Traffic Light Detection for Information Systems and Telecommunications using CNN

Ramya Manasw V[1*], *SankaraBabu B*[1], *Madhavi Karanam*[1], *Soujanya R*[1], *Deepa Bisht*[2], *Kalyan Rao H*[1]
[1]Gokaraju Rangaraju Institute of Engineering and Technology, Department of Computer Science, Hyderabad, Telangana, India
[2]Uttaranchal Institute of Management, Uttaranchal University, Dehradun, India.

Abstract. Information and telecommunications using traffic signals plays major role in computer vision. Distinguishing objects which are in more modest size is a difficult task. We centre around an exceptional case: Detection and the classification of traffic lights in road sees and gives a guidance to the regulator for semi-autonomous and completely self-governing vehicles. We are introducing a profound learning approach for precise traffic signal discovery in adjusting a Single Shot Detection (SSD) approach SSD performs object proposition creation and order utilizing one single CNN. The first SSD battles in recognizing smaller objects, that which is fundamental for Traffic Light Detection (TLD). By our transformations it is feasible to recognize protests a lot more modest than the ten pixels without expanding the input picture size.

1. Introduction

Security is a significant part of computerized self-driving vehicles. Security isn't just for drivers and travellers yet in addition for walkers, different vehicles, and bikes. To get inescapable acknowledgment, the security issues should be cleared to get the full- fulfilment of individuals. Self-driving vehicles are currently being a master-piece of our transport organization. The quick improvement of car innovation centre to give us the best security includes and Automated Driving Systems (ADS) in vehicles can deal with the entire work of driving when the individual needs the vehicle to change to auto drive mode or at the point when the individual isn't certain of driving. Self-driving vehicles and trucks that will drive us, rather than us driving them, will become a reality. To accomplish every one of these things, object location is important. Detecting the object is broadly used in autonomous vehicles as a significant programming framework to distinguish objects as walkers, autos, traffic sign sheets and so forth in the event that of smashed and drive the situation is much more dreadful. The driver would lose his control and they will in general hit different vehicles, public and will not stop at traffic signals and this prompts major mishaps and even passing. To stay away from these situations, our vehicle framework should be independent. For these things to occur, the vehicles ought to have eyes, not the genuine eyes however, the cameras. Among these significant articles, traffic signals are likewise one of them. Identification of traffic signals is vital since it incorporates security of the general population. Since the above investigation shows that people will generally commit errors in observing the standards which incorporates keeping the traffic signal principles, prompts the portion of object recognition in self-sufficient vehicles. Previously, a wide range of calculations were utilized to distinguish traffic signals. De Charetteet proposed a three stages technique. At first, spotlight identification is executed in the dark level picture by utilizing formal hat morphological administrator to feature focused energy spots. At that point, a versatile layout matcher was utilized and detection is performed on the required areas of the traffic signal. Mu proposed a picture handling-based strategy of Change in the picture from red-green-blue (RGB) shading space to Holden Special Vehicles (HSV) shading space.

*Corresponding author: ramya.manaswi48@gmail.com

TLD accepts a caught picture as information and produces the bounding boxes as the result to be taken care of into the Classification model. We utilize Tensor Flow Object Detection(TFOD) API, which is an open source structure based on top of Tensor Flow to develop, prepare and convey object detection models. The Object Detection API additionally accompanies an assortment of discovery models pre-prepared on the COCO dataset that are appropriate for quick prototyping. In particular, we utilize lightweight model: ssd_mobilenet_v1_coco that depends on Single Shot Multi box Detection (SSD) structure with negligible alteration. Despite the fact that this is a universally useful recognition model, we see that this model addressed issues well, accomplishing the harmony between great bounding box precision and quick running time.

Subsequent to finding the bounding box for the traffic signal, we crop the picture to just incorporate the traffic signal, resize it to 32x32, and give this to the classification step. For training pictures, we utilize a mix of similar object detection and web scrapping. We use Keras (with Tensor Flow as backend) for preparing with a sum of 2,000 pictures (10% of which as approval/test tests). Aside standardization, we don't utilize other image augmentation methods. We prepared for 25 epochs and can accomplish >99% approval/test accuracy.

We take a two-stage ML based methodology in traffic signal classification. That is, the traffic signal Detection module comprises of two CNN based models:

1. Traffic Light Detection (step of localization) and

2. Traffic Light Classification.

We assembled our own model to guarantee the area and order of this test and we made our own street circuit and traffic signal sign to prepare and test the model and we gave directions as per the discovery of light.

1. Vehicle warns driving unit to slow down when yellow signal detected.

2. Vehicle just stops and will not provide any signal when red signal is detected.

3. Vehicle gives option to start when the green light is detected.

2. Literature Review

Tiagrajah V. Janahiraman proposed that strategies in AI for identifying traffic signals and order are supplanted by the new upgrades of ML(Machine Learning) object detection techniques by accomplishment of the building convolutional neural Network (CNN), which is a part of ML. This paper presents a ML approach for identification of traffic signal by looking at two object detection models and by assessing the adaptability of the Tensor Flow Object Detection Framework to tackle the real-time issues. [1]

Mahesh. G, Satish Kumar. T proposed that Autonomous vehicles are the booming research areas to come. They will open up the route for future progressed frameworks where PCs are relied upon to get control over the dynamics of driving. These vehicles are fit for detecting their surroundings and moving practically with no human interaction. The fundamental objective of this research is to detect a traffic signal continuously for independent vehicles. Aside from taking choices to explore in the correct way the Autonomous Vehicles (AV) significant assignment is to identify traffic signals, so it can submit to the traffic rules with adequate accuracy. The work completed in this examination utilizes two Artificial Intelligence strategy, these methods are analysed in achieving the assignment of traffic signal recognition in continuous. The two models that are planned and carried out are CNN and Deep Convolution Inverse Graphics Network (DCIGN). The outcomes plainly show that DCIGN out execution CNN by 8%. [2]

Self-driving vehicles can possibly reform metropolitan portability by giving practical, protected, helpful and blockage free movability. This vehicle self-sufficiency as a utilization of AI has a few difficulties like reliably perceiving traffic signals, signs, muddled path markings, walkers, and so forth these issues can be overwhelmed by utilizing the innovative improvement in the area of Deep Learning (DL), Computer Vision because of accessibility of Graphical Processing Units (GPU) and cloud stage. This paper is to propose a profound neural network-based model for solid discovery and acknowledgment of traffic signals utilizing move learning. The technique joins utilization of region-based CNN(R-CNN) Inception model V2 in Tensor flow for move learning. The model was prepared on dataset containing various pictures of traffic lights as per Indian Traffic Signals which are recognized in five kinds of classes. The model achieves its target by detecting the traffic signal with its right class type. The accuracy obtained by this model is 99%. [3]

ZhenchaoOuyang proposed that because of the inaccessibility of Vehicle-to-Infrastructure (V2I) correspondence in flow transportation frameworks, TLD is as yet thought to be a significant module in independent vehicles and Driver Assistance Systems (DAS). To beat low adaptability and precision of vision-based heuristic calculations and high force utilization of ML based techniques, we hereby propose a lightweight and constant traffic light identifier for the self-driving vehicle stage. Our model comprises of a heuristic traffic light detecting model to recognize all kinds traffic signals, and a lightweight CNN classifier to order the outcomes acquired. Disconnected simulations on the GPU worker with the gathered dataset and a few public datasets show that our model accomplishes higher normal exactness and less time utilization. By incorporating our indicator module on NVidia etson TX1/TX2, we direct on-street tests on the two full-scale self-driving vehicles stages (a vehicle and a transport) in typical rush hour gridlock conditions. Our model can accomplish normal recognition precision of 99.3% (mRttld), 99.7% (Rttld) at 10Hz on TX1 and TX2, separately. The on-street tests additionally show that our traffic signal recognition module can accomplish < ±1.5m mistakes at stop lines when we are working with other self-driving models. [4]

Aneesh A N proposed that self-driving vehicles are getting more advanced nowadays because of its protected, helpful and blockage free transportability. A large portion of the constant difficulties for AV driving like perceiving traffic signals, traffic signs, people on foot are being precisely tended to by the more up to date best in class calculations dependent on DL. On-going innovative interventions in cloud computing and the accessibility of very good quality cloud-based GPU speed up the advancement of AI calculations considerably. For continuous detection and acknowledgment of traffic signals, we propose Retina Net (a DNN design) based model through ML. Profound neural Network Retina Net was utilized as model and the framework was executed in Keras with Tensor Flow backend in Google Collaborator cloud state. The Retina Net model was prepared and assessed on Bosch Small Traffic Light Dataset containing traffic signal pictures of goal 1280 by 720 pixels, which falls under four kinds of classes. The model accomplished improved precision of detection and classification than other ML strategies for continuous activity. [5]

Karsten Behrendt proposed that Dependable TLD and classification is important for autonomous driving in metropolitan conditions. Right now, there are no situations that can dependably see traffic lights continuously, without map-based data, and in adequate distances required for smooth metropolitan driving. We propose a total framework comprising of a traffic signal locator, tracker, and classifier dependent on ML, sound system vision, and vehicle odometry which sees traffic signals progressively. Inside the extent of this work, we present three significant commitments. The first is a precisely marked traffic signal dataset of 5000 pictures for preparing and a video arrangement of 833 edges for assessment. The dataset is distributed as the Bosch Small TL Dataset and utilizes our outcomes as gauge. It is right now the biggest openly accessible named traffic signal dataset and incorporates names down to the size of just 1 pixel in width. The subsequent commitment is a TLD which runs at 10 edges each second on 1280 by 720 pictures. While choosing the certainty edge that yields equivalent blunder rate, we can identify traffic signals as little as pixels in width. The third commitment is a traffic signal tracker which utilizes sound system vision and vehicle odometry to process the movement gauge of traffic signals and a NN mentioned movement gauge. Got validation set accuracy for about 99%. [6]

The traffic issues in Lebanon are a significant concern for the population, according to Mohammad Belal Natafgi. Traffic emergency is supported by the increasing number of vehicles that exceed the limits of the streets, the failure of public transportation foundations, and the non-versatile traffic signal frameworks. Gridlocks afflict many Lebanon streets due to static green and red occasion's allocations that are insensitive to the current situation. The solution to this problem is a framework that adjusts to the types of traffic powerfully and updates traffic lights as well. Using real-time increasing traffic information, this paper implements a flexible traffic signal framework utilizing support learning. The framework is prepared and tested using a reproduction apparatus. In this device, the neural organization is able to collaborate with the intersection and reproduce it. Comparing the proposed model with the genuine traffic signal framework, the proposed model showed a 6.282% decrease in normal line lengths and a 56.37% decrease in normal lining times. [7]

Sahar Araghi proposed that Gridlock is one of the serious issues in present day urban communities. Q-learning and NN are applied here to set sign light occasions and limit all out delays. It is expected that a crossing point acts along these lines to an insightful specialist figuring out how to set green occasions in each cycle dependent on traffic data. Here, a correlation between Q-learning and NN is introduced. In Q-learning the hang of, considering consistent green time requires an enormous state space, making the learning interaction essentially inconceivable. Rather than Q-learning techniques, the NN model can undoubtedly set the proper green time to fit the traffic interest. The presentation of the proposed NN is contrasted and two conventional options for controlling traffic signals. Re-enactment results demonstrate that the use of the proposed strategy incredibly diminishes the absolute postponement in the network contrasted with the elective techniques [8].

https://doi.org/10.1051/e3sconf/202343001068

3. Methodology

3.1. Research Objective:

The objective of the research was to efficiently detect the TL efficiently form the input images and also to classify detected images based on light in the image detected. Based on the detected colour the caption of an autonomous car will be done. The images are given as input. There will be a camera placed and that camera will capture the images as per the view. Based on the input the working of the algorithm is done. There is a specific process which is designed for the purpose of implementing this concept.

The entire procedure of our proposed system of TLD is explained in the diagram format.

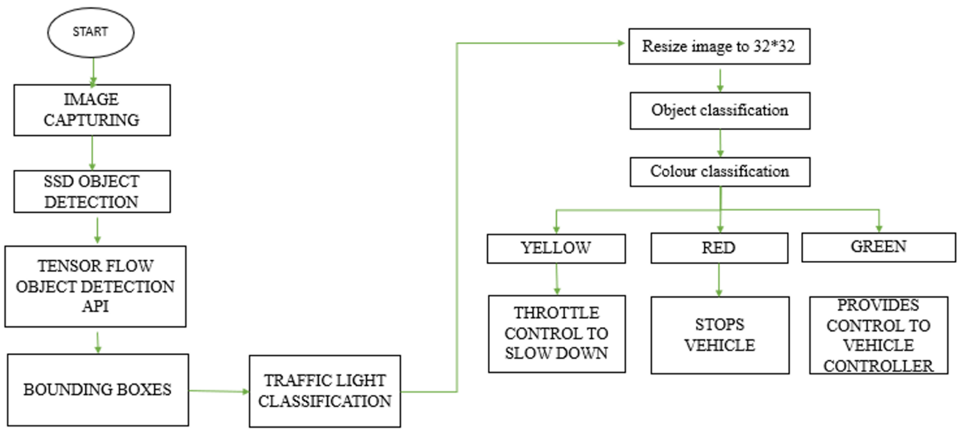

Fig.1.Algorithm Flow chart

3.2. Dataset:

Implemented this proposed system with the help of COCO dataset. The sample images in the dataset are the real time images of the road lanes and traffic lights. This dataset is embedded to the algorithm so that the proposed system will work. There will be many objects in the input data i.e., dataset. Our proposed system will only detect the necessary objects only with the help of ROI (region of Interest).

Fig.2. Traffic Light Dataset

3.3 Image Capturing:

In this step an autonomous vehicle will be moving on the road and there will be a camera placed at the driver place towards the road. The placed camera will be continuously taking pictures so that every object around the autonomous vehicle will be covered. This step is named as capturing of images. In the images proposed system will detect every object with the help of SSD and will recognize every object. If there are any traffic lights detected then the colour of the light will be detected and the colour will be classified and based on the colour the action of the vehicle will be done.

3.4. SSD Object Detection:

SSD is an algorithm which is used to detect multiple objects in the images at a time. At one shot only SSD will detect multiple objects in the image given. SSD is particularly very fast and works with high accuracy. It is a seep and high accurate object detection algorithm. The high accuracy and speed were acquired by the algorithm because of the following reasons:
1. Like RCNN SSD eliminates bounding boxes.
2. It predicts offsets in the bounding boxes and object categories with the help of convolutional filter.
High detection exactness in SSD was accomplished by utilizing numerous cases or channels with various sizes, and angle proportion for object location. It likewise applies these channels to various component maps from the later phases of an organization. This performs location at different scales. SSD with VGG16 based network is the standard for CNN architecture. For high quality image classification is done in SSD. VGG16 is used in SSD for the purpose of feature extraction. Expectation for the jumping boxes and certainty for various items in the picture is done not by one but rather by numerous element guides of various sizes that address different scales

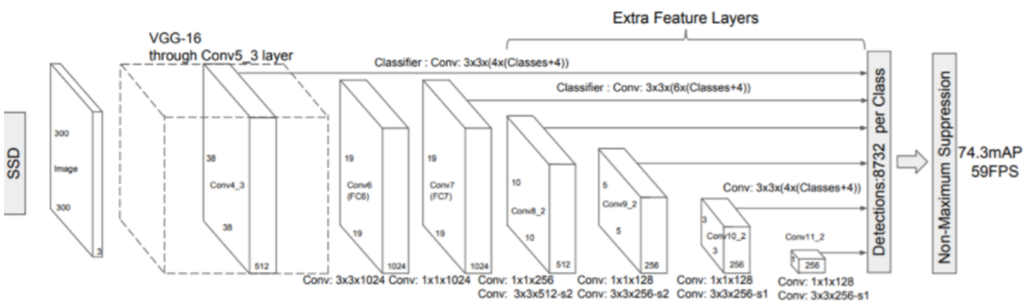

Fig.3.SSD Architecture

Continuously diminishing convolutional layers diminishes the element map size and increment the profundity. The profound layers cover bigger responsive fields and build more conceptual portrayals. This is useful is recognizing bigger items. Beginning convolutional layers cover more modest responsive fields and are useful in identifying more modest articles present in the picture. Contribution to SSD is an info picture with ground truth bouncing boxes for each article in the pictures as demonstrated underneath VGG-16 is the base organization that plays out the element extraction. Convolutional layers assess boxes of various viewpoint proportions at every area in a few element maps with various scales as demonstrated beneath. Multiboxes resemble anchors of Fast R-CNN. We have various default boxes of various sizes, angle proportion across the whole picture as demonstrated underneath. SSD utilizes 8732 boxes. This assists with discovering the default box that most covers with the ground truth jumping box containing objects.

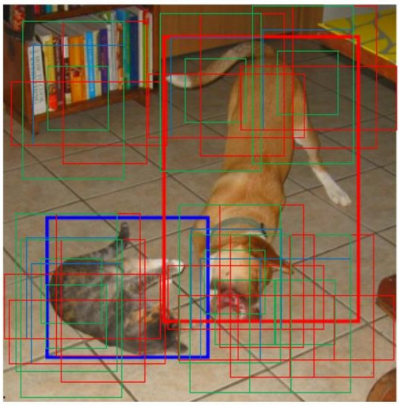

Fig.4. Default boxes

Matching Strategy:

During preparing time the default boxes are coordinated over viewpoint proportion, area and scale to the ground truth boxes. We select the crates with the most elevated cover with the ground truth jumping boxes. IOU (crossing point over joining) between anticipated boxes and the ground truth were to be more prominent than 0.5. We at last get anticipated box with greatest cover with ground truth.

3.5. Tensor Flow Object Detection API:

A Tensor Flow object detection API allows you to create a deep learning network to solve object detection problems. Model Zoo, their framework, already has pre-trained models. These models are trained on COCO, KITTI, and Open Images datasets. These can be used for the inference when only the categories in those datasets are desired. Additionally, they can be employed to initialize models when training on a new dataset. The various architectures that are used in the pre-trained model are given in this table.

Model name	Speed	COCO mAP	Outputs
ssd_mobilenet_v1_coco	fast	21	Boxes
ssd_inception_v2_coco	fast	24	Boxes
rfcn_resnet101_coco	medium	30	Boxes
faster_rcnn_resnet101_coco	medium	32	Boxes
faster_rcnn_inception_resnet_v2_atrous_coco	slow	37	Boxes

Fig.5. Architectures of pre-trained model

3.6. Bounding Boxes:

Bounding box is a non-existent square shape that fills as perspective for object recognition and makes a crash box for that article. Information annotators draw the square shapes over the pictures, illustrating that object of interest, inside every picture by characterizing X, Y organizes. This makes it simpler for AI calculations to discover what they're searching for, decide crash ways, and preserves important figuring assets. Bouncing boxes are perhaps the most well-known picture explanation methods in profound learning. Contrasted with other picture handling strategies, this technique can diminish expenses and increment comment proficiency. To recognize an article in a picture, the PC has to understand what it is and where it is. Accept self-driving vehicles for instance. An annotator will draw bouncing boxes around different vehicles and name them. This aides train a calculation to comprehend what vehicles resemble. Commenting on items like vehicles, traffic lights, and people on foot makes it feasible for independent vehicles to move occupied roads securely. Self-driving vehicle insight models depend intensely on bouncing boxes to make this conceivable.

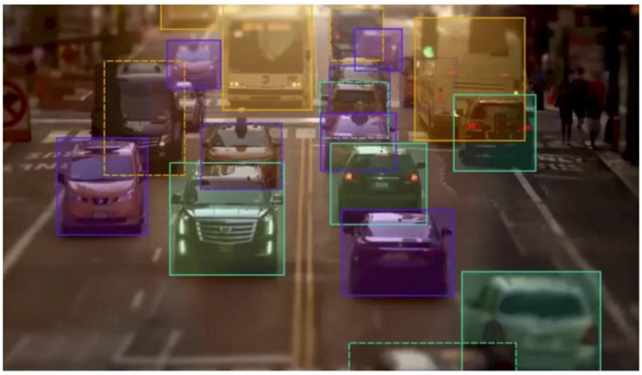

Fig.6. Bounding boxes

Bounding boxes will be generally used in healthcare systems, Agriculture, ecommerce, insurance claims, and Self-driving cars. These bounding boxes will be used in all the above areas in order to train the algorithm in order to identify the patterns.

3.7. Traffic Light Classification:

We utilize a basic CNN for our characterization. It comprises of three convolutional layers with (3x3 portion), the last two of which are trailed by a maximum pooling layer, a straighten layer, and two completely associated layers. For preparing pictures, we utilize a blend of web scratching, test system picture catching. With the exception of standardization, we don't utilize other picture expansion methods. We prepared for 25 ages and can accomplish >99% approval/test precision. We take a two-stage profound learning-based methodology in traffic signal characterization. That is, the traffic signal identification module comprises of two CNN based models pair:

1. Traffic signal recognition (restriction) and

2. Colour classification (light).

We assembled our own model to guarantee the identification and grouping of this analysis and we made our own street circuit and traffic signal sign to prepare and test the model and we gave directions as per the discovery of light.

•When the yellow sign is recognized the vehicle cautions the driver to back off.

•When the red sign is recognized the vehicle pauses and doesn't give any choice to continue and cautions to stop.

•When the green light is identified the vehicle gives an alternative to the vehicle regulator.

3.8. Resize Image to 32X32:

When image is captured by the camera and detected by the SSD then the image is further resized in order to reduce the complexity and increase the accuracy, speed. The image will be resized to 32X32. Image is resized in order to increase the speed, processing time, get desired accuracy.

3.9. Object Classification:

Picture order includes foreseeing the class of one article in a picture. Object localization alludes to recognizing the area of at least one item in a picture and drawing flourishing box around their degree. Item discovery joins these two errands and limits and orders at least one article in a picture.

3.10. Colour Classification:

After classifying the objects which were detected, now after detecting and classifying only relevant images are relevant objects will be considered and the objects of TRD will be given for colour classification phase. First colours are detected. Then after detection colours, colours will be classified according to the colour detected. There will be three colours detected. There are red, green, yellow.

 https://doi.org/10.1051/e3sconf/202343001068

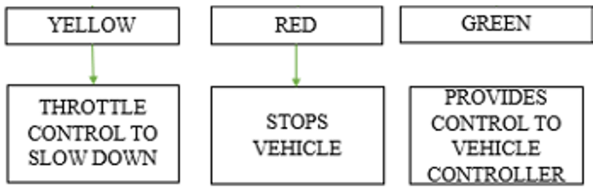

Fig.7.Classification

3.11. Functions used:

Accuracy:

The primary aim of the proposed model is building one specific model that will detect the traffic light from an autonomous car and classify the light using CNN and to do functionality as per the colour of the traffic light that has detected. The accuracy of the system is measured as how many times system has detected and done right for total no of detected colours.

Noise Removal Technique (NRT):

The initial step of pre-processing for the segmentation technique is the Noise Removal Technique (NRT). The dataset we are using has already been filtered to remove noise. Therefore, our images have already been pre-processed based on our own NRT. As part of the image enhancement process, some dark images present in the dataset. Histogram specification that is employed during the pre-processing stage, each image is determined by that image from the dataset. This figure was determined from 1% and 99% quarters. So the resulting image was normalized.

Data Augmentation (DA):

Data Augmentation strategy is utilized to handle variations of article sizes and shapes utilizing shearing, zoom in, zoom out, flipping, editing and so forth Utilization of information expansion makes the model more interesting to different info object sizes and shapes. This assistance improve the precision of the model. Each preparation picture is arbitrarily inspected by one of the accompanying choices utilize the whole unique info picture.

Region of Interest (ROI):

AROI is segment of the image on which channelling or other manipulation needs to be performed. Create a parallel coverage to characterize the ROI. This is a pair of images of the same size as the image that needs to be measured, with the pixels that characterize the ROI set to 1 and all remaining pixels set to 0. Multiple ROIs in an image can be characterized. Districts can be geographic in nature, such as polygons surrounding touching pixels, or they can be characterized by various authorities. In the latter case the pixels are not really contiguous.

Fig.7. Region of Interest

CNN Configuration:

Here simple CNN is used for the purpose of classification. There are three CNN layers with 3X3kernel and the last two were followed by maxpooling layer, two full connected layers, a flatten layer.

4. Results

The received results were obtained by the training the ML algorithm with the COCO data that we have got. After getting the bounding boxes relevant features were extracted and traffic lights will be detected. Here by applying SSD for detection and CNN for classification, got 100% accuracy and there is no data loss.

5. Conclusion

Future work was expected to make it completely computerized. Our investigations proposed that profound learning approach in object identification assignments conveys a phenomenal outcome. In this we can ready to accomplish higher exactness through profound learning approach. In this task we have done identification of traffic signals in two stages precisely I e., location of lights and characterization of lights.

References

1. T. V. Janahiraman and M. S. M. Subuhan, "*Traffic Light Detection Using Tensor flow Object Detection Framework,*" in proceedings of the IEEE 9th International Conference on System Engineering and Technology, (ICSET), (2019), pp. **108-113**, doi: 10.1109/ICSEngT.2019.8906486."
2. M. S. K. T. G, (IJITEE, 2020).
3. R. Kulkarni, S. Dhavalikar and S. Bangar, "*Traffic Light Detection and Recognition for Self Driving Cars Using Deep Learning*," in proceedings of Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), (2018).
4. Z. Ouyang, J. Niu, Y. Liu and M. Guizani, "*Deep CNN-Based Real-Time Traffic Light Detector for Self Driving Vehicles,*" in proceedings of IEEE Transactions on Mobile Computing, (2020).
5. A. N. Aneesh, L. Shine, R. Pradeep and V. Sajith, "*Real-time Traffic Light Detection and Recognition based on Deep RetinaNet for Self Driving Cars*," in proceedings of 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT)(2020).
6. K. Behrendt, L. Novak and R. Botros, "*A deep learning approach to traffic lights: Detection, tracking, and classification* "in proceedings of IEEE International Conference on Robotics and Automation (ICRA), (2017).
7. M. B. Natafgi, M. Osman, A. S. Haidar and L. Hamandi, "*Smart Traffic Light System Using Machine Learning*," in proceedings of IEEE International Multidisciplinary Conference on Engineering Technology (IMCET), (2018).
8. S. Araghi, A. Khosravi, M. Johnstone and D. Creighton, "*Intelligent Traffic Light Control of Isolated Intersections Using Machine Learning Methods*", in proceedings of IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, (2013).
9. Chandrika Lingala, and Karanam Madhavi, "*A Hybrid Framework for Heart Disease Prediction Using Machine Learning Algorithms* ", in proceedings of E3S Web of Conferences, ICMED (2021).
10. V. Tejaswini Priyanka, Y. Reshma Reddy, D. Vajja, G. Ramesh and S. Gomathy "*A Novel Emotion based Music Recommendation System using CNN"*.in proceedings of 7th International Conference on Intelligent Computing and Control Systems (ICICCS),(2023).
11. Chandrika Lingala, and Karanam Madhavi et.al, "*A Survey on Cardiovascular Prediction using Variant Machine learning Solutions.*" In proceedings of E3S Web of Conferences ICMED (2021).