Automating Traffic Lights using Deep Learning

[1] **Tejaswini Shailesh,[2]Huda Mirza Saifuddin,[3]Pratheek G Aithal**

[13]Student,[2]Assistant Professor
[123]Department of Information Science and Engineering,
[123]B N M Institute of Technology,Bengaluru, India

Abstract— Efficient traffic management is necessary to avoid traffic jams which affect wide areas. With the increase in vehicles, the traditional control strategies are incapable of clearing heavy traffic which leads to long traffic queues and prolonged waiting times. Another challenge faced is that of emergency vehicles having to wait for a long time due to traffic congestions and blocks. It can be a life or death situation for any person as each and every second counts. The proposed system is designed with an aim to improve traffic clearance at intersections along with giving precedence to emergency vehicles as soon as it detects a siren sound. The system includes the use of pre-trained model RetinaNet to detect and count the number of vehicles and classify them. Enhanced Ratio based algorithm is applied to generate green signal timings. In order to detect the sirens from emergency vehicles, Convolutional Neural Networks (CNN) has been used.

***Index Terms*—Audio detection, CNN, Deep Learning, Intelligent traffic light control, Object Detection.**

I. INTRODUCTION

In recent years, with the growing population, the number of vehicles on the road has also increased drastically. The ever increasing flow of traffic in large metropolitan cities has been a cause of severe problems such as traffic jams, increased waiting time and fuel consumptions at intersections. Additionally, it is even a cause of concern for pollution. In a world where everyone is in a rush to reach their destination, traffic congestion is responsible for slowing people down and making them less productive. Solving this problem has been a challenge worldwide.

Another challenge faced is that of emergency vehicles having to wait for a long time due to traffic congestions and blocks. It can be a life or death situation for any person as each and every second counts. Predicting the oncoming of ambulance and clearing the traffic for it has also been a major issue.

The existing traffic signals operate on a set of predetermined rules with a fixed duration and require minimal human intervention. These Pre-Timed Control Systems (PTCS) are easy to install but are effective only in situations with low traffic densities and this rigid behavior does not handle exceptions that lead to congestion.

Using modern technology such as Deep Learning, the current traffic control system can be automated, thus ensuring smoother flow of traffic throughout the city. A Traffic Density-based Control System (TDCS) can be introduced which uses information on current traffic. The main purpose of the project is to reduce traffic in an intersection and provide priority for emergency vehicles. This is done by computing total traffic load at a particular junction, which is then further used for real time traffic control by generating green light timing of the traffic signal. The lane with relatively high traffic density is allowed to move for a longer time.

The system also keeps a track of siren sounds. When there is a siren detected, the lane having the emergency vehicle is granted green signal immediately. This ensures that the emergency vehicle reaches its destination in time and thus minimizing the delay caused by traffic congestion.

II. LITERATURE SURVEY

Over the decades, there has been a lot of research on traffic signal control systems. Various researches have proposed systems with different methods with different accuracies. The aim of the project is to control traffic light time based on the traffic density to achieve a higher efficiency and also to give precedence to emergency vehicles when they are detected. GPS technology, Reinforcement learning with Q learning algorithm, YOLO and SORT algorithms for video analysis, ANN and Fuzzy logic controller and ReliefF algorithm with SVM classifier are used in the following papers proposed by various researchers.

Tanwar. R, et al., [1] proposed Global Positioning System which will convey the details of traffic flow and by considering that, time slots are provided to traffic lights i.e. entire controlling and overseeing will work like Real Time System. Various vehicle frameworks that are presently in utilization are some type of Automatic Vehicle Location (AVL), which is an idea for deciding the geographic area of a vehicle and transmitting this data to a remotely found server. To achieve vehicle following constantly, an in-vehicle unit and a respective server is utilized. The data is transmitted to a following server using GSM/GPRS modem on GSM framework by utilizing cellular telephone instant message or by direct TCP/IP association with following server through GPRS. The follow-up server likewise has a GSM/GPRS modem that gets vehicle zone data by means of GSM system and stores this data in a database.

In order to establish a connection with traffic lights and for analyzing GPS output, a data center is designed which will cover all main traffic lights of some specified zones. The role of a data center will be to analyze traffic signals. Data center consist of number of monitor (equal to the number of traffic lights in measured area) which will display the volume of traffic over every traffic light. GPS output will be analyzed in the data center using a predefined mechanism with the help of a system and timing coordination message for traffic light will be generated.

Natafgi, et al., [2] developed an adaptive traffic light system using Q-learning algorithm for reinforcement learning and tested using real data from Lebanese traffic. For each road, the sensors are installed at 2 locations - at the ending of the road near the intersection, called exit point, and at a location earlier in the road, called the entry point, where representative queue length can be obtained for that road.

A queue data structure is used in conjunction with the sensors to determine the queue lengths and queuing times for each lane. Each vehicle detected by the entry point sensors is added to the queue data structure and its entry time, queue length at entry, and velocity are recorded. When it reaches and passes the exit point, it is removed from the queue. The velocity is calculated by the average length of the car divided by the time it took to cross the sensor. The queue length is determined by the number of elements in the queue data structure. For training and testing the system, a software simulation tool is used. This tool can simulate the traffic intersection and allows the neural network to interact with it. Compared with the actual traffic light system, the proposed model displayed a reduction in average queue lengths by 62.82% and in average queuing time by 56.37%.

Sharma. M, et al., [3] aims to detect and track vehicles on a video stream and count those vehicles going across a predefined line. By using the count of vehicles on each side of the traffic light, we have optimized the traffic lights by assigning them time according to the traffic behaviour in real-time. If there is less traffic or more traffic than usual, the model will optimize the light by increasing or decreasing the duration of the light. YOLO and SORT algorithms was used on the live video feed to get the vehicle count in real-time.

Once the objects are detected and tracked over different frames, a simple mathematical calculation is applied to count the intersections between the vehicle's previous and current time frame positions with a defined line.

The amount of time wasted by drivers lingering at traffic signals was reduced by 40, and to traverse the targeted area reduced by 25%. By evaluating and comparing YOLO with other existing models on various grounds as far as YOLOv3 is a good detector. It's fast and accurate also. But by using YOLO, the purpose of identifying objects on the roads is solved.

Suhweil, Y, et al., [4] developed an automatic algorithm to control traffic light time based on artificial intelligent techniques and images for cars on traffic lights. This algorithm is validated by comparing its results with manual outcomes. Applying the following proposed algorithm in transportation system will regulate traffic flow and reduce travelling and waiting time wasted on roads. Algorithm starts by acquiring and preparing images for cars on traffic light, then segment images into different parts using sliding window, and predict if each part is a vehicle or not. Using artificial neural network, the total number of vehicles is updated in vehicle counter to control traffic light time using Fuzzy Logic Controller. The sample data results show an error of about 4% (average error).

Al-Ostath, N. et al., [5] proposed a model called ETL (Emergency Traffic Light) control system. The main goal behind this system is to provide a smooth flow for emergency vehicles like ambulances to reach their destinations in time, thus minimizing the delay caused by traffic jams. The ETL control system will control the traffic signal lights in the path of the emergency vehicles, stopping conflicting traffic, and allowing the emergency vehicle a path to help in reducing their response time.

Radio frequency is used for communication and the system contains 2 main parts - the sender and the receiver. The driver of the emergency vehicle has a set of control buttons to control the traffic lights, which when pressed, a radio frequency is transmitted. A radio frequency transceiver, in the sender, passes the signal to the radio frequency transceiver in the receiver which then changes the traffic light colour. This is a complete traffic lights intersection model in which each traffic light was operated for every 10 seconds.

Carmel, D. et al., [6] considers a set of acoustic features and the use of ReliefF algorithm to select only the ones that best differentiate between alarms and other sounds. Audio frames of 100ms were used. Careful feature selection is a key issue when designing an effective audio signal detector and this algorithm helps do so. ReliefF receives an input dataset with n instances and p features belonging to 2 known classes. The key idea of ReliefF is to estimate feature importance according to how well their value distinguishes among neighbouring instances. An SVM classifier is used as a detector. On the tested dataset consisting of several dozen alarm sounds and several dozen background noises, the proposed technique shows an accuracy of 98% per audio frame. This technique is not limited to particular alarms and can detect most electronically generated alerting sounds within 200ms.

III. PROPOSED METHODOLOGY

The project uses image dataset consisting of vehicles and another audio dataset consisting of siren and non-siren sounds. The implementation is divided into 3 modules as follows – vehicle detection, siren detection for ambulance and traffic light controller.

A. Dataset

The project uses image dataset taken from GitHub and has around 2000 images of both dense and sparse traffic consisting of all kinds of vehicles. It is used in vehicle detection and for simulating traffic.

An audio dataset consisting of around 200 emergency and non-emergency sounds has also been used for training and testing siren and non-siren sound detection.

B. Vehicle Detection

ImageAI provides very powerful and easy way to use classes and deep learning models such as RetinaNet to perform Image Object Detection and Extraction. RetinaNet is a one-stage object detection model composed of a backbone network and 2 task-specific subnetworks and is trained on the COCO dataset. resnet50_coco_best_v2.1.0.h5 is used as the backbone and is responsible for computing convolutional feature map over the input image.

The model is able to detect and classify motorcycles, cars, buses and trucks.

C. Siren Detection for Ambulance

This module majorly has 2 parts – training the model and using live audio to detect and classify siren and non-siren sounds.

1) Training the Model

The capabilities of Convolutional Neural Network is exploited on audio signals by converting them into respective Mel spectrogram. The features from the graph is used by the library librosa to build a Convolutional Neural Network having the features mentioned in Fig. 1. The Convolutional Neural Network has 8 layers, 3 of which uses tanh activation function and one uses softmax. The optimizer used is the Adam optimizer.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 16, 8, 64)         640

 max_pooling2d (MaxPooling2D  (None, 8, 4, 64)         0
 )

 conv2d_1 (Conv2D)           (None, 8, 4, 128)         73856

 max_pooling2d_1 (MaxPooling  (None, 4, 2, 128)        0
 2D)

 dropout (Dropout)           (None, 4, 2, 128)         0

 flatten (Flatten)           (None, 1024)              0

 dense (Dense)               (None, 1024)              1049600

 dense_1 (Dense)             (None, 2)                 2050

=================================================================
Total params: 1,126,146
Trainable params: 1,126,146
Non-trainable params: 0
```

Figure 1. Proposed CNN model

2) *Classifying the Audio*

Once the model is built and trained, it is saved and loaded using the Keras library. A real-time audio is then taken to test and predict if the sound is of class emergency or non-emergency.

D. Traffic Light Controller

From the vehicle detection module, the vehicle count is taken and an enhanced ratio based algorithm is applied to get the green signal timings for each of the lanes. The maximum time for each cycle is set to 120 seconds, which is then divided amongst the 4 lanes. In case of special conditions where the number of vehicles is less than 10, the green signal timings is double the number of vehicles in that lane.

```
for i, lane in enumerate(lanes):
    if lane < 10:
        time_per_lane[i] = lane * 2
# for the remaining lanes, ratio is calculated for the remaining time
time_left = max_round_time - sum(time_per_lane)
total_vehicles = sum(lanes)
for i, lane in enumerate(lanes):
    if(time_per_lane[i] == 0):
        time_per_lane[i] = int((percentage(lane, total_vehicles) * time_left))
```

Figure 2. Enhanced Ratio based algorithm

The time for the remaining lanes is calculated based on the ratio of the number of vehicles on one lane to the total number of vehicles at the junction multiplied by the remaining time.

The traffic light controller controls the timing of the traffic signal based on the traffic density at each lane. It also keeps checking for the siren sounds of any approaching emergency vehicles with the help of the siren detection model trained earlier. When a siren sound is detected, a flag is set to that lane where the siren was detected and the lane is immediately granted green signal.

IV. IMPLEMENTATION DETAILS

The object detection is performed using ImageAI, RetinaNet resnet50_coco_best_v2.1.0.h5 model. The siren detection is performed by building a convolutional neural network using the Python's Keras library and TensorFlow as backend. The audio is recorded using the Python's sounddevice library. Python's multiprocessing library is used for integrating the modules and sharing the resources between the processes.

The simulation of traffic light is done using turtle, which is a pre-installed Python library that enables users to create pictures and shapes by providing them with a virtual canvas.

V. RESULTS

RetinaNet resnet50_coco_best_v2.1.0.h5 model is a standard and efficient model for object detection. It provides multiple bounding boxes and their categories simultaneously.

Fig. 3 shows the heat map of the confusion matrix generated when the model for siren detection was tested. The training accuracy was found to be 98%.

The project on the whole, was tested under various scenarios of normal and sparse traffic conditions. Siren sounds were provided and detected at various points of time to generate immediate green signal for that lane. Fig. 4 and Fig. 5 shows the simulation of the system for non-emergency and emergency scenarios.

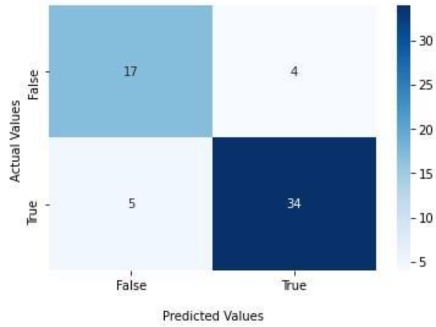

Figure 3. Confusion Matrix for the trained model

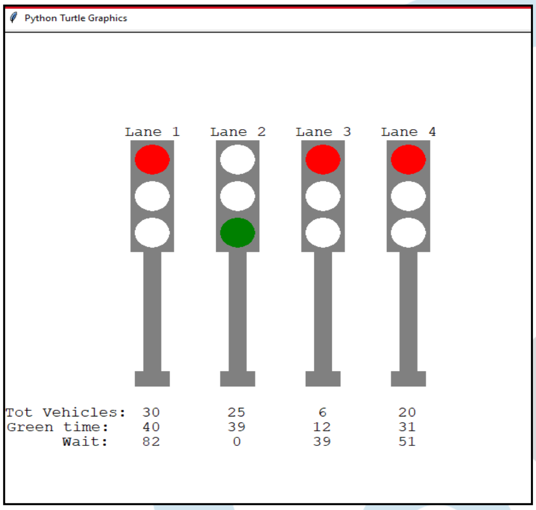

Figure 4. Traffic light simulation under Normal condition

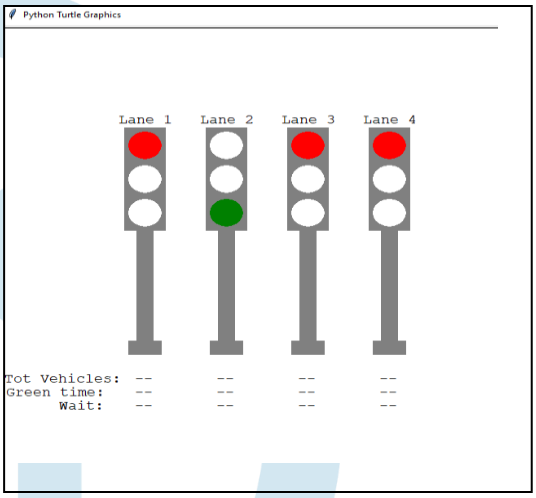

Figure 5. Traffic light simulation with priority to emergency vehicle

VI. CONCLUSION

The process implemented in regards to this project is useful to improve upon the traffic conditions faced by the people on a daily basis; such as the traffic congestion at the intersections, the long waiting times and even the delay for emergency vehicles. The vehicles are detected using the RetinaNet Model and classified into various categorical classes. RetinaNet is majorly used for image classification and object detection and proves to be a highly efficient model when coupled with ResNet50 as the backbone.

The Convolutional Neural Network model trained on the audio dataset containing siren sounds, detects and predicts if the vehicle is emergency or non-emergency, achieving a testing accuracy of around 86%. This accuracy can be improved by increasing the dataset size as well covering more cases. Once the siren sounds is detected, the traffic light is pre-empted to give priority to the lane with the emergency vehicle.

The system has scope for future enhancements in the way of increasing the accuracy of the model, which can be achieved by increasing the size of the dataset. Additionally, different categorical classes maybe added under emergency vehicles such as the fire-truck, police car and any other government vehicles which can also be further detected through image dataset.

VII. REFERENCES

[1] Tanwar, R., Majumdar, R., Sidhu, G.S. and Srivastava, A., 2016, January. “Removing traffic congestion at traffic lights using GPS technology.” In 2016 6th International Conference-Cloud System and Big Data Engineering (Confluence) (pp. 575-579). IEEE.

[2] Natafgi, M.B., Osman, M., Haidar, A.S. and Hamandi, L., 2018, November. “Smart traffic light system using machine learning.” In 2018 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET) (pp. 1-6). IEEE.

[3] Sharma, M., Bansal, A., Kashyap, V., Goyal, P. and Sheikh, T.H., 2021. “Intelligent Traffic Light Control System Based On Traffic Environment Using Deep Learning.” In IOP Conference Series: Materials Science and Engineering (Vol. 1022, No. 1, p. 012122). IOP Publishing.

[4] Suhweil, Y. and Al Yaman, M., 2017, October. “Smart controlling for traffic light time.” In 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT) (pp. 1-5). IEEE.

[5] Al-Ostath, N., Selityn, F., Al-Roudhan, Z. and El-Abd, M., 2015, July. “Implementation of an emergency vehicle to traffic lights communication system.” In 2015 7th International Conference on New Technologies, Mobility and Security (NTMS) (pp. 1-5). IEEE.

[6] Carmel, D., Yeshurun, A. and Moshe, Y., 2017, August. “Detection of alarm sounds in noisy environments.” In 2017 25th European Signal Processing Conference (EUSIPCO) (pp. 1839-1843). IEEE.