IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

A Review on Traffic Sign Recognition Focused on Deep Learning Models

Rishi Mohan Awasthi, M.Tech Scholar, Department of Computer Science & Engineering, Rameshwaram Institute of Technology & Management, Lucknow, India.
Shyam Dwivedi, Assistant Professor, Department of Computer Science & Engineering, Rameshwaram Institute of Technology & Management, Lucknow, India.

Abstract—**In the area of object detection in road situations, significant advancements have been made in recent years. Nevertheless, it focuses mostly on vehicles and people as its targets. In order to achieve this goal, we conduct research into the detection of traffic cones, an important object category for the effects of roads and their upkeep. In order to develop a solution for the efficient and quick identification of traffic cones, the YOLOv5 algorithm is used in this work. The goal of this work is to discover a solution. The YOLOv5 is capable of accomplishing a high level of detection accuracy.**
We will be able to construct a data-driven traffic sign recognition and detection system with a high detection accuracy with the help of this data set and a YOLOv5 that is flawless. This system will also have a high performance ability in the training and recognition procedures. This not only reduces the likelihood of accidents happening, but also makes it easier for drivers to focus on the road ahead of them rather than trying to keep track of every single traffic sign. This article aims to provide an effective strategy for the detection and recognition of various traffic signs used in India.

***Index Terms*—**Traffic Sign Recognition, Intelligent Transport System, Road Accidents, Deep Learning, Object Detection, etc.

I. Introduction

One of the crucial aspects of the Intelligent Transport System (ITS), which prioritizes both the protection of passengers and the optimization of traffic flow, is the Traffic Sign Recognition System (TSR). When travelling on congested roads, the detection of a traffic sign can alert a motorist to pay attention to the current speed limit and any other road signs that are displayed on the screen of the console.

The fact that road signs come in a wide variety of shapes and colours makes it so that drivers can quickly recognize all of the traffic signs without any difficulty. This makes the situation appear to be quite simple. However, the legitimacy of the detected and recognized sign by electronic devices might vary for a variety of reasons, including the shape, dimensions, height, and orientation of these road signs. There are times when they are partially concealed, which can lead to a false detection being made. In India, there are three distinct categories of traffic signs. These road signs include those that are required or Regulatory, Cautionary or Warning, and Informative. Table 1 provides an overview of the numerous indications.

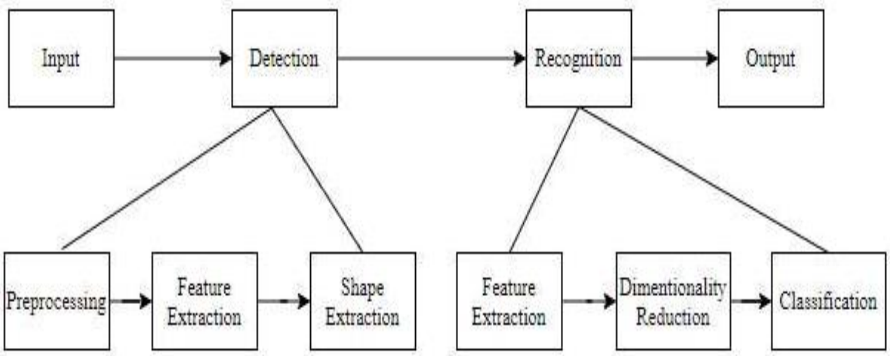
As can be seen from Table 1, the road signs have been developed with specialized hues and contours that are entirely distinct from the surrounding natural environment. This is done to ensure that drivers are able to spot and identify the signs quickly and easily. They are constructed using shapes that are just two-dimensional, such as circles, rectangles, and triangles. According to the research of P. Mukhopadhyay et al. [1], the colour red was chosen for them because it is the colour that is most easily distributed by air molecules. In addition to this, it has the ability to see through cloud cover and precipitation and may be detected from great distances.

Nonetheless, there are a few issues that arise while processing a picture of a real-time traffic sign. A low camera resolution, an imperfect sign state, inadequate illumination, occlusion, and adverse weather circumstances such as precipitation, whether it be rain, snow, or fog, are some of the primary challenges that must be overcome. The colour analysis and form extraction of signs are both susceptible to being influenced by all of these constraints [2]. As a result, these problems need to be appropriately handled if they are going to be adequately detected and recognized when it comes to traffic signs. The TSR system is built with a sequential methodology that consists of three stages (i) the preprocessing stage, (ii) the detection stage, and (iii) the recognition stage. Figure 1 provides an illustration of the numerous steps involved in the process of traffic sign recognition.

Table 1: Indian Traffic Sign Categories

Type of Sign	Sign	Type of Sign	Sign
Mandatory	 Right Turn Prohibited  Speed Limit  No Entry  No honking	**Warning**	 Left Hand Curve
			 Road Narrow from Both Sides
		Informative	 Parking
			 Hospital Ahead

Figure 1: General block diagram followed by TSR algorithm



Here, during the stage known as preprocessing, the image quality is improved so that it can be supplied to the stages that come after it in an effective and timely manner. During the detection stage, the image is combed through in search of the various traffic signs; thus, Regions of Interest (ROI) are determined by making use of the colour or form characteristics of the sign. Following that, the recognition stage will do an analysis on these ROIs. After the signs have been successfully detected, the information is subsequently wirelessly transferred from one car to another vehicle in the immediate vicinity.

Color-based image detection and shape-based image detection are the two primary categories that have been established for the image detection algorithms. According to the research conducted by Arturo Escalera et al. [4], the method of colour segmentation is the one that is most frequently utilized for the detection of traffic signs. It is possible to employ a variety of colour spaces in order to locate the colour of interest and differentiate it from the background. Images obtained from a camera, on the other hand, are hypersensitive to the chromaticity of the light that was incident upon them as the colour of the sign becomes less distinct. Thus, various background conditions need to be met in order to conduct an accurate analysis. A huge library of photos of traffic signs is necessary. The shape of the image can also play a role in the detection process. Only when there are frequent shifts in the light illumination can the shape detection function well. Moreover, shape-based detection works on grayscale images, which may help to meet global cost criteria if this is taken into consideration. After converting the real-time collected images to grayscale, the present work makes use of a shape-based detection technique in order to account for the many cost parameters associated with the design and implementation processes. However, the resolution of the image affects the ability to distinguish objects based on their colours or shapes. The recognition of images requires a large image collection since the dataset is comprised of a variety of images that have distinct visual characteristics. Images recorded in real time are typically blurry and noisy due to the nature of the process. In spite of this, numerous suggestions for addressing the problems with the TSR system can be found across the available research.

Given a particular classification, the goal of Traffic Sign Recognition (TSR) is to pinpoint the exact location of traffic signs within digital still images or video frames [5]. In its most basic version, TSR approaches make use of visual information such as the shape and colour of various traffic signs. The conventional TSR algorithms, on the other hand, have a number of shortcomings that become apparent in real-time tests. One of these is that they are easily constrained by the driving conditions, which can include lighting, camera angle, obstruction, driving speed, and so on. Additionally, it is extremely challenging to accomplish multitarget detection, and because recognition is so slow, it is easy to miss visual objects. [6]

With continual progress of computer hardware, the limitation of artificial neural networks has been well relieved, which has led machine learning into a golden moment of development. One category of approaches to machine learning is known as deep learning [7]. While information is being processed, the neuronal structure of the human brain can be modelled using something called a deep neural network. It has the potential to improve the robustness and generalization of the algorithms if this neural network model is used to extract the effective features from the road image rather than the traditional TSR algorithms. This model uses a neural network to model the road, and it is much better than the conventional TSR algorithms [8].

The study achievements in TSR not only avert traffic accidents and safeguard drivers, but also help review traffic signs on roadways efficiently and precisely, which reduce unneeded manpower and resources. In addition, it also provides technical support for unmanned and auxiliary driving. Thus, the research work based on deep learning has significant significance and is helpful to our daily lives.

As can be seen in figure 2, the research field of TSR is broken up into two primary subfields on the basis of the algorithms that are applied [9].

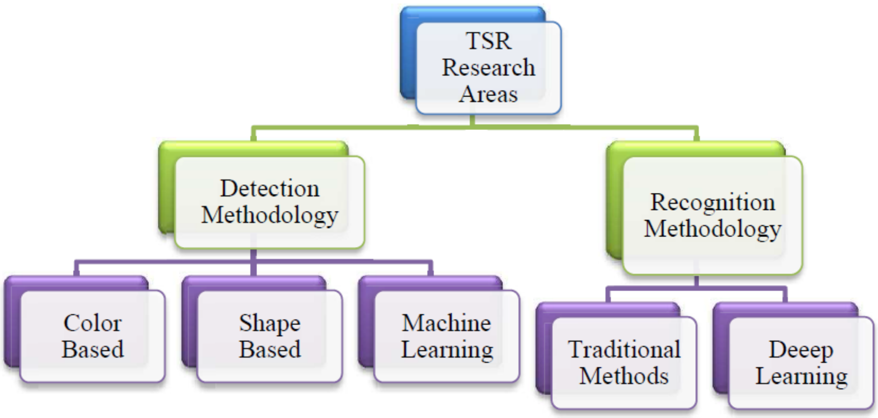


Figure 2: Several TSR Research Areas

The three main categories of detection algorithms are color-based, shape-based, and machine-learning based. The two main categories of recognition algorithms are traditional methods and methods based on deep learning.

II. LITERATURE REVIEW

This section provides a comprehensive assessment of the research done on the various methods of detecting traffic indicators. These methods of detecting traffic signals can be broken down into two distinct categories: detection based on colour, and detection based on shape. The analysis of the different recognition methods is covered in the second section of this section.'

Color Based Detection

The usage of a specific colour space in conjunction with colour segmentation is required for accurate detection while performing colour segmentation. Several studies believe that using color-based approaches rather than shape-based detection is the best way to mitigate the effects of background texture. Nevertheless, the most significant drawback of color-based techniques is that they are sensitive to variations in lighting, which might make the detection less reliable. Katsamenis, I. et al. [10] suggested using the colour clustering approach as a color-based detection tool. However, they came to the conclusion that clustering can also offer challenges, particularly if the colours are either excessively or under clustered.

The HSV colour space was utilized for the colour detection that was carried out by Li, G., Xie, and H. [11]. They came to the conclusion that the colour saturation varies significantly between the videos captured from the various devices. Prabhakar, G., and Kailath [12] conducted research to determine the effect of light on the colour of traffic signs during the day and at night. They came to the conclusion that the colour of roadside images could become distorted due to the light, which could have an impact on the quality of the images. On the other hand, they stressed the fact that the RGB component differences for traffic signs are unaffected by the external illumination. The authors [13] placed a strong emphasis on colour segmentation due to the unreliability of shape identification in urban environments. Voulodimos, et al. [14] have developed an algorithm that uses colour information to determine the presence of potential traffic signs. Researchers carried out a statistical analysis to ascertain the threshold values for the several parameters they were interested in. However, all of the segmentation algorithms that are based on colour thresholding require their thresholds to be adjusted multiple times throughout the process. As a result, colour segmentation must invariably involve the establishment of a threshold.

Shape Based Detection

According to R. Biswas et al. [15], color-based detection would be ineffective throughout the night because of the presence of all the artificial lights. Only when there are frequent shifts in the light's intensity level is shape detection effective. In addition, shape-based detection is effective even when applied to grayscale images, which may make it easier to satisfy global cost criteria. Edges, gradient features, template matching, or the Hough transform are the typical methods used to derive the shape.

The majority of edge detection strategies are dependent on either the contour or edge information. The detection of signs has been discussed by Liu, H., and Liu, Y. [16] who worked on grayscale images and used an algorithm that matched rectangular patterns. The regions in which the indications were found were categorized as belonging to a single scan window. In order to locate a circle within an image, the ellipse equation was utilized as well. The detection of rectangular and circular pictures of 640x480 pixels of traffic signs utilizing rectangular and circular detection method was explained by Timofte, R., et al. [17]. Yet, accurate results from shape detection utilizing edge features can only be obtained during the daytime.

It is possible to extract forms using machine learning algorithms such as support vector machines (SVMs) and neural networks (NNs). Because of their superior capability to recognize forms, SVMs and NNs are frequently utilized in the process of sign identification. Each individual collection of indicators is used to train a neural network. The addition of additional indicators, on the other hand, requires additional training of the network and, as a result, the manual selection of training samples. Although SVMs are resistant to rotation, translation, and the partial obscuration of road signs, training them takes a significant amount of time and requires huge datasets.

III. METHODOLOGY

YOLOv5

In order to recognize traffic cones, the system that is being shown makes use of the roadwork image dataset. Each image was inputted into the YOLOv5 algorithm in the correct manner. YOLO is an abbreviation that stands for "You only look once," and it refers to a method for target detection that is based on a regression approach and makes use of neural networks to give real-time object detection. YOLO was developed by Google. The fact that it finishes the prediction of the categorization and position information of the objects according to the computation of the loss function is the source of its value; as a result, it transforms the target detection problem into a regression problem solution [18]. This method identifies the most cutting-edge detecting technologies that are currently available and then optimizes their implementation such that it conforms to industry standards. For the purpose of our implementation, we make use of YOLOv5, which, of all the YOLO algorithms, has the best performance. It is a suitable lightweight detector that can balance detection accuracy and model complexity under the constraints of processing platforms with limited memory and computation resources. It is based on the PyTorch framework, and its functionality comes from the fact that it is a suitable lightweight detector [19].

There are three components that make up the architecture of the model YOLOv5: The Backbone is (i) CSPDarknet, (ii) PANet is the Neck, and (iii) YOLO Layer is the Head. The data are first fed into CSPDarknet for the purpose of feature extraction, and then they are transferred to PANet for the purpose of feature fusion. In conclusion, the results of object detection are output by the YOLO Layer (i.e., class, score, location, size). Figure 3 provides a visual representation of the model's underlying structure.

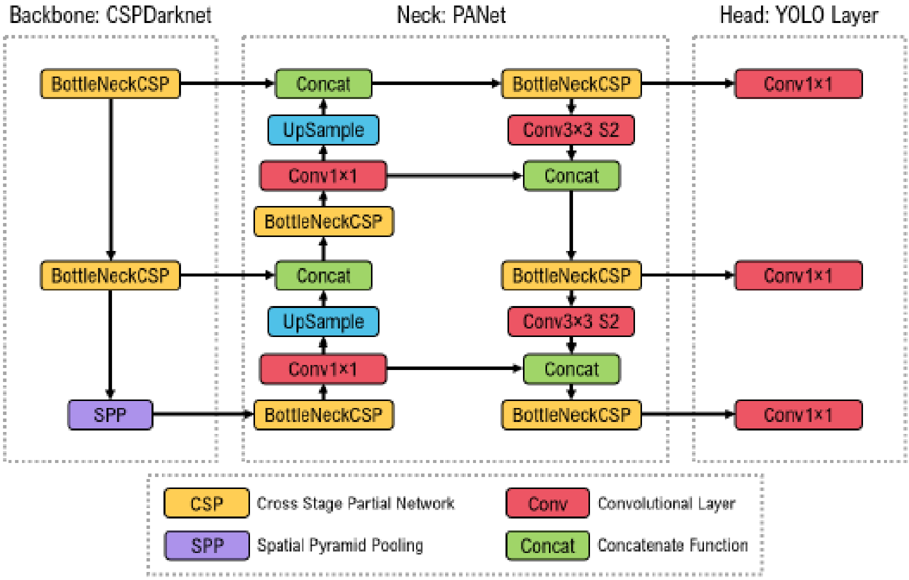


Figure 3: Architecture of YOLOv5 model

IV. Dataset description

One can extract the pertinent highlighted features after the pre-processing stage is complete. The relevant features in a traditional FER system are facial features. Since the accuracy of the system greatly depends on the quality of these characteristics, computer vision researchers have created a number of feature extraction methods. The following feature extraction methods are the most widely used in the reviewed works:

For the purpose of this investigation, we chose eight categories of traffic signs that have a significant impact, both in terms of awareness and protection. Because there were so few traffic signs along the road, we decided to take pictures of the signs rather than recordings of ourselves driving. We have separated them into two groups, and both of these groups are obtained from the public areas of our city by utilising the cameras that are mounted on our mobile devices.

Our database contains total 2,182 images of traffic signs that have the following labels: "No U-turn" (271 images), "Road bump" (329 images), "Road works" (294 images), "Watch for children crossing" (176 images), "Crosswalk ahead" (313 images), "Give way" (317 images), "Stop" (286 images), and "No entry" (286 images) (196 images).

A small number of the images that make up our raw data set were taken with the landscape orientation of the camera. To begin, we utilised a piece of software known as JPEG Autorotate to change the orientation of the photographs to that of a portrait. Following then, because ultrahigh-definition photographs require an excessively lengthy amount of time for training, we downscaled the image while maintaining the same aspect ratio between its width and height. As a result, we adjusted all of the photos in our dataset to have dimensions of 1128 × 2016 and 1536 × 2048.

V. Conclusion

In this paper, we introduced and tested a YOLOv5 algorithm for traffic cone recognition over a multisource roadworks image dataset. The dataset included images of roadworks from multiple sources. The method that is utilized makes use of a deep learning framework, with traffic cones functioning as an example of an object detection scenario. The model was able to do the identification assignment with flying colours and come out on top with excellent marks. The combination of additional ephemeral items that are associated with the upkeep and development of road networks should be part of the work that will be done in the future.

References

[1] P. Mukhopadhyay, J. Kaur, L. Kaur, A. Arvind, M. Kajabaje, J. Mann, V. Vallery, A. Gopal, A. Ish, R. Pochont, and S. Chakravarty,"Ergonomic design analysis of some road signs in India. Information Design Journal (IDJ)", vol. 20, no. 3, pp. 220-227, September. 2013.

[2] M. Shi, H. Wu, and H. Fleyeh, "Support vector machines for traffic signs recognition In IEEE International Joint Conference on Neural Networks, Hong Kong, China, pp. 3820-3827, June 2008.

[3] A. Escalera, L. E. Moreno, M. A. Salichs, and J.M. Armingol, "Road traffic sign detection and classification", IEEE transactions on Industrial Electronics, vol. 44, No. 6, pp. 848-859, December 1997.

[4] D.G. Lowe, " Istinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, no.2, pp. 91 110, Nov 2004.

[5] R. E. Haas, S. Bhattacharjee and D.P. Moller, "Advanced driver assistance systems. In Smart Technologies,", Singapore, Springer, pp. 345-371. 2020

[6] J. Son, H. Yoo, H. Kim and K. Sohn, "Real-time illumination invariant lane detection for lane departure warning system", Experts Systems with Applications, vol. 42, no. 4, pp. 1816-1824, 2015

[7] V. Saini and R. Saini, " driver drowsiness detection system and techniques: a review International Journal of Computer Science and Information Technologies, vol. 5, no.3, pp.4245-4249, 2014

[8] Xu S, Niu D, Tao B, Li G (2018) Convolutional neural network based traffic sign recognition system. In International Conference on Systems and Informatics (ICSAI), pp 957-961.

[9] R. Qian, B. Zhang, Y. Yue, Z. Wang, D. Coenen, "Robust Chinese traffic sign detection and recognition with deep convolutional neural network," IEEE 11th International Conference on Natural Computation, pp. 791-796, January 2016.

[10] Katsamenis, I., Protopapadakis, E., Doulamis, A., Doulamis, N., Voulodimos, A.: Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation. In: International Symposium on Visual Computing. pp. 160–169. Springer (2020)

[11] Li, G., Xie, H., Yan, W., Chang, Y., Qu, X.: Detection of road objects with small appearance in images for autonomous driving in various traffic situations using a deep learning based approach. IEEE Access 8, 211164–211172 (2020)

[12] Prabhakar, G., Kailath, B., Natarajan, S., Kumar, R.: Obstacle detection and classification using deep learning for tracking in high-speed autonomous driving. In: 2017 IEEE region 10 symposium (TENSYMP). pp. 1–6. IEEE (2017)

[13] S. Houben, J. Stallkamp, J. Salmen,M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," in Proc. IEEE IJCNN, 2013, pp. 1–8.

[14] Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E.: Deep learning for computer vision: A brief review. Computational intelligence and neuroscience 2018 (2018)

[15] R. Biswas, H. Fleyeh, M. Mostakim, "Detection and Classification of Speed Limit Traffic Signs," IEEE World Congress on Computer Applications and Information System, pp. 1-6, January 2014. G. Antipov, SA Berrani, JL Dugelay, "Minimalistic CNN-based ensemble model for gender prediction from face images," Elsevier, January 2016.

[16] Liu, H., Liu, Y., & Sun, F. (2014). Traffic sign recognition using group sparse coding. Information Sciences, 266, 75–89. doi:10.1016/j. ins.2014.01.010

[17] Timofte, R., Zimmermann, K., & Van Gool, L. (2011). Multiview traffic sign detection, recognition, and 3D localisation. Machine Vision and Applications (November):1–15. doi:10.1007/s00138-011-0391-3.

[18] Garg P, Chowdhury DR,More VN (2019) Traffic sign recognition and classification using YOLOv2, Faster R-CNN and SSD. International Conference on Computing, Communication and Networking Technologies.

[19] Kuznetsova A, Maleva T, Soloviev V (2020) Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images. International Symposium on Neural Networks.