IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 26, Issue 6, Ser. 1 (Nov. – Dec. 2024), PP 23-28
www.iosrjournals.org

Road Sign Detection Using R-CNN

Ranit Pal
Department Of Cse,
Chandigarh University,
Punjab, India

Arvind
Department Of Cse,
Chandigarh University,
Punjab, India

Rahul Bhandari
Department Of Cse,
Chandigarh University,
Punjab, India

Varun Kumar Choudhary
Department Of Cse,
Chandigarh University
Punjab, India

Jaydeep Kumar
Department Of Cse,
Chandigarh University,
Punjab, India

Veer Bajpai
Department Of Cse,
Chandigarh University,
Punjab, India

Abstract
The progression of automatous automobile navigation systems relies on the study of Traffic Sign Recognition (TSR). Because of the need to constantly analyze the situation to make decisions, all autonomous driving systems lack the capacity to run complex image processing algorithms as well as large-scale computing responsibilities. However, they need the efficient calculations of input data in order to generate outputs in the shortest time possible. This is due the fact that there are a lot of traffic signs, advertisements, parked cars, individuals and many other vehicles and objects both in motion and stationery in urban environments. While the strategies discussed in literature for the implementation of ITS have been proposed in varied forms, few have been modelled and analyzed at most at moderate speeds, or with regards to rural or urban roads. Hence, the aim of this paper is to give a background of the main problem and the available techniques to solve those issues, and present a detailed approach to tackle real-time recognition problems in urban traffic sign detection.
Keywords: *Road Sign Detection, Machine Learning, R-CNN, Image Classification, Deep Learning.*

Date of Submission: 06-11-2024 Date of Acceptance: 16-11-2024

I. Introduction

Modern applications like self-driving cars encompass features of software well understood as Automatic Traffic Sign Detection and Recognition (ATSDR). Over the last couple of years, many approaches to the automatic detection as well as recognition of traffic signs and signals have been proposed, early of which Kamada et al. (1990). From a survey by Zhu et al (2016a & 2016b), it was identified that various research works have indicated encouraging results in reconstructing the TSDR with the help of computer vision technology for develop autonomous driving system. However, a significant part of the studies has been limited to a small set of categories associated with Automotive Driving Assistance Systems (ADAS), according to Timofte et al. (2011). As a result, a significant number of TSDR benchmarks focus on particular segments of traffic signs and are primarily oriented towards recognition tasks (Zaklouta & Stanciulescu, 2012; Zhu et al., 2016a, 2016b). The large differences in the shapes and sizes of traffic signs, especially those not covered in these standards, present difficulties in detection of these signs.

A vast majority of TSDR studies have been done outside the Indian context and only a few studies have taken India's social context into account. Some public benchmark datasets exist, for example, the Tsinghua-Tencent 100K dataset (Zhu et al., 2016a, 2016b), GTSRB (Stallkamp et al., 2012), GTSDB (Houben et al., 2013), and LISA (Mogelmose et al., 2012), but they essentially originate from foreign traffic conditions. Different traffic sign models implement deep neural networks especially Convolutional Neural Networks (CNNs) during traffic sign identification and recognition. Even though researchers have presented the algorithms such as Mask R-CNN (Aziz et al., 2020) and Fast R-CNN they have been tested in foreign nations and its applicability in the Indian roads cannot be confirmed.

Therefore, this paper puts forward an advanced model, known as Refined Mask R-CNN (RMR-CNN), to detect and recognize traffic signs that are relevant to Indian highways primarily. The improvement of the original Mask R-CNN, RMR-CNN has certain additional steps that include shape detection, select ROI, color probability and some parameterization to improve the detection precision. All these enhancements mean a higher accuracy level, whereby more traffic sign identification is bound to be accurate.

Based on the identified objectives, the primary objective of this study is to; for measures of minimizing road accidents in India. It is imperative to pinpoint that vehicle accidents do not only cause of deaths but also widely affects families. Traffic sign recognition which is an important aspect in ADAS can greatly help to reduce automobile accidents. However, the area of interest based on traffic signs in the Indian field is still relatively uncovered, and no attempts have been made to evaluate contemporary Deep learning algorithms like Mask R-CNN or Fast R-CNN for traffic sign recognition in the Indian condition.

Using a custom dataset and a deep learning approach, this study aimed to:
- Identify them and explain them.
- Assess the reliability of traffic signs detection.

Though, there is only relatively few large datasets exist for Indian traffic signs, there are numerous large datasets available for countries other than India. The study area is vast and limitless, and there is no single source or database that has standardized information regarding all types of Indian traffic signs which appears to be a setback. This variability impacts both the detection rate and the speed of TSDR methods, which advocates for reduction of this gap to enhance traffic sign detection and recognition capabilities.

Related work

This section aims to give a brief literature review regarding related researches on traffic sign detection made abroad. Luo et al. (2018) propose a three-step data-driven system for text-based and symbol-based communication that uses a camera mounted on a car to show an example. The three phases included post-processing, ROI extraction, and ROI refinement or classification, to name but a few. The issue that was mainly problematic with their proposed assignment was in fact the additional time spent on post- processing as a whole.

While giving an overview of the limitations and challenges of the TSDR system, a critical component of ADASs, Mammeri and his team found that the system was only effective at a frequency of 2Hz, which would explain why moving cars could not recognize traffic signals because of vibrations and oscillations of the low-resolution camera.

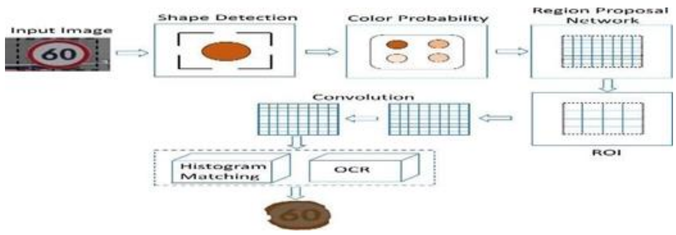
The CNN traffic-sign identification method herein described will be released in parallel by the authors of Lee and Kim (2018). It identifies the traffic signs, determines the location and border of the signs. Although it was accurate, specifying people's locations by city level did not meet Lee's group's requirements for high quality images. Hu et al. (2016) concentrated on three categories of items: motor bikes, automobiles and traffic signals. They also suggested that one learning-based detection should be employed to detect all the three groups based on learning.

The time it took to operate reduced since they chose a traffic sign detector model which was created using fewer sub-detectors relative to other models. However, further characteristics developed just for the detection of other items have made it possible to bring down the run times of the detections. The method explained by Greenhalgh and Mir Mehdi (2012) divided the signal-containing portion of the picture into maximum stable extremal zones, including that conducive for detection of signals under varying climatic conditions. Sign detection uses highly trained SVM classifiers with the HOG features. Meanwhile, concerning the accuracy 86% of proper documents was found, as for remembrance only 80% of important papers was returned. Subsequently, Greenhalgh and Mir Mehdi (2015 LM) used a scene structure in order to determine, where a traffic sign would likely appear in an image.

II. Data Analysis And Results

Traffic sign detection using Mask R-CNN

In the following discussion, we describe the enhancements made to the proposed RMR-CNN system that has been designed for traffic sign detection. First, the Mask R-CNN algorithm applied for traffic sign detection is considered, and the main changes made are described. Second, the model is examined with particular focus on the parameters' tuning and how the network is fine-tuned to achieve better results. Last but not the least, modifications done in the Mask R-CNN model structure together with changes done in the training and testing data are discussed to improve detection abstraction. The refined Mask R-CNN model is shown in module:1.

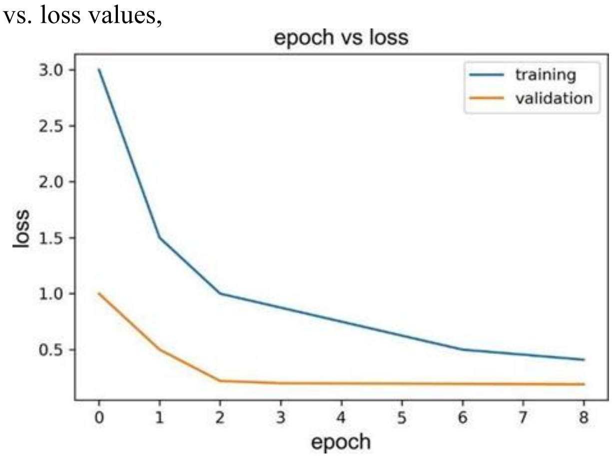


Module: 1 Mask R-CNN is improved by form identification, illumination-based color probability adjustment for TSR and histogram matching, and OCR to enhance the TSD and TSR.

Using Fast R-CNN for Traffic sign detection

The performance of the proposed RMR-CNN is measured by evaluating the detection performance of Fast R-CNN and Masked R-CNN on image recognition. But one of the most significant disadvantages regarding the usage of R- CNN is that the training time is quite extensive, where several seconds for every image are spent on the recognition and detection. This limitation gave rise to another model known as Faster R-CNN as the problem is addressed here as shown below. A CNN version is implemented in Girshick's (2015) Fast R-CNN architecture where instead of the standard pooling layer, there is an ROI pooling layer. The Fast R-CNN model uses the output of the ROI Pooling layer through both its BB-regression and SoftMax branches. For instance, the SoftMax branch, which is part of the current system, can determine how every ROI corresponds to other several categories in addition to the background categories described above. Substituting the data in the Fast R-CNN model, the test accuracy was about 93% when tested on this set of data, it however is not applicable for real world applications.

Module:2 shows the epoch vs. loss values,

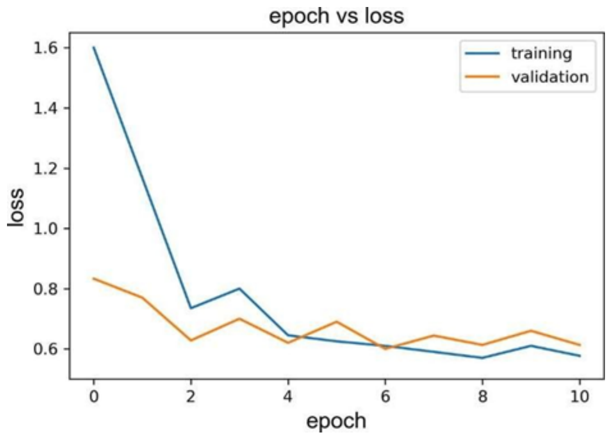


Module 2. This graph shows the epoch vs Loss.

Used Refined Mask R-CNN in Traffic Sign Detection In their work titled "Mask R-CNN", He et al in 2017 presented important enhancement with the title with the second model known as Fast R-CNN. Detection models actually involve two key aspects by and large. The first is the Region Proposal Network (RPN) which, given an image, indicates areas of interest or locations where the object of interest might be in that image in the form of bounding boxes. The second part is called Fast R-CNN, which is a modification of the conventional R-CNN that uses region based convolutional neural networks to then categorize the contents of the proposed boxes. This is particularly true considering that Mask R-CNN uses the Feature Pyramid Network (FPN) as suggested by

Lin et al. (2017), which maintains the useful features at the lower scale. Unlike Fast R-CNN that supports the use of the VGG 16 network, Mask R-CNN optimizes its architecture with the aid of a residual network. While tested on the smaller custom data set in this study, Mask R-CNN reported a test accuracy of 94%. After training and validation, the final log accuracy value was also 0.94 percent.

As shown in the module 3, epoch vs. loss values transactions depict that the validation data lost less than the training dataset lost.



Module. 3. Here shows the graphical form of epoch and loss.

Data Augmentation

One of the key factors for having high accuracy in deep learning models is the amount of data we have in the training dataset. Due to the existence of millions of parameters to be learnt, a big dataset is required for achieving accurate parameter tuning in these models. The tens of thousands of photos can, however, not normally be accumulated real time. This brings us to data augmentation which would help in feeding more photos into the set by manipulating the existing ones. Eight methods of data augmentation were used in this project:

1. Image Flipping: A clockwise or counter clockwise rotation can be created on the picture. The pixel values from adjacent regions are given to any successive blank spaces.
2. Flipping a horizontal axis: For the mirror image along the y-axis of a given picture, the orientation of the picture is reflected backwards.
3. Brightness Adjustment: Originally, image brightness is often reduced, and by subtracting a certain value from the intensity of each pixel, the image's required brightness is achieved.
4. Taking a Closer Look: A 25X enlargement makes the picture look as if it was clipped from the original. This method can therefore be combined with brightness reduction in order to increase the separation of the improved image from the standard.
5. Shift to the right: If some right-side pixels are erased and rest of the pixels are stretched towards right side then the picture is shifted towards right. The neighboring pixel values are placed in the vacant space on the left of the image.
6. Left Shift: Cutting off the pixels from the left side and then mirroring the image over will do the same as right shift. The adjacent pixel values take up all the space left blank at the right side as shown in the first case.
7. Up Shift: And that's how erasing pixels at the top and shifting the rest of the material upwards results in the image being shifted upwards. The following empty area at the bottom is completed by using adjacent pixel values from the top.
8. Bottom Adjustment: Thus, when we subtract pixels from the bottom and move the remaining pixels, the total image shifts downwards. Pixel values of adjacent areas at the lower part complete the blank area located at the top.

III. Literature Review

The paper "Traffic Sign Recognition Based on a Deep Learning Algorithm" presents an advanced approach to improving traffic sign detection, which is critical for intelligent transportation systems. The study emphasizes the importance of detecting traffic signs for vehicle decision-making and control, especially when dealing with small and inconspicuous signs, which often lead to low detection accuracy.

Limitations of Traditional Methods: The paper discusses the shortcomings of previous approaches, particularly their difficulty in detecting small-scale objects. These models often rely on Mean Squared Error (MSE) as a loss function, which proves ineffective for small target recognition.

Enhancements in the Algorithm: The proposed method integrates a spatial pyramid pooling structure, allowing for a more efficient fusion of local and global features. This is particularly valuable for recognizing smaller signs. An additional feature prediction scale (152 × 152) is introduced to better detect shallow features, significantly improving the detection of small objects.

Loss Function Optimization: Instead of relying on MSE for bounding box regression, the paper utilizes the distance- IoU (DIoU) loss function. This method accounts for the distance between the object and anchor, overlap rate, and scale, leading to more stable bounding box predictions.

Dataset and Augmentation Techniques: The authors make use of the Tsinghua–Tencent 100K (TT100K) dataset, recalculating anchors with K-means clustering to manage class imbalance. Additionally, they apply data augmentation techniques to balance the dataset and ensure a fair representation of traffic sign categories.

Use of RCNN: Along with the deep learning algorithm, the paper explores the use of the RCNN model to enhance detection capabilities. The combination of both methods allows for more robust detection of various traffic sign types, particularly small targets.

Performance Evaluation: The improved model demonstrates a significant 8.4% increase in mean average precision (map), with a 10.5% improvement in small target detection. This highlights the algorithm's effectiveness in addressing challenges such as small object detection and dataset imbalance.

IV. Conclusion

This study offered a realistic approach to real-time localization of Indian RTSSs with deep learning in different locations, orientation, position, and size. This work develops RMR-CNN from the Mask R-CNN that is the Mask R-CNN is enhanced model. It offers better parametrical values, more enriched data, and better designed architectural plan. In this case, a novel and kind specific data set was collected in real time in order to effectively train and evaluate the provided RMR-CNN model. Aside from DA, several changes to the typical CNN model corroborated the feasibility and effectiveness of the proposed RMR-CNN framework for training adequate numbers of Indian traffic signs while maintaining the classification's precision, speed, execution time, and learning ability. This was also supported by a significant decrease on the miss rate and false positive rate as reported in the study. These recently created CNN kinds were tested using different types of traffic signals including real ones that were used and included in the dataset. The 3% error rate mentioned above is the result of three different mistake types: broad viewing angle, occlusion, and resemblance to rest traffic signs. With the use of a stereo camera used to capture a scene with the same traffic sign from multiple angles, the possibility of making a mistake is minimized. In this regard, the authors believe that there is a need to employ the RMR-CNN technique to further enhance performance, reduce the miss rate and false positive rate, and perform sign recognition and detection of the obscured, messy and foggy traffic signs as suggested in Majid & Heaslip (2016).

References

[1] Haque, W. A., Arefin, S., Shihavuddin, A. S. M., & Hasan, M. A. (2021). Deepthin: A Novel Lightweight Cnn Architecture For Traffic Sign Recognition Without Gpu Requirements. Expert Systems With Applications, 168, 114481.

[2] Dewi, C., Chen, R. C., Jiang, X., & Yu, H. (2022). Deep Convolutional Neural Network For Enhancing Traffic Sign Recognition Developed On Yolo V4. Multimedia Tools And Applications, 81(26), 37821- 37845.

[3] Hijji, M., Iqbal, R., Pandey, A. K., Doctor, F., Karyotis, C., Rajeh, W., ... & Aradah, F. (2023). 6g Connected Vehicle Framework To Support Intelligent Road Maintenance Using Deep Learning Data Fusion. Ieee Transactions On Intelligent Transportation Systems.

[4] Ahmed, S., Kamal, U., & Hasan, M. K. (2021). Dfr-Tsd: A Deep Learning Based Framework For Robust Traffic Sign Detection Under Challenging Weather Conditions. Ieee Transactions On Intelligent Transportation Systems, 23(6), 5150-5162.

[5] Mall, P. K., Narayan, V., Pramanik, S., Srivastava, S., Faiz, M., Sriramulu, S., & Kumar, M. N. (2023). Fuzzynet-Based Modelling Smart Traffic System In Smart Cities Using Deep Learning Models. In Handbook Of Research On Data-Driven Mathematical Modeling In Smart Cities (Pp. 76-95). Igi Global.

[6] Tursynova, A., Omarov, B., Tukenova, N., Salgozha, I., Khaaval, O., Ramazanov, R., & Ospanov, B. (2023). Deep Learning-Enabled Brain Stroke Classification On Computed Tomography Images. Computers, Materials & Continua, 75(1), 1431-1446.

[7] Bi, Z., Yu, L., Gao, H., Zhou, P., & Yao, H. (2021). Improved Vgg Model-Based Efficient Traffic Sign Recognition For Safe Driving In 5g Scenarios. International Journal Of Machine Learning And Cybernetics, 12, 3069-3080.

[8] Tursynova, A., Omarov, B., Sakhipov, A., & Tukenova, N. (2022). Brain Stroke Lesion Segmentation Using Computed Tomography Images Based On Modified U-Net Model With Resnet Blocks. International Journal Of Online & Biomedical Engineering, 18(13).

[9] Zhao, Q., Yang, L., & Lyu, N. (2024). A Driver Stress Detection Model Via Data Augmentation Based On Deep Convolutional Recurrent Neural Network. Expert Systems With Applications, 238, 122056.

[10] Kumar, G. K., Bangare, M. L., Bangare, P. M., Kumar, C. R., Raj, R., Arias-Gonzáles, J. L., ... & Mia, M. S. (2024). Internet Of Things Sensors And Support Vector Machine Integrated Intelligent Irrigation System For Agriculture Industry. Discover Sustainability, 5(1), 6.

[11] Kendzhaeva, B., Omarov, B., Abdiyeva, G., Anarbayev, A., Dauletbek, Y., & Omarov, B. (2021). Providing Safety For Citizens And Tourists In Cities: A System For Detecting Anomalous Sounds. In Advanced Informatics For Computing Research: 4th International Conference, Icaicr 2020, Gurugram, India, December 26–27, 2020, Revised Selected Papers, Part I 4 (Pp. 264-273). Springer Singapore.

[12] Jayapal, P. K., Muvva, V. R., & Desanamukula, V. S. (2023). Stacked Extreme Learning Machine With Horse Herd Optimization: A Methodology For Traffic Sign Recognition In Advanced Driver Assistance Systems. Mechatronics And Intelligent Transportation Systems, 2(3), 131-145.

[13] C. Niu, Y. Li And H. Li, "Strawberry Pest Identification Based On Image Gray Histogram", Jiangsu Agricult. Sci., Vol. 45, Pp. 169-171, 2017.

[14] S. Zhang, S. Zhang, C. Zhang, X. Wang And Y. Shi, "Cucumber Leaf Disease Identification With Global Pooling Dilated Convolutional Neural Network", Comput. Electron. Agricult., Vol. 162, Pp. 422-430, Jul. 2019.

[15] Y. Guo, P. Su And Y. Wu, "Robot Target Detection And Spatial Localization Based On Faster R-Cnn", J. Huazhong Univ. Sci. Technol. (Natural Sci. Ed.), Vol. 46, No. 12, Pp. 55-59, 2018.

[16] B. S. Ghyar And G. K. Birajdar, "Computer Vision Based Approach To Detect Rice Leaf Diseases Using Texture And Color Descriptors", Proc. Int. Conf. Inventive Comput. Inform. (Icici), Pp. 1074-1078, Nov. 2017.