is the time which will be predicted using reinforcement learning algorithms. The reason for reinforcement learning is based on the previous vehicular density the current prediction would be done. This allows us to understand the trends and the peak hours of traffic to set the time of the signal accordingly.

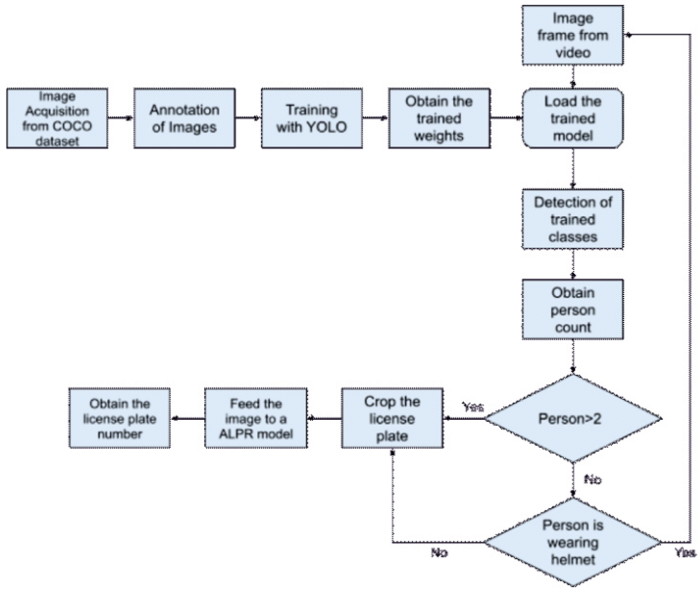

Fig. 2 : Block Diagram of Helmet and Triple Riding Detection System

For real-time triple riding and detecting helmets, speed and accuracy are necessary tools. Here, a model based on DNN - YOLO (You Only Look Once) comes into picture which is a state-of-the-art system of detecting real-time objects, which predicts the future based on a single network assessment. Detecting objects itself is a method for finding instances of a specific class in a picture or motion video, such as animals, humans, and others. Using pre-trained models of detecting objects, the API of Pre-Existing Object Detection makes it simple to identify things. This model detects several objects which are not useful to us, and therefore we will create and train our model for only the required classes. Five objects need to be detected for triple riding detection implementation, helmet detection, and extracting and recognising number plates, which are: Person sitting on the two-wheeler, the two-wheeleritself, helmet, no helmet, and the license plate. A bespoke model of detecting objects is required. The dataset is a series of photos that include the items of the classes that need to be identified. The custom model is then trained using the dataset. The model may then be used to detect these things after it has been trained. All of the photographs that were obtained, together with their dimensions are fed for training. The model uses the ground truth of the needed classes to extract the characteristics of each class from each picture. For withdrawing and hoarding features, we employ a deep learning classifier based on CNN to detect those characteristics in other pictures. When the picture is supplied to this trained model, it is important to detect the needed set.

The processed photos are sent into the model YOLOv3, which is used to train for the bespoke classes. The model is then loaded using weights. Then, the image is given as the input. All the five classes trained are detected by the model. From this, we obtain info about the individual riding the bike. If there are more than two people riding the bike or the individual doesn't have a helmet on, we can withdraw the license plate from the picture. When the traffic violating bike rider is found out, the associated set is detected. This is achieved by checking if the coordinates of the helmet set does not lie inside the people set or not and if the number of individuals is more than two. Similarly, the supported bike and number plate are detected using the same steps. Once found, the number plate coordinates are cropped and stored as a new picture. The OCR will take the license plate which is extracted as an input, in order to recognize all the text and strings in machine encoded format. Thus resulting in the license plate numbers that are predicted with value reflecting its accuracy. Based on the accuracy obtained, the highest is saved for further use.

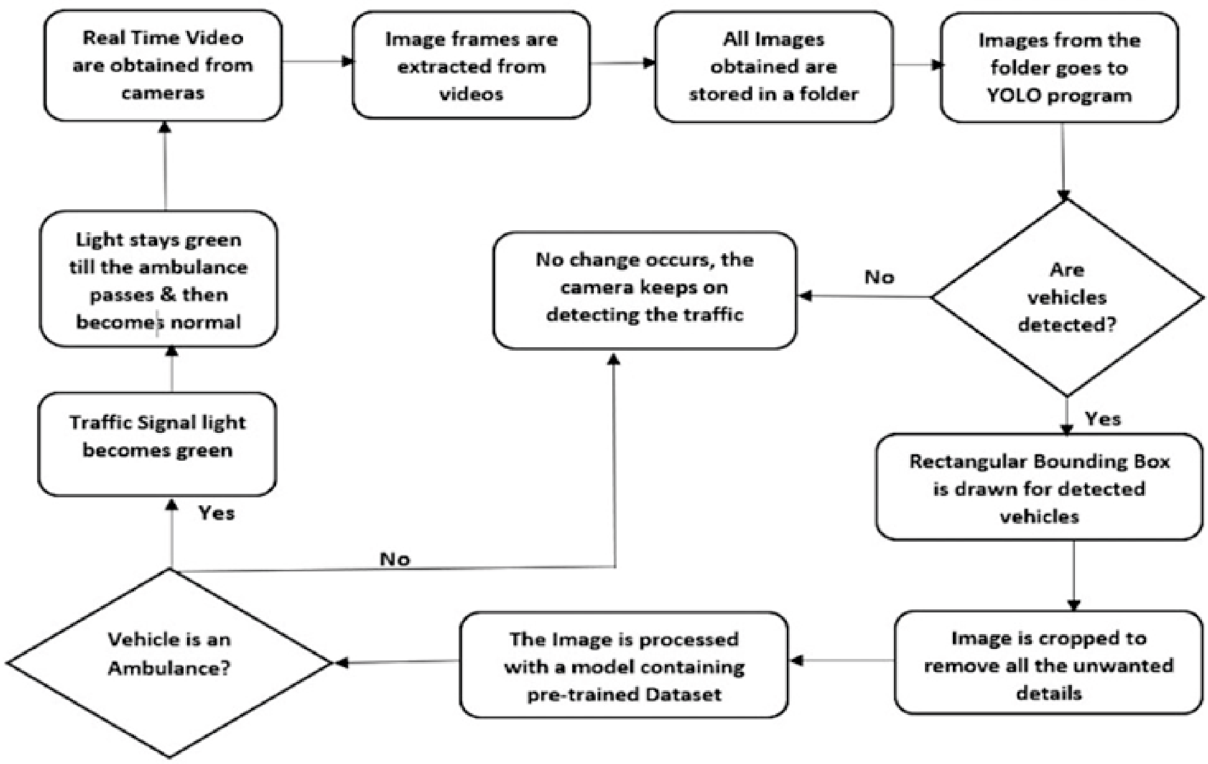

Fig. 3 : Block Diagram of Ambulance Detection System

The camera captures video snippets which would be the input, extracting the frames to get images. The images obtained will be stored in a folder. The YOLO (You Only Look Once) program takes the images from the folder and detects all the vehicles present. It uses COCO (Common Objects in Context) to detect which is developed by Microsoft. YOLO is a technique based on neural networks for real time object detection. YOLO is a real-time object identification technique that uses neural networks. YOLO detects objects similar to a regression problem that reports all the probabilities. The YOLO algorithm is based on three techniques, firstly using the residual blocks the image will be divided into various grids of dimension S cross S. Secondly an outline is drawn to highlight the object in an image which is represented by a bounding box. Every bounding box consists of the width, height, box center, and the class it belongs to. Thirdly, using intersection over union (IOU) YOLO will construct the output box which will surround objects.

If vehicles are detected in the image, then a bounding box is drawn surrounding the detected vehicle. Further, the images of detected vehicles will be cropped according to the bounding boxes. In order for the system to detect vehicles, it should be trained to recognize one. Using the deep learning platform TensorFlow, a CNN on a dataset of images of ambulances is trained. Using Keras, the TensorFlow algorithm is used to train the dataset. The image on being passed to main function, where the trained models reside will be called, to detect if the image is an ambulance. Yes is returned by the main function if it qualifies as an ambulance. Which makes the traffic signal light of that lane turn green and red for all the other lanes at the intersection of roads. The light will remain green until the ambulance passes. And then the lights are back to their previous state. If the main function returns no, there is no change in the traffic signal lights, and incoming traffic is used to detect the further vehicles, repeating the entire process in a cyclic fashion.

4. CONCLUSION

The proposed traffic management system overrides the traditional traffic control systems in terms of installation and maintenance cost, delay in result and accuracy. Reducing waiting time and congestion leads to fewer traffic signal violations which in turn leads to fewer accidents. This system of traffic coordination will also provide data for future road designsand constructions. Image processing is a field that is getting more and more accurate over the years. It allows the machine to become self-sufficient as no external data is required. Each traffic light junction will act as a separate entity as it will be customized to that junction only. This increases the efficiency drastically. With the additional features mentioned. The proposed system has the potential to really improve the traffic condition in our country if implemented properly. In future, improvements in the algorithm used and higher camera quality can further increase the efficiency.

ACKNOWLEDGEMENT

The authors wish to acknowledge JSS Academy of Technical Education, Bengaluru, for providing the facilities to carry out the research work.

REFERENCES

[1] Pranav Maheshwari, Deepanshu Suneja, Praneet Singh, and Yogeshwar Mutneja, “Smart Traffic optimization using image processing” in *Proc. IEEE 3rd Int. Conf. MOOCs*, Amritsar, India, Oct 2015, doi: 10.1109/MITE.2015.7375276

[2] Meisam Razavi, Mehdi Hamidkhani, Rasool Sadeghi, “Smart Traffic Light Scheduling in Smart City Using Image and Video Processing”, in *Proc. IEEE 3rd Int. Conf. Internet of Things and Applications (IoT)*, Isfahan, Iran, April 2019, doi: 10.1109/IICITA.2019.8808836

[3] Elaf J. Al Taee, “The proposed Iraqi vehicle license plate recognition system by using Prewitt Edge Detection Algorithm”, in *Proc. Journal of Theoretical and Applied Information Technology*, vol. 96, no. 10 May 2018

[4] Satbir Singh, Baldev Singh, Ramandeep, Baljit Singh, and Amitava Das, “Automatic Vehicle Counting for IoT-based Smart Traffic Management System for Indian urbanSettings”, in *Proc. IEEE 4th Int. Conf. Internet of Things: Smart Innovation and Usages (IoT-SIU)*, Ghaziabad, India, April 2019, doi: 10.1109/IoT-SIU.2019.8777722

[5] Varsha Srinivasan, Yazhini Priyadharshini Rajesh, S Yuvaraj, and M Manigandan, “Smart traffic control with ambulance detection”, in *Proc. IOP Conf. Series: Materials Science and Engineering 2nd Int. Conf. Advances in Mechanical Engineering*, vol. 402, Kattankulathur, India, March 2018

[6] Venkatesh H, Shrivatsa D Perur, and Jagadish M C, “An Approach to Make Way for Intelligent Ambulance Using IoT”, in *Proc. Int. Journal of Electrical and Electronics Research*, vol. 3, issue. 1, March 2015, pp. 218-223

[7] M. Abirami, G. Archana, R. S. Deepika, and M. Keerthana, “GPS Based Traffic Signal Control System for Ambulance Using Machine Learning ", in *Proc.*

[8] Noel Charlie, Yashaswini Ashok, and Shanta Biridar, “Automatic Helmet Detection System on Motorcyclists Using YOLOv3”, in *Proc. Int. Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 8, issue. 5, May 2020

[9] Shreya Bhagat, Dhwani Contractor, Sonali Sharma, Tanu Sharma, and Mrs. Ketki C Pathak, “Cascade Classifier based Helmet Detection using OpenCV in Image Processing” in *Proc. National Conf. Recent Trends in Computer and Communication Technology (RTCCT)*, May 2016, pp. 195-200

[10] Prajwal M. J., Tejas K. B., Varshad V., Mahesh Madivalappa Murgod, and Shashidhar R, “Detection of Non-Helmet Riders and Extraction of License Plate Number using Yolo v2 and OCR Method”, in *Proc. Int. Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, issue. 2, December 2019

[11] Basamma Umesh Patil, Ashoka D.V., Ajay Prakash B. V., “Optimization of Hyperparameters in Machine Learning Techniques for Air Quality Predictive Analysis” , in International Journal on Information Technologies and Security, vol 13, issue 3, September 2021

***