Traffic Light Detection and Recognition for Self Driving Cars using Deep Learning

Ruturaj Kulkarni
Dept. of Electronics and Telecommunications
Pune Vidyarthi Griha's College of Engineering and Technology
Savitribai Phule Pune University
Pune, India
ruturajkulkarni4@gmail.com

Shruti Dhavalikar
Dept. of Electronics and Telecommunications
Pune Vidyarthi Griha's College of Engineering and Technology
Savitribai Phule Pune University
Pune, India
dh.shruti25@gmail.com

Sonal Bangar
Dept. of Electronics and Telecommunications
Pune Vidyarthi Griha's College of Engineering and Technology
Savitribai Phule Pune University
Pune, India
sonalsbangar@gmail.com

***Abstract*—Self-driving cars has the potential to revolutionize urban mobility by providing sustainable, safe, convenient and congestion free transportability. This vehicle autonomy as an application of AI has several challenges like infallibly recognizing traffic lights, signs, unclear lane markings, pedestrians, etc. These problems can be overcome by using the technological development in the fields of Deep Learning, Computer Vision due to availability of Graphical Processing Units (GPU) and cloud platform. In this paper, we propose a deep neural network based model for reliable detection and recognition of traffic lights using transfer learning. The method incorporates use of faster region based convolutional network (R-CNN) Inception V2 model in TensorFlow for transfer learning. The model was trained on dataset containing different images of traffic signals in accordance with Indian Traffic Signals which are distinguished in five types of classes. The model accomplishes its objective by detecting the traffic light with its correct class type.**

Keywords— Self Driving cars, Autonomous Driving, Traffic Light detection, Deep Learning, ADAS, Faster R-CNN, Inception-V2, TensorFlow, Object Detection, Classification.

I. INTRODUCTION

In the recent years, several technological firms and universities have shown keen interest in autonomous vehicles or self-driving cars by investing immense resources both technically and financially for the research and development. Some of them with ongoing extensive research are Alphabet's subsidiary Waymo, PilotNet of NVIDIA, Delphi with an MIT based start-up nuTonomy, General Motor's Cruise, Tesla 'S' model of Tesla Motors, BMW in collaboration with BAIDU, Ford's Agro AI, etc. So, the future of autonomous driving is near for its mainstream provided it overcomes the technological, practical challenges along with economic, social, legal acceptance.

The combination of traffic signals and road signs determine a visual language that forms set of rules whose interpretation aids for disciplined driving. In the field of autonomous driving the perception of traffic discipline has high industrial potential. The vision system helps to analyze current traffic situation on the road, danger and difficulties near the vehicle, warn and help them for safe, convenient and healthier navigation by providing the useful information. But easier said than done the vision-based object detection and recognition in traffic scenes is still a major challenge to be successfully overcome by the autonomous driving industry. Variation due to lighting and weather condition, random and dynamic scenarios at any moment contributes to the challenges for object detection function.

Deep learning integrating Computer Vision has the potential to engender reasonably affordable, robust solutions for autonomous driving industry. In 2012, Krizhevsky et al., rejuvenated interest in CNNs by showing significant progress in accuracy of image classification at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [1] [2]. Preceding to the use of CNNs, majority of the pattern recognition tasks were accomplished using a hand-crafted feature extractor accompanied by a classifier [3]. By localizing objects with a deep network and training a high-capacity model with only a small quantity of annotated detection data, the use of CNN gives remarkable better object detection performance than the systems based on simpler HOG-like features [4].

So, for Traffic Light (TL) detection problem in the self-driving cars, we have presented a deep learning based solution that reliably detects TL and recognizes the type of TL according to Indian traffic light system which outperforms traditional image processing methods. The arrangement of rest of the paper is as follows: Section II describes the literature survey which enlightens about various existing methods. Section III gives detailed information about the implementation of proposed method. Section IV demonstrates the experimental results. Section V concludes the overall work enlightening about the future work that can be carried out further.

II. RELATED WORK

To stay updated with ongoing research and monitor present-day movements in TL detection, a literature survey has been carried out exploring new and creative approaches towards TL Detection. Methods for TL detection are often classified as image processing based, machine learning based or map-based techniques [5]. In the image-processing based method, a single or multiple amount of actions or operations are performed on the image in order to achieve a particular resultant. Dwi H. Widyantoro and Kevin I. Saputra achieved TL detection and recognition with the help of Color Segmentation and Circle Hough Transform [15] whereas Guo Mu achieved the same with RGB to HSV conversion, filtering, histogram of oriented gradients (HOG) features and support vector machine (SVM) [6]. Swathy S Pillai engineered a system detecting tail lights for analysing traffic during night using various morphological operations like thresholding, filtering, extraction, etc [14]. Zhenwei Shi discussed about Adaptive Background Suppression filters as a fast and robust method for TL detection under different illumination conditions [7].

978-1-5386-5257-2/18/$31.00 ©2018 IEEE

Though image processing approach is quite straight and uncomplicated, it undergoes critical phases such as thresholding, filtering. Fine miscalculations or slight deviations from standards in these phases may lead to ambiguous outcomes which is strictly undesirable in sensitive cases of TL detection. To tackle this pitfall, machine learning based methods and algorithms are tried over singly or in combination with ample processing techniques to prune the misleading directions. For example, Keyu Lu proposed Generalized Haar Filter based Convolution Neural Network (CNN) feasible to be deployed for object detection in traffic scenes [9]. Seokwoo Jung developed CNN-based traffic sign recognition algorithm where extraction of traffic sign candidates is performed in first stage and classification with LeNet-5 CNN architecture takes place in further stage [8]. Gwang-Gook. LEE and Byung Kwan PARK attained faithful outcomes for TL recognition by combining conventional approach of image processing with Deep Neural Network (DNN) as a promising classifier [10]. Also, Karsten Behrendt put forth deep learning approach incorporating stereo vision and vehicle odometry for TL detection, tracking and classification [11]. Masamitsu Tsuchiya revealed an efficient Hybrid Transfer Learning method for object detection giving aids to unleash the untried learning methods [12].

To obtain optimized results in this learning based technique, one must collect variety of large training datasets and train the model for significant amount of time. Addressing this bottleneck, in recent years, map-based techniques are employed to achieve accurate outcomes. TL detection and recognition oil the wheels of Advanced Driver Assistance Systems (ADAS) in turn the make of INTELLIGENT vehicle.

III. Methodology

In this paper, the execution of proposed system is distributed into following manner: collecting images of Indian traffic lights, pre-processing images to generate the dataset, training the CNN for detection and recognition of traffic lights and finally validation of model through experimental results.

A. Dataset

A collective set of required images according to Indian streets is not yet published in standard format for evaluation of the model. Hence, the collection of images and video sequences is done by using a 16 Megapixel camera which supports image resolution of 4616 × 3464 pixels and video recording resolution of 1920 × 1080 pixels at frame rate of 30 fps. In total, 1237 frames have been taken in the daylight conditions captured in the urban streets and suburb roads of Pune, Maharashtra, India.

The pre-processing of these selective frames is done by resizing them to size of 600 × 800 pixels which are then used for labelling. There are five classes namely red, yellow, straight, left, right which are used for classification of TL. Each frame is labelled manually based on the type of the class it contains. Fig.1 shows these labelled images used as dataset for training and testing of the model.

Fig. 1. Examples of labelled images according to dataset

B. Traffic Light Detection model

The system incorporates transfer learning based pre-trained method wherein Faster R-CNN-Inception-V2 model is used. Transfer learning is a machine learning technique where there is enhancement in learning of a new task by channeling the knowledge through a related task that has formerly been learned. It is a method in deep learning that allows to eliminate the extensive computational and time resources that are essential to develop neural network models by using pre trained models.

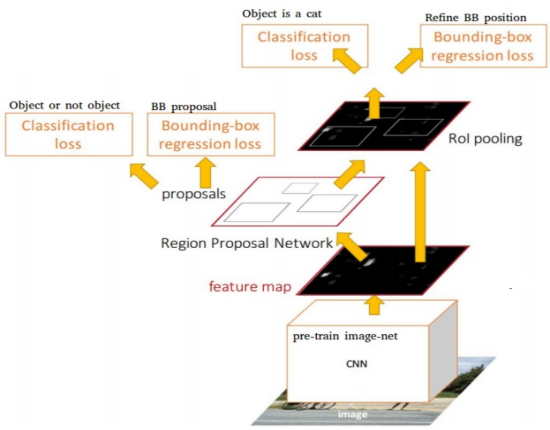

Fig. 2. R-CNN architecture [13]

There are several object detection architectures available like Single Shot Multibox Detector (SSD), Faster Region-based Convolutional Neural Networks (R-CNN), Region-based Fully (R-FCN) which incorporates feature extractors like ResNet-101, Inception-V2, Inception-V3, MobileNet, etc. The selection of architecture and feature extractor is tradeoff between speed and accuracy that your application needs.

For the Traffic Light detection considering the application requirement and available computational resources, Faster R-CNN Inception-V2 model is used which serves the accuracy and speed tradeoff. The model is trained on the above mentioned dataset where loss is reported at each step of training. The model is trained on the NVIDIA GEFORCE 940M GPU using TensorFlow.

IV. Results

After training the model for 120,000 iterations which took nearly 12 hours for our used hardware it reported a loss in the range of 0.01 which is better compared to specified acceptable loss which is in the range of less than 0.05.

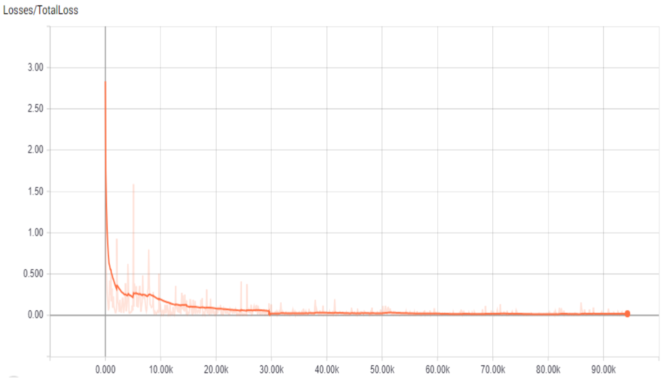

Fig. 3. Loss graph of model

The model successfully detects the traffic light and classifies it according to its type.

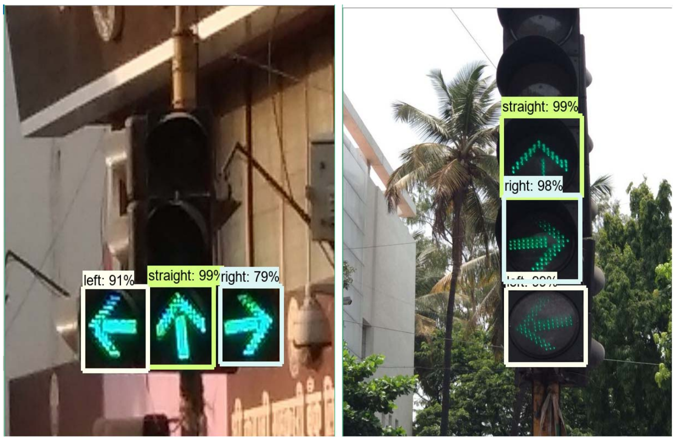

Fig. 4. Green TL classified as straight, right, left

Fig. 5. Detection of TL in noisy environment and unclear frames

Fig. 6. Yellow TL and green in noisy environment

V. Conclusion

In the field of self-driving cars or autonomous driving, the proposed work serves as a module for navigation system. The use of Faster R-CNN Inception-V2 model via transfer learning improves the accuracy which makes the system reliable for real time application. The outcomes with bounding boxes provide guidelines for real time control actions of the vehicle.

The dataset created for the system covers various use-cases according to the Indian Signal System. So, the work done in this paper paves the way for realization of self-driving cars on Indian streets.

In advancement to this, the system can also be optimized for safe driving despite unclear lane markings. Also, it can be equipped with the ability to respond to spoken commands or hand signals from law enforcement or highway safety employees.

Acknowledgment

The authors would like to thank Prof. Dr. Prasanna Shete for his constant support and encouragement throughout the entire duration.

References

[1] Ross Girshick, Jeff Donahue, Trevor Darrell and Jitendra Malik, "Region-based Convolutional Networks for Accurate Object Detection and Segmentation" in IEEE Transactions on Pattern Analysis and Machine Intelligence, 10.1109/TPAMI.2015.2437384.

[2] J. Deng, A. Berg, S. Satheesh, H. Su, A. Khosla, and L. FeiFei, "ImageNet Large Scale Visual Recognition Competition 2012 (ILSVRC2012)," http://www.image-net.org/challenges/LSVRC/2012.

[3] Mariusz Bojarski, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Prasoon Goyal, Lawrence D. Jackel, Mathew Monfort, Urs Muller, Jiakai Zhang, Xin Zhang, Jake Zhao, Karol Zieba, " End to End Learning for Self-Driving Cars", NVIDIA Corporation, Holmdel, NJ 07735.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[5] V. John, K. Yoneda, Z. Liu, Senior Member, IEEE, and S. Mita, "Saliency Map Generation by the Convolutional Neural Network for Real-time Traffic Light Detection using Template Matching" in IEEE Transactions on Computational Imaging DOI 10.1109/TCI.2015.2480006.

[6] Guo Mu, Zhang Xinyu, Li Deyi, Zhang Tianlei, An Lifeng, "Traffic light detection and recognition for autonomous vehicles" in The

Journal of China Universities of Posts and Telecommunications, DOI: 10.1016/S1005-8885(15)60624-0.

[7] Zhenwei Shi, Zhengxia Zou, and Changshui Zhang, "Real-Time Traffic Light Detection with Adaptive Background Suppression Filter" in IEEE Transactions on Intelligent Transport Systems, DOI 10.1109/TITS.2015.2481459.

[8] Seokwoo Jung, Unghui Lee, Jiwon Jung, David Hyunchul Shim, "Real-time Traffic Sign Recognition System with Deep Convolutional Neural Network" in 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAl) 2016, 978-1-5090-0821-6/16/$31.00.

[9] Keyu Lu, Jian Li, Xiangjing An, Hangen He and Xiping Hu, "Generalized Haar Filter based CNN for Object Detection in Traffic Scenes" in 13th IEEE Conference on Automation Science and Engineering (CASE) 2017, 978-1-5090-6781-7/17/$31.00.

[10] Gwang-Gook. LEE, Byung Kwan PARK, "Traffic Light Recognition Using Deep Neural Networks" in IEEE International Conference on Consumer Electronics (ICCE) 2017, 978-1-5090-5544-9/17/$31.00.

[11] Karsten Behrendt, Libor Novak, Rami Botros, "A Deep Learning Approach to Traffic Lights: Detection, Tracking, and Classification" in IEEE International Conference on Robotics and Automation (ICRA) Singapore 2017, 978-1-5090-4633-1/17/$31.00

[12] Masamitsu Tsuchiya, Yuji Yamauchi, Hironobu Fujiyoshi, Takayoshi Yamashita, "Hybrid Transfer Learning for Efficient Learning in Object Detection" in Second IAPR Asian Conference on Pattern Recognition 2013, DOI 10.1109/ACPR.2013.8

[13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", arXiv:1506.01497v3 [cs.CV] 6 Jan 2016

[14] Swathy S Pillai, Radhakrishnan B, Dr.L. Padma Suresh, "Detecting Tail Lights for analyzing traffic during night using Image Processing Techniques" in International Conference on Emerging Technological Trends [ICETT] 2016.

[15] Dwi H. Widyantoro & Kevin I. Saputra, "Traffic Lights Detection and Recognition based on Color Segmentation and Circle Hough Transform" in International Conference on Data and Software Engineering 2015.