IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Deep Learning Technique for Real-time Traffic Light Detection by Automated Vehicles

Priyanka S. N[1*], Shashidhara H. S[2]

[1]Dept. of ISE, Ramaiah Institute of Technology, Bangalore-54
[2] Dept. of ISE, Ramaiah Institute of Technology, Bangalore-54

**Corresponding Author: priyankasn02@gmail.com, Tel.: +91-9148329764*

Available online at: www.ijcseonline.org

Accepted: 23/Jul/2018, Published: 31/July/2018

Abstract— In urban environment the trusted traffic lights identification and classification is very important for automated driving vehicles. In urban driving presently there is no such systems that has to become aware of dependable traffic lights in real time and we can find the traffic lights in automated vehicles without map based information and enough distance require for regular surface in urban driving. Here we suggest a complete system made up of traffic light detector, tracker, and classifier based on deep learning, stereo vision, and vehicle odometer which become aware of traffic lights in real time The first is a precisely marked traffic signals data sets contains 5000 guiding pictures and 8335 motion pictures for development. Tiny traffic signals dataset is distributed by Bosch Company is used as a basic tool. The traffic signal identification which works at 10 frames per second on 1280*720 images is the second achievement. Third achievement is a traffic signal which identifies utilizes stereo vision and odometer of vehicles to calculate the moment computation of traffic signals.

Keywords: Convolution Neural Network , Odometer, Stereo vision, , Traffic Light, GPS Detector, AdaBoost Algorithm, RGB Conversion, Pixels, Grey scale Images, Dimensional Matrix

I. INTRODUCTION

In the present world automated vehicles are playing an important role. Many researches have-been done in the traffic light detection for automated vehicles. The vehicle must recognize traffic lights and should work under the traffic management system. There are some drawbacks in the previous research ambiguity of image such as decorated lightings in the image, blurriness in the traffic signals etc.., Some research has only done related to suspended traffic signals experiment during dull weather, overcast sky with light rain etc. The aim of this paper is "detection of traffic lights, classification and tracking of automated vehicles" used to overcome the drawbacks present in the current technology. Automated vehicles [1] are driver less vehicles which uses sensors to sense the surrounding environment and drive the vehicles without manual efforts. Eg, Google self-driving cars, motorway pilot, Traffic jam pilot, Piles peak etc. One of the features of the automated vehicles is to sense the traffic light and guide the vehicle accordingly. It means the automated vehicle must be able to stop during red signal and proceed during green signal. This feature is defined as detection of traffic lights. Suspended traffic lights are much easier to detect traffic lights in automated vehicle because the background of suspended traffic lights is usually static in color, and normally it includes a sky area. Supported traffic lights are which having backgrounds like buildings, trees and any images or things etc. Supported traffic lights are difficult to identify in automated vehicles.

The paper is organised as follows. Section I contains the introduction of different traffic light detection in automated vehicles. Section II contains related work of proposed topic, Section III describes the system overview, Section IV describes the traffic light classification, Section V describes the results and evaluation, Section VI describes the conclusion.

II. RELATED WORK

Lindner F. et al. explained the different methods to detect and recognize traffic lights in his paper "Robust Recognition of traffic signals" using Different methods such as colour detector, with the help of GPS detector and using AdaBoost algorithm in urban environment. Most of the existing work was done on suspended traffic lights. Several algorithms were used previously such as colour segmentation using hidden Markova models for detection of traffic lights. In this paper however, as opposed to this strategy the main identifier which meets real-time limitations in their exploration is a mixture of "costly colour sensor" [1] and differential GPS delineate. In this system the detection of traffic lights is more flexible but less depended on the sensor quality. This paper recommends

not using colour in the detection step to meet up the global requirements as it is costly.
Many existing studies used spot light detection, colour threshold, map information and template matching to detect traffic light and their classification [2]. [2] Supported both suspended and supported traffic light systems. The main issue with this system is positioning of the traffic lights because of which it is difficult for 2-D based vehicles to determine traffic signals during cross roads.
Using the current technology of machine learning and deep neural network such as deep learning used in image classification [3], pixel-precise object segmentation and end to end object detection the performance can be increased. A deep neural network forms the neurons and hidden layers in between input layer and output layer. Neural network creates 4*4-pixelregions that can be in network. Based on connected network the output probabilities, cropped traffic lights in images are fitted for detecting. [3] Reports a review of 91.4% with a crossing point over association of 0.25 for movement lights bigger than 8 pixels in width.
The disadvantages of deep neural networks presently are they need huge amount of training data set. Existing systems they used existing Lara dataset for traffic light detection but this dataset has only lower resolution images so accuracy is insufficient.
To solve these problems, we suggest a system for detection, tracking and classification of traffic lights for automated vehicles, which make use of deep learning. The detection of traffic lights is passed out by an end to end trained neural network which is used to detect traffic lights as small as 3*10 pixels. The tracker detects on-board sensing of autonomous vehicles, as it makes use of stereo imagery [4][5] for triangulating the position of traffic lights in the odometer information and 3D world to calculate the relative movement of traffic light with respect to the vehicle. The tracked location is then adjusted by a neural network which is trained exactly for that purpose.

III. SYSTEM OVERVIEW

Our main aim is to detect supported and suspended traffic lights in autonomous vehicles in urban, rural and semi-urban environments. Our traffic light detection method consists of 8 stages as appeared in fig-1. First step is bounding box detector that finds traffic lights in the images. After detecting the traffic light in the images, we use classifiers that remove false positives and predict the lights states. Detected traffic light is then tracked over multiple frames and at last confirmed by the classifier once again. Within the following, we will get in every of these steps.

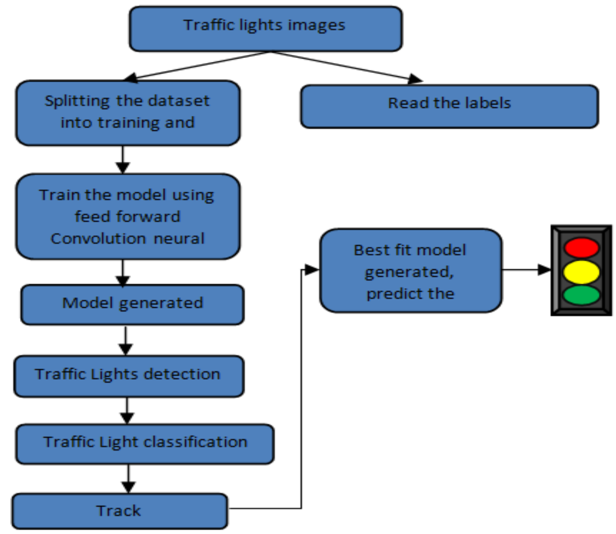

Fig-1: Traffic light detection, classification and tracking steps.

Traffic light dataset contains 13427 camera pictures at a resolution of 1280*720 pixels and contains regarding 24000 annotated traffic lights. The notation contains bounding boxes of traffic lights along with current state of traffic light.

3.1 Traffic Light Images – Splitting Into Testing And Training Data Sets

To train the data, machine learning approach inherently concludes deep learning concepts which are necessary part of the existing work. Here we are using Bosch small traffic lights dataset, to tacking, classifying and detecting the traffic lights we are using labelled dataset which was accurately detects the RGB colour images which are included in dataset. A traffic lights dataset image has 1280 *720 pixels. Robert Bosch small traffic light dataset contains 22790 images in that we are splitting 14236 images for training and 8334 images for testing purpose. The 12 bit HDR images are taken from the camera using a red-clear-clear blue filter and as recreated 8-bit RGB colour images. For debugging and also for training purpose RGB images are taken. Some drawbacks are there while processing RGB conversion. A number of the converted images may include artefacts and the colour spreading might also seem unusual. Totally 10,760 traffic lights are labelled images out of 22790 images.

3.2 Training The Dataset Using Convolution Neural Network And Detection Of Traffic Lights.

Since the (Robert Bosch Traffic Light) dataset has labels in the images hence, the system considers supervised learning technique. Supervised learning is a type of machine learning technique. Training algorithm of a neural network, supervised learning is best fit to use. Because in supervised learning technique already known desired output while training, one of the input image is given to the net's input layer. The label dissemination of the various traffic signal states is massively inclined towards the three common traffic signal states which square measure 'Red', 'Green', 'Yellow'. If frequency of captured camera and the traffic light refresh rate differences, then more traffic lights often look as to be off. If this is the condition, then we elected for labelling them as 'off' as a replacement of their existing state. Different types of labels of traffic lights such as red, red straight, red straight left, red left, red right, yellow, green, green straight, green straight left, green straight right, green left, green right, off. The data set for testing is collected from California Avenue university [6][7]. The stereo video data sets collected are then labelled every 15.5 seconds which is contrary to training data sets.

3.2.1 Convolution Neural Network:

Convolution neural network is a deep learning class which is the best technique for image processing. Image is converted into grey scale image. We have used convolution neural network to classify input image into three categories Red, Green and Yellow. Initially every traffic input image is shown as matrix of pixel values. There are two components considered for images: channel and grey scales.

Channel represents particular component of an image. Generally, a digital image will have three channels Red, Green and blue. These colours are represented as 2-dimensional matrix each with a value ranging from 0 to 255 pixels. Contrary to channel images **Greyscale** images will have 1 channel which is represented by one 2-dimensional matrix. The value of this matrix will range from 0 to 255 where 0 indicates black and 255 will indicate white.

Fig-2: (a) Traffic light color image. (b) Gray scale image.

(c) Cropped traffic signal image. Convolution

Convolution networks here are used to derive features from the images. Here the input data is divided into small squares and their features is learnt by convolution techniques. Every image is treated as a matrix of pixel values. Suppose consider a filter of 5*5 matrixes where its values are filled with 0 to 255. 0 indicates black and 255 indicates white. Consider one more testing image of size 3*3 matrixes. We stride the filter image over the testing image by 1 pixel each. Now using ReLU approach we compute the values of the newly formed convolution matrix. First, we will calculate the product of first value from training matrix and testing matrix and then we will calculate the sum of 1 in the entire matrix and fill it in the convoluted matrix. These filters act as feature to detect the traffic lights from the testing images. Stride one filter over one image gives one feature to the output layer similarly striding 2000 filters gives 2000 features. These features are then combined to form high level traffic features.

Pooling is the process of reducing the dimensionality of the convoluted matrix. Here we use Max pooling approach where a spatial neighbourhood of n*n matrix is defined and slide over the convoluted matrix. This spatial matrix takes the maximum value from that window and slides by 2 cells. But pooling will make sure of keeping the important information from the convolution matrix.
Some of the important functions of pooling are:

- Pooling helps in reducing dimensionality of the input images into required traffic features.
- Pooling helps in controlling over fitting and reduces the number of computations.
- Pooling is resistant to variations in the input images. Even though there are disturbances in the input images the performance of pooling is not changes.
- Scaling helps in locating even small traffic lights from the images even though their scalability is very small.

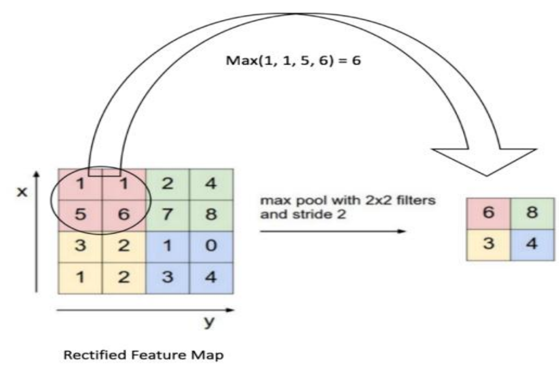

Fig-3: Max pooling

Fig 3 shows the series of convolution and pooling steps applies repeatedly on the data set to obtain the final traffic feature. Convolution is again applied on the pooling output using feature matrix and then ReLU is performed again individually on these features finally max pooling is applied. Finally, at the end of this traffic light detection phase we will be able to predict the traffic feature.

ReLU (Rectified Liner Unit):

ReLU has been used under after the convolution operation. It is also called as non-linearity. ReLU applied per pixel and replaced all the negative value into zero. ReLU is non-linearity in our ConvNet. It also act as a activation function to allow the convolution neural network to train the dataset faster.

F(x) = Max (0, x) and range: [0, infinity].

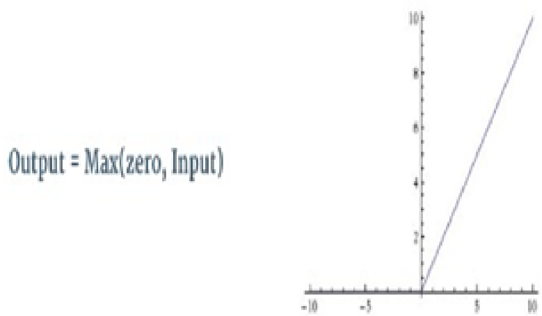

Fig-4: The ReLU operation

IV. TRAFFIC LIGHT CLASSIFICATION

Traffic lights classification is done using neural networks. First traffic lights are detection using traffic light detection phase with different heuristics and colour thresholding, then the detected lights are classified using different neural layers. All traffic light images are extended and rescaled into 20px wide and cropped into 64*64 pixels. Nearly 22px of context on left and right extra context helps for classification. Different classifiers are used to decide between “Red”, “Green” or “yellow” and reduce number of false positives.

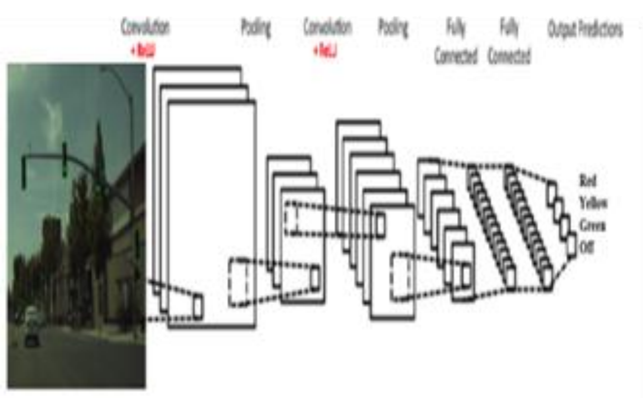

Fig 5: Traffic feature extraction using convolution neural network. Tracker

Tracker is an object that can be applied after classification which is optional. Tracking is used to find the precise region of the traffic lights. Tracker is used to solve multiple target tracker problems because there may be similar features that looks like the traffic lights for e.g.., a car break light may be detected as a traffic feature. To avoid this kind of misconceptions we use tracker to precise the required features

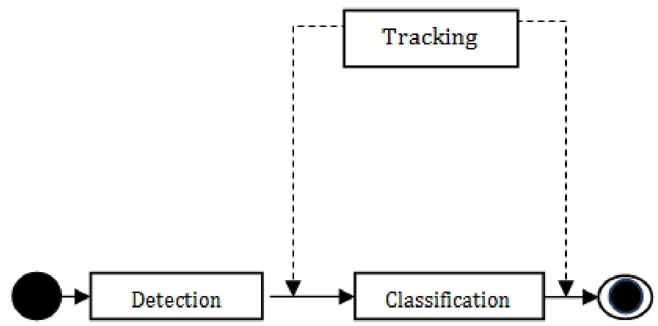

Fig-6: Tracking traffic light flow.

V. RESULTS AND EVALUATION

The wide range of methodologies and measurements has been utilized to assess detector execution. For evaluation on the Robert Bosch small traffic dataset we recommend utilizing: precision and recall, which are categorised in condition [1], and [2]. TP abbreviation for true positive, FP abbreviation for false positive and FN abbreviation for false negatives. The TPs, FPs and FNs ought to be assessed on a per frame basis.

$$Precision = \frac{TP}{TP+FP} \qquad \text{eq(1)}$$

Precision is the ratio of the correct traffic light detections to the total of detections made by the detector.

$$Recall = \frac{TP}{TP+FN} \qquad \text{eq(2)}$$

Recall is the ratio of correct traffic light detections to the actual number of traffic lights. For showing and assessing the general framework execution, we propose producing an accuracy review curve and utilizing the area-under-curve (AUC) as execution measure.

A. VALIDATION ACCURACY

The figure 7 shows the validation accuracy[1] and validation lose[2] for the first 10 convolution steps applied on the Robert Bosch small traffic light data set. From the graph it can be predicted that the accuracy of the performance is increasing from one layer to another layer. Here the x axis denotes number of layers the images have been processed. Y axis shows the accuracy percentage for the data set.

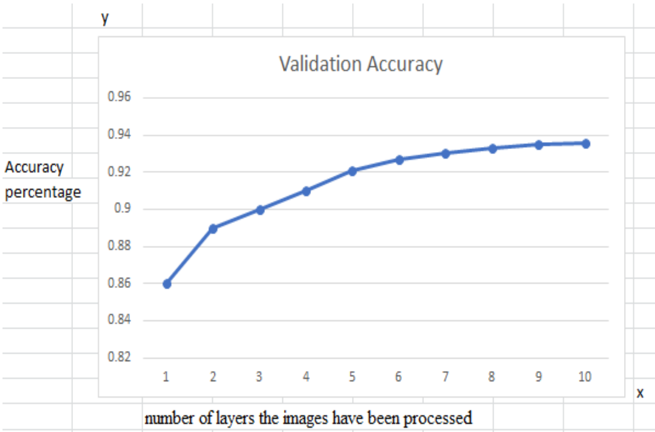

Chart -1: validation accuracy

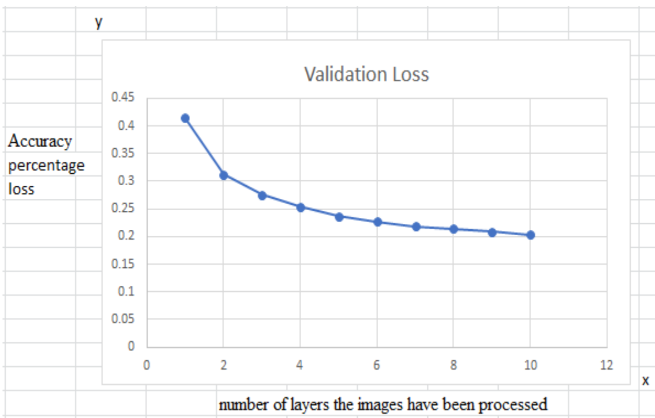

Chart -2: validation loss

VI. CONCLUSIONS

Deep learning techniques are invoked to detect traffic lights. Data set images of size 1280*720 pixels are considered. Bosch traffic lights data set is considered for evaluation of IOU of 0.5. We train the neural network for traffic light detection compressing to 4*12 pixels. Then at the second stage of neural network identifies traffic light states and identifies false positive. Now we choose odometer, deep learning techniques and stereo visions in place of detectors. This method achieved a high precision and accuracy even in challenging conditions. The implemented experiment provided the accuracy of 94.4%.

Fig -7: Traffic Light Detector in Above Images

ACKNOWLEDGEMENT

We are indebted to Robert Bosch for providing the required Traffic Light dataset for experimenting.

REFERENCES

[1] C. Wang, T. Jin, M. Yang, and B. Wang, "Robust and real-time traffic lights recognition in complex urban environments," International Journal of Computational Intelligence Systems, vol. 4, no. 6, pp. 1383–1390, 2011.

[2] R. de Charette and F. Nashashibi, "Real time visual traffic lights recognition based on spot light detection and adaptive traffic lights templates," in Intelligent Vehicles Symposium, 2009 IEEE. IEEE, 2009, pp. 358–363.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolution neural networks," in Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States., 2012, pp.1106–1114.

[4] V. Haltakov, J. Mayor, C. Unger, and S. Ilic, "Semantic segmentation based traffic light detection at day and at night," in German Conferenceon Pattern Recognition. Springer, 2015, pp. 446–457.

[5] M. Diaz-Cabrera, P. Cerri, and P. Medici, "Robust real-time traffic light detection and distance estimation using a single camera," ExpertSystems with Applications, vol. 42, no. 8, pp. 3911–3923, 2015.

[6] M. P. Philipsen, M. B. Jensen, A. Møgelmose, T. B. Moeslund, and M. M. Trivedi, "Traffic light detection: a learning algorithm and evaluations on challenging dataset," in Intelligent TransportationSystems (ITSC), 2015 IEEE 18th International Conference on. IEEE, 2015, pp. 2341–2345.

[7] G. Siogkas, E. Skodras, and E. Dermatas, "Traffic lights detection in adverse conditions color, symmetry and spatiotemporal information. "in VISAPP (1), 2012, pp. 620M-627.

[8] N. Fairfield and C. Urmson, "Traffic light mapping and detection, "in Robotics and Automation(ICRA), 2011 IEEE international conference on. IEEE, 2011, pp. 5421-5426.

[9] M. Diaz-Cabrera, P.Cerri, and J. Sanchez-Medina, "Suspended Traffic lights detection and distance estimation using color features", in 2012 15th international IEEE conference on Intelligent Transportation Systems. IEEE, 2012, pp. 1315-1320.

Authors Profile

Priyanka Navalagund, She received B.E degree in Information Science from VTU University. She is pursuing M.Tech in Software engineering at Ramaiah Institute of Technology, Bangalore. Her research interest is in Machine Leaning, Software Engineering, Cloud Computing, Data Science, Software Testing.

Shashidhara H S received B.E. Degree in Computer Science & Engineering from Bangalore Institute of Technology, Bangalore University in 1999, and M.Tech. Degree in Computer Science & Engineering from M S Ramaiah Institute of Technology, Visvesvaraya Technological University in 2004. He is a Part-time Ph.D. research scholar at Jawaharlal Nehru Technological University, Hyderabad. He has around 16 years of work experience in academic and industry. He has worked as a developer for M.M. Softech, Bangalore. Currently, he is working as Associate Professor in Department of Information Science & Engineering, M S Ramaiah Institute of Technology, Bangalore. He has published research papers in national and international conferences and Journals. He has organized several seminars, workshops, delivered lectures in training programmes and given some guest lectures. His research interests include various techniques of Data Mining and Application of those techniques in the field of Bioinformatics. He is recognized by Visvesvaraya Technological University as a subject expert of Microprocessor and has given a series of lectures on microprocessors and assembly language programming which was transmitted live through EDUSAT. He was also a subject expert of Software Architectures when the subject was introduced first in the university and delivered lectures again through EDUSAT..

