International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.021 ⵣ Peer-reviewed & Refereed journal ⵣ Vol. 13, Issue 3, March 2025
DOI: 10.17148/IJIREEICE.2025.13327

Indian traffic sign detection and recognition using deep learning

Mr. M. Lokesh[1], Mrs. P. Menaka, M.C.A., M.Phil., (Ph.D.),[2]

Department of Information Technology, Dr. N. G. P. Arts and Science College,
Coimbatore, Tamil Nadu, India[1]

Assistant Professor (SG), Department of Information Technology, Dr. N. G. P. Arts and Science College,
Coimbatore, Tamil Nadu, India[2]

Abstract: Traffic signs are essential in regulating traffic on the road, guiding drivers, thus helping to avoid injuries, property damage, and deaths. Managing traffic signs with automatic detection and recognition is a significant component of any Intelligent Transportation System (ITS). In the age of self-driving vehicles, the importance of automatic detection and recognition of traffic signs cannot be emphasized enough. This paper introduces a deep-learning-driven autonomous approach for the identification of traffic signs in India. The automatic detection and recognition of traffic signs were designed using a Convolutional Neural Network (CNN)- Refined Mask R-CNN (RM R-CNN)-based end-to-end learning framework. The proposed concept was evaluated using an innovative dataset featuring 6480 images that included 7056 instances of Indian traffic signs categorized into 87 classes. We provide multiple enhancements to the Mask R-CNN model in terms of both architecture and data augmentation. We introduce multiple improvements to the Mask R-CNN model, both in terms of architecture and data augmentation. We have examined particularly difficult Indian traffic sign categories that have not been documented in earlier research. The dataset for training and testing our proposed model is gathered by taking images in real-time on Indian roads. The evaluation findings show an error rate of less than 3%. Additionally, the performance of RM R-CNN was contrasted with traditional deep neural network architectures like Fast R-CNN and Mask R-CNN. Our proposed model attained a precision of 97. 08%, which surpasses the precision achieved by the Mask R-CNN and Faster R-CNN models.

Keywords: Traffic Signs, Intelligent Transportation System (ITS), Refined Mask R-CNN (RM R-CNN), Indian Traffic Signs Dataset, Deep Learning, Real-time Image Processing, Data Augmentation, Precision Rate.

I. INTRODUCTION

This research presents an improved version of Mask R-CNN (RMR-CNN) aimed at traffic sign detection and recognition specifically designed for Indian roads. Although considerable research has been carried out worldwide on traffic sign detection (TSD) and recognition (TSR), the majority of datasets and models derive from foreign roads, restricting their use in India. The absence of a substantial, standardized dataset for Indian traffic signs has posed a challenge in creating precise systems for Indian road conditions. This study fills that void by developing a tailored dataset of manually collected Indian traffic signs, classified according to types and real-world variations such as scaling, orientation, and lighting.

The RMR-CNN model improves upon the standard Mask R-CNN by incorporating various pre-processing techniques, including shape detection, region of interest (ROI) selection, and adjustments to colour probabilities. These enhancements are designed to boost accuracy in detecting and identifying traffic signs across the varied conditions typical of Indian roads. The model was evaluated against other deep learning models, such as Fast R-CNN and Mask R-CNN, for both traffic sign detection and recognition tasks.

II. INDIAN ROAD TRAFFIC SIGNAL CATEGORIES

Road traffic signals hold significant importance for the safety of vehicles, drivers, passengers, and pedestrians, as well as limiting property damage. Road signs in India are sorted into three categories:

a) Mandatory or Regulatory signs
b) Cautionary, Precautionary, or Warning signs
c) Informatory signs

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⵣ Peer-reviewed & Refereed journal ⵣ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13327

Mandatory signs are those that users have no alternative but to comply with. Ignoring these signs is regarded as illegal and a punishable offense. Cautionary or Warning signs notify vehicle users of potential slowdowns, roadwork, hazards, or accidents ahead. Informatory signs are highly beneficial for road users, indicating whether they are heading in the correct direction, traveling on the right route towards a desired destination, etc. They also provide information about roadside services for travelers—food/restaurants, lodging, rest areas, toilet gas/diesel stations, etc.The customized dataset created for this paper's research contains a total of 6480 images, with 7056 instances of traffic signs, organized into 100 categories based on the traffic sign instances depicted in the images. Out of the 100 categories, 13 traffic sign categories were sourced from standard datasets such as Deutsche Forschungs Gemeinschaft (DFG), where the traffic signs match those of Indian traffic signs. A total of 4544 images, from the overall 6480, were sourced from real-world conditions, classified into 87 traffic sign categories; the remaining 1936 images were obtained from the Jharkhand (a state in northern India) Police Department, India Mart, and slide share websites. These three are public websites; only images not restricted by copyright were utilized. Each category comprises a minimum of 32 images containing traffic signs. Approximately 70. 12% of the total images are high resolution (4128 2322px, 774 1032px, 960 1280px), while the remaining 29. 8% of images are low resolution (225 225px, 200 200px). There are 200 images where 2 instances of traffic signs are located, 40 images with 3 traffic sign instances, and 32 images where 4 traffic sign instances are present. Roughly 97% of the traffic sign instances are distinctly visible, leaving only 3% with poor visibility. Sample images from the custom dataset are displayed in Figs. 1–2.

Images in the customized dataset are classified into 2 categories—one with signs exceeding 30 pixels resolution, and the other with signs surpassing 50 pixels resolution. There are a few instances where the bounding box is less than 30 pixels, which occurred in images with multiple instances of the same traffic sign; these instances are disregarded during both training and testing phases. A significant majority of all traffic sign instances exceed 30 pixels. The rationale for excluding instances with fewer than 30 pixels is due to downsampling in faster R-CNN and Masked R-CNN, where the 32 32 pixels are represented using 1 1 pixel in the feature map. The standard test split ratio of 80:20 was applied in this RMR-CNN model, ensuring that at least 6 images from each category are used in the custom dataset, with 32 being the minimum quantity of images in each category. The division of the dataset into training and test datasets was conducted randomly. This approach resulted in 5664 images for the training dataset and 1412 images for the test dataset.

Redefined Mask-RCNN for traffic sign detection

This section contains a conversation about the suggested RMR-CNN system, aimed at detecting traffic signs, with various enhancements. Initially, the Mask R-CNN algorithm utilized for traffic sign detection is briefly outlined; subsequently, modifications in the parameter values to customize the Mask R-CNN to our specifications are illustrated, followed by advancements in architecture and data augmentation of the Mask R-CNN. fig. 3.

Pre-processing for better accuracy

Three preliminary processing stages are executed before implementing the Mask R-CNN algorithm for sign detection and recognition: shape detection, ROI, and colour probability.

Shape Detection: The initial preprocessing step aims to identify shapes within an image captured from live video. After locating the shape, both coloured and grayscale images are obtained. Initially, the image data from the camera is transformed from colour to grayscale. Subsequently, an OpenCV counter detection technique is utilized to identify the counter values. Using the captured counter values, the area is computed. Based on the area of the counters, the form of the traffic sign is identified according to the threshold parameters set by the user. These images are subsequently forwarded to the ROI module.

Region Of Interest (ROI): The region of interest plays a crucial role in pinpointing the precise location of the traffic sign. Most traffic signs in India are displayed in the shapes of triangles, rectangles, and circles; thus, the proposed RMR-CNN model is able to identify these shapes as ROI in the image. An image cannot be accurately assessed by the recognition model if it contains undesirable areas. This could also result in erroneous predictions, or it may require additional time to predict a larger area. Three algorithms are employed to determine ROI within an image. If the ROI is circular, the Hough Circle package is applied; for triangle and square detection, Counter and Edge detection algorithms are utilized. A distinct code, implemented to find ROI for each shape, highlights the region of the circle, rectangle, or triangle shapes with a specific colour. The image is cropped by applying a threshold of 5%, to extract the image from the ROI without any loss. The cropped images are transformed into colour images, preceding the next step, colour probability [subsection 'D']. Colour probability is utilized to determine the percentage of RGB pixel values present in the image.

Colour Probability: In this phase, the Red Green Blue (RGB) value for each pixel in the sample image is computed and saved in a new dataset. Various lighting situations such as direct sunlight hitting a traffic sign, sunlight illuminating

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.021 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 3, March 2025
DOI: 10.17148/IJIREEICE.2025.13327

the rear of the traffic sign, and conditions of dim light, are examined to determine the spectrum of RGB pixel values. Using the acquired dataset, the count of pixels that are red, black, and white is calculated through real-time assessment of range from the image dataset. For each pixel value in the image, the counter is incremented for red, black, and white colours, according to the range in which it resides. We calculated the spectrum of RGB pixel counts from the dataset while the image was positioned towards the sunlight and while it was positioned away from the sunlight. Following the colour probability assessment, the image will be forwarded to the model. A threshold value was established from the training dataset for the image to be transmitted to the model.

Fig. 1. The initial collage of the custom dataset. There are 54 pictures, several of which are close-ups, showing different Indian traffic signs in different orientations, placements, and lighting conditions.

Fig. 2. The second collage of the custom dataset. This collage displays 49 pictures of different Indian traffic signs at different distances, in different orientations, placements, and light conditions.

ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13327

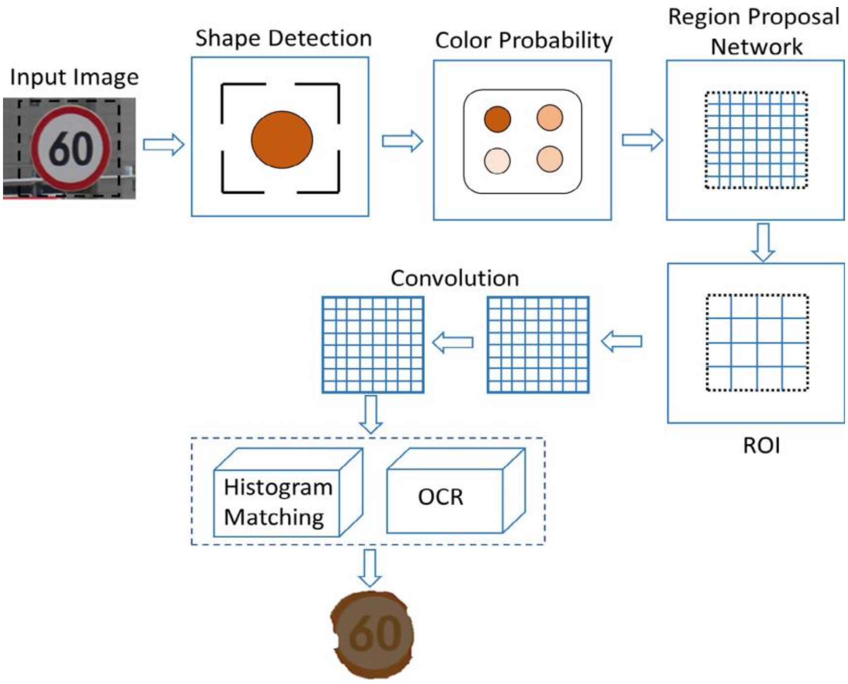

Fig. 3. R-CNN model with refined mask. For TSD and TSR, Mask R-CNN is suggested, which combines form detection, colour probability depending on lighting conditions, histogram matching, and optical character recognition (OCR).

III. DATA AUGMENTATION

The key element for achieving high accuracy is the volume of the training dataset. Since deep learning models contain millions of parameters that can be learned, it is essential to have a vast dataset that assists in effectively learning these parameters. It is not feasible to acquire tens of thousands of required images in real-time. Data Augmentation offers a means to enhance the size of a dataset through variations in certain parameters. A total of 8 data augmentation techniques were employed, which are listed as follows:

Image Rotation: the image is rotated in either a clockwise or anticlockwise direction. The resulting blank pixels are filled with values from adjacent pixels.

Horizontal Flipping: The image is mirrored along its y-axis.

Brightness Change: The brightness of the image is adjusted, typically reduced, by subtracting a constant value from all pixel values.

Zoom In: The image is magnified by an additional 25%, giving it the appearance of a cropped version. This may also result in a decrease in brightness, further differentiating it from the original image.

Right Shift: The image is shifted to the right by removing a segment of pixel values from the right edge and moving the image accordingly; the resulting empty pixels on the left are filled with values from adjacent pixels.

Left Shift: The image is shifted to the left by removing a segment of pixel values from the left side and adjusting the image to the left; the resulting empty pixels on the right are filled with values from adjacent pixels.

Top Shifting: The objects within the image are shifted upwards. This is accomplished by removing a segment of pixel values from the top and relocating the lower pixel values upwards, while the resulting empty pixels at the bottom are filled with values from adjacent upper pixels.

Bottom Shifting: The objects within the image are shifted downwards. This is achieved by removing a segment of pixel values from the bottom and repositioning the upper pixel values downwards, while the ensuing empty pixels at the top are filled with values from adjacent lower pixels.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13327

Traffic sign detection using Mask R-CNN

A detailed account of Mask R-CNN, as an enhanced version of Fast R-CNN, was provided in He et al. (2017). Each of the detector models consists of 2 components that are linked through convolutional connections. RPN, the initial component of the detector model, (Abbas and Singh, 2018) takes the input image and proposes or generates rectangular boxes known as bounding boxes. Bounding boxes are used to likely detect an object. The second component of the detector model is the Fast R-CNN, which is a region-based CNN. Fast R-CNN identifies the objects contained within the bounding boxes. Mask R-CNN employs Feature Pyramid Network (Lin et al. , 2017) that assists in preserving the key features of low resolution. Visual Geometry Group (VGG) 16, a foundational network included in Fast R-CNN, is substituted with a residual network in Mask R-CNN (Simonyan and Zisserman, 2015). The Mask R-CNN model, trained and validated on the custom dataset, achieved an accuracy of 0. 94. Fig. 4 illustrates a graph of the epoch versus loss values, demonstrating the decreased loss of the validation data relative to the loss of the training dataset.

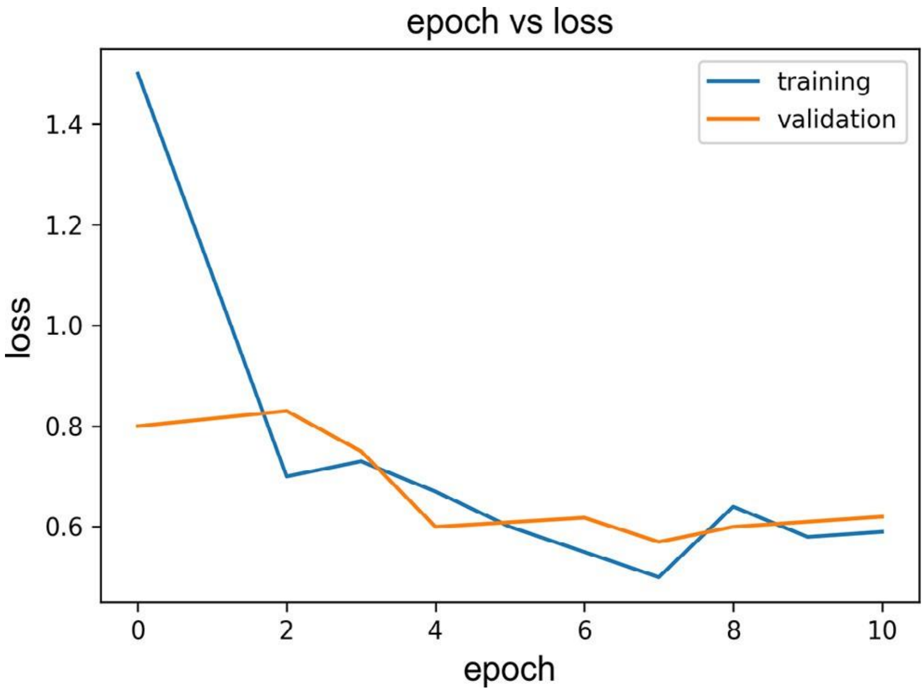

Fig. 4. Epoch vs. loss values for mask R-CNN. The graph's y-axis displays the loss value, while the x-axis displays the epoch value.

Traffic sign recognition

This section presents Optical Character Recognition (OCR)-centric techniques for character recognition and Histogram Matching apropos sign recognition.

Optical character recognition

Classic Mask R-CNN can be utilized for identifying traffic signs that do not have any textual elements. Resume OCR was used on the masked image to extract the enclosed numbers or text characters. This detail is essential, as every circular or triangular traffic sign necessitates accurate identification of the traffic sign (Mainkar et al. 2020). The fundamental model introduced by Mainkar's group was improved [through 'easyocr'] to meet the demands of this study for reduced latency and precise accuracy.

Evaluation

The hyperparameters defined in our model include the number of epochs, batch size, and the minimum and maximum dimensions of the images in the dataset. We have not implemented any optimization techniques but have utilized a physical method where we experiment with various values and evaluate the prediction accuracies on the test dataset. The number of epochs is configured to 20, while the batch size is adjusted to 100 images per batch. The minimum image dimension required is 32 × 32 pixels and the maximum image dimension is 2048 × 1080 pixels. A large batch size will enhance computation speed by leveraging GPU parallelism, but excessively large batch sizes may lead to poor generalization. Utilizing an epoch value that exceeds the necessary amount will lead to overfitting.

DOI: 10.17148/IJIREEICE.2025.13327

IV. RESULTS AND CONCLUSION

This paper put forward a feasible deep learning approach for the identification and recognition of Indian road traffic signs which demonstrated good results under varying conditions such as light changes, orientation changes, and scale changes. The paper presents RMR-CNN, which is an enhanced version of Mask R-CNN and incorporates advancements in architecture, data augmentation, and refinements in parameter values. An innovative custom dataset was acquired in real-time for the training and validation of the proposed RMR-CNN model. In addition to data augmentation, several modifications to the standard CNN model acted as evidence for the RMR-CNN method's precise, efficient, rapid detection, and learning ability of a significant number of Indian traffic signs. The noted performance enhancements were confirmed by significant reductions in the miss rate and false positive rate. The suggested RMR-CNN model excelled beyond both the Fast R-CNN and Mask R-CNN models regarding precision, recall, and F-measure.

The developed CNN variant strategy was effectively validated on a wide range of traffic sign categories. The error rate of almost 3% is primarily attributed to the resemblance with other traffic signs, occlusion, and broad viewing angle. The error rate can be minimized by utilizing multiple instances of the same traffic sign using a stereo camera. The authors propose that the RMR-CNN approach should be further explored for refining the detection and recognition of dirty, unclean traffic signs as indicated in Majid and Heaslip (2016), as well as for improved performance in reducing the miss rate and false positive rate.

REFERENCES

[1]. Singh, D.S.N., and Abbas, S.M. (2018). quicker R-CNN for region-based object recognition and classification. Computational Intelligence & Communication Technology (CICT) 4th International Conference, 2018, pp. 1–6. It can be found at https://doi.org/10.1109/CIACT.2018.8480413.

[2]. Avramovic´, A., Stojnic´, V., Sluga, D., Tabernik, D., and Skoċaj, D., Ilc, N., 2020. Region focussing and parallelisation are used in neural-network-based traffic sign identification and recognition in high-definition photos. Access 8, 189855–189868, IEEE. 10.1109/ACCESS.2020.3031191. https://doi.org.

[3]. Sheikh, U.U., Ayub, S., Haji Salam, M.S.B., and Aziziz, L. (2020). A thorough analysis of the architecture, tactics, applications, and current trends in generic object detection based on deep learning. Access 8, 170461–170495, IEEE. 10.1109/ACCESS.2020.3021508 (https://doi.org).

[4]. Lu, S., and Chen, T. (2016). Using support vector regression and discriminative AdaBoost, traffic signs may be detected accurately and efficiently. 4006–4015 in IEEE Trans. Veh. Technol. 65 (6). 10.1109/TVT.2015.2500275, https://doi.org.Girshick, R., (2015). Fast R-CNN. IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440-1448, doi: 10.1109/ ICCV.2015.169.

[5]. Greenhalgh, J., Mirmehdi, M., 2012. Real-time detection and recognition of road traffic signs. IEEE Trans. Intell. Transp. Syst. 13 (4), 1498–1506. https://doi. org/10.1109/TITS.2012.2208909.

[6]. Greenhalgh, J., Mirmehdi, M., 2015. Recognizing text-based traffic signs. IEEE Trans. Intell. Transp. Syst. 16 (3), 1360–1369. https://doi.org/10.1109/ TITS.2014.2363167.

[7]. Guo, J., You, R., Huang, L., 2020. Mixed vertical-and-horizontal-text traffic sign detection and recognition for street-level scene. IEEE Access 8, 69413–69425. https://doi.org/10.1109/ACCESS.2020.2986500.

[8]. He, K., Gkioxari, G., Dollár, P., Girshick, R., (2017). Mask R-CNN. In: 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp. 2980- 2988, doi: 10.1109/ICCV.2017.322.