Automated Traffic Sign Recognition via CNN Deep Learning

K.V.Narasimha Reddy
Dept of CSE,
Narasaraopeta Engineering College,
Narasaraopeta-522601, India
narasimhareddynec03@gmail.com

Yenuganti Narendra
Dept of CSE,
Narasaraopeta Engineering College,
Narasaraopeta-522601, India
yenugantinarendra123@gmail.com

Medam Adi nagamanendra Reddy
Dept of CSE,
Narasaraopeta Engineering College,
Narasaraopeta-522601, India
medamadi111@gmail.com

Avula Ramu
Dept of CSE,
Narasaraopeta Engineering College,
Narasaraopeta-522601, India
avularamu2004@gmail.com

Dodda Venkata Reddy
Dept of CSE,
Narasaraopeta Engineering College,
Narasaraopeta-522601, India
doddavenkatreddy@gmail.com

Sireesha Moturi
Dept of CSE,
Narasaropeta Engineering College,
Narasaraopeta-522601, India
sireeshamoturi@gmail.com

***Abstract*—The rapid development of the road traffic systems is a very important component of the nation's infrastructure, and that is reflected in growing importance of the traffic safety. Traffic Sign Recognition (TSR) is an important field of research because traffic offenses, particularly the disregard for traffic signs, are a major contributor to accidents. This paper gives a comprehensive review of the latest developments in the area of traffic sign detection and recognition techniques with attention to the applications of CNNs. This paper discusses the need for traffic signs to be detected under complex conditions and proposes an architecture based on modified CNN that helps reduce the processing time and increase accuracy. Improving recognition accuracy in realistic environments is the motivation behind the CNN model, which has been architected for real-time training as well as target identification.Experiments suggest that this solution outperforms present intelligent driving systems and the state-of-the-art performance achieved by the image processing algorithms currently in use and traffic sign datasets.**

***Index Terms*—Traffic signs, Convolutional Neural Network, Keras.**

I. INTRODUCTION

The computer vision and machine learning problem of "Traffic Sign Recognition" (TSR) aims to recognize and classification of the traffic signs from image or video streams. In essence, it is meant to develop models and algorithms that are capable of automatically recognizing and interpreting many types of traffic indicators, ranging from yield signs and stop signs to speed restrictions among others[1][2][3]. Due to this, TSR is essential in supporting drivers to follow traffic laws, hence it ought to be kept up to date with contemporary road safety. Given the nature of this technology, human errors on roads could be greatly minimized and incidences of road accidents could be avoided. These technologies also reduce vehicle emissions and fuel consumption by encouraging people to respect the set speed and other provisions of traffic law. In cities, which often suffer from high pollution levels, this is particularly important. Using video feed coming from cameras mounted on vehicles, TSRs are made to recognize traffic signs promptly. For such systems, image processing techniques are used for sign recognition and classification according to form, color, and symbols. The procedure followed for the classification of signs is comparison-based: In this, the system's database forms a predefined set of templates against which the detected signs are compared. Traffic sign recognition technologies are fast becoming a critical necessity in light of ever-burgeoning traffic congestion and accidents. These devices identify and decode traffic signs, hence giving motorists information and guidelines on what actions to undertake [4].Some of the major benefits of Traffic sign recognition include its ability to push forward road safety as it ensures that crashes are prevented and reductions in deaths occur with fast notifications of traffic signs. Alam et al. [5][6] had developed SURF, which in prior work could be referred to as a feature identification and description approach for traffic sign recognition.It is suggested in the paper [7] that the best strategy for traffic sign identification is to integrate grid search, SVM classifier, and HOG features. There are many techniques to develop a traffic sign recognition system, like machine learning algorithms on CNN [8] [9], template matching [10], rule-based systems [11] [12] etc.

II. LITERATURE REVIEW

Future technologies used to incorporate cloud computing, artificial intelligence, and Internet of Things will eventually help improve the intelligent driving technologies. TSR is therefore an essential element in intelligent transportation systems since the whole safety of driving depends on the understanding of traffic signs and the realization of appropriate avoidance and moving actions. An effective TSR system can ensure the safety of driver by fast and proper provision of real-time traffic information to drivers to make proper decisions or let the vehicle drive on its own to avoid danger. It is in this concern that traffic sign detection technology is at present being emphasized on a research involving important traffic sign identification technologies. Through the evolution of traffic sign identification, we have categorized common detection techniques into four major groups, such as color-based techniques, shape-based techniques, and deep learning-based techniques.

A. Detection based on color-feature

In general, there are three kinds of traffic signs: warning, indication, and prohibition. According to the categories, each type has quite distinct color characteristics. On the RGB color space, Akatsuka et al [5]. marked out the red, yellow, and blue colors with threshold segmentation algorithms to finish the traffic sign detection within a threshold range of distinct colors. Based on this, in the 1980s.A sign-detection system based on RGB color was developed. Zhangka et al. counted all color details of sign images, calculated the red, yellow, and blue thresholds of statistical distribution, and extracted the global color characteristics. Huang Zhiyong et al. used RGB mapping on a row of three-component differences, which supports derivational empirical thresholds, splitting of recognition indications, and avoiding multiplications fully in this technique. B. Using form and feature for detection.

B. Shape-feature-based detection

Anandhalli and Baligar [13] utilized a Harris corner detector in order to locate triangle and rectangle symbols in the ROI, while locating corners in a defined control region. Li [14] exploited edge information during the detection of occluded traffic signs during driving like circle, triangle, and rectangles were identified within the image and masked over 95% of all traffic signs by the use of the nonparametric shape detector with form properties of scale invariant edge turning angles.From the multi-feature fusion of the traffic sign recognition methods, Jin et al. [15] proposed a two-module detector as follows: the first module uses the commonality of symbol borders to extract ROIs and the second module uses HOG and SVM to validate the effectiveness of the produced ROIs to detect the traffic signs of dataset. Zheng et al[16] had proposed sliding window detection(SWD). where this approach integrates channel feature classifier to search for the presence of traffic signals on different sizes. Gim et al. [16] has come up with a system that has two coarse filter modules specifically for the required and prohibited signs in the GTSDB; the first module is HOG and LDA-based. Both modules use an SVM classifier and operate with huge windows; the second uses a small sliding window. Although these attempts prove to be successful in traffic sign recognition with graphical approaches, they are not very good when it comes to complicated situations (like dim lighting, partly covered signs, etc.), and especially bad in recognizing signs with different orientations or perspectives.

C. Detection using deep learning

It utilizes learning and training to identify traffic signs and extract characteristics, which is extremely distinct from previous techniques. Relatively, the deep learning-based approach applies a better accuracy level and more robust generalization ability compared to the classical traffic sign detection approaches. Features acquisition through big data, strong feature-expression capabilities, insensitivity to influences from outside factors such as illumination and occlusion, it is actually a fundamental approach in itself. The algorithm of target detection has high detection accuracy; its basis is Prospective Region extraction. Girshick et al [5].designed a target detection system based on a Prospective Region, which is also popularly referred to as RCNN (Regions with CNN characteristics). True identification of objects and separation of semantic content requires a deep feature hierarchy which used Convolutional Neural Network (CNN). This it categorizes item suggestions, achieving an accuracy of object identification. However, it consumes much computational time and space as it repeats the extraction of every qualifying region and stores it in its memory. Meanwhile, region stretching applied by RCNN pool together every Prospective Region into a 227 × 227 size thus reducing the accuracy for the detection and further reducing the quality of the features CNN extraction. FPN combines, depth feature ,shallow feature to produce predictions on the several scales of feature pyramid scale. Candidate areas are extracted by layering an RPN network over the feature pyramid to achieve this. Such a concept raises the semantic quality of the shallow feature map and amplifies the precision in detecting tiny targets. One is YOLO network, which enforces a single view regarding detection-not as many classification or regression tasks but applies regression on the entire image for predicting the coordinates of the bounding boxes and class probabilities.

III. Methodology

A. Traffic Sign Dataset

Over fifty thousand photos were used from the GTSRB set specially designed for the traffic sign classification. There are forty-three classes into which the various traffic signs fall,like stop sign or no entrance signs, or speed restriction. For the reasons that the photos were taken in a variety of real-world scenarios, involving various lighting, shadows, occlusion, and distances, the classification process is harder and more demanding compared to real case-scenarios. Convolutional Neural Networks are frequently applied for designing machine learning models in order to identify traffic signs. The dataset consists a training set and a test set. Over 39,000 images comprise the training set, broken down into 43 subdirectories, each of which corresponds to a unique type of traffic sign. Thus,the samples in classes are imbalanced. The test set consists of more than 12,000 images and the other file named Test.csv that carries the path for each picture along with its corresponding label for it. Scaling, normalization, as well as other preprocessing techniques can enhance the model, addressing the irregular sizes in the dataset, irregular illumination and orientation. Two of the challenges of the dataset are class imbalance, where some classes contain a lot more images than others, and heterogeneity in image environments, such as varying lighting, occlusion, or weather conditions. Due to these characteristics, GTSRB constitutes an excellent benchmark to build and test advanced models for traffic surveillance and driverless cars. The following table gives an overview of some prominent features of the dataset:

TABLE I
PROPERTIES OF THE DATASET USED FOR TRAINING AND TESTING

Attribute	Description
Number of Images	Over 50,000
Number of Classes	43
Image Resolution	Variable (resized to 32x32 for model input)
Training Set Size	39,000+ images
Test Set Size	12,000+ images
Label File	Test.csv for test images, includes file paths and labels
Challenge Factors	Class imbalance, varying lighting, occlusion, angles
Real-world Conditions	Captured in various environmental conditions

Fig. 1. All types of Images in Dataset

Fig. 2. Visualization of sample traffic sign images from the dataset representing various classes

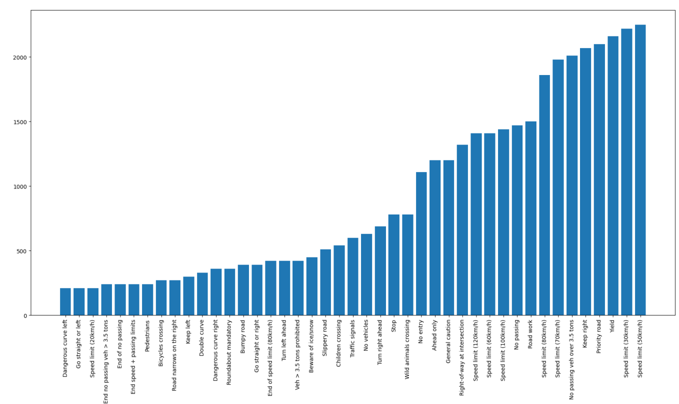

Fig. 3. Number Of Images in Each Class

B. Preprocessing Techniques

Gaussian Blur: Gaussian blur is one of the picture processing techniques that reduce noise of an image and detail. The picture is processed with a Gaussian function to reduce the high-frequency noise and keep the key edge and structure. This is very useful for traffic sign identification as it filters out all unwanted sounds leaving unique features of the traffic sign itself. It becomes soft and distributes itself in all possible directions from the central point such that extraneous features which may confuse the classification algorithm are minimized.

HSV Conversion: Another alternative to the RGB paradigm is the HSV, which can be particularly helpful in jobs involving picture analysis. Value gives the lightness of the color, while saturation is how intense the color is, and hue represents how it is represented. Conversion of the image into HSV helps recognize traffic signs more efficiently as this algorithm can focus more effectively on the signals' characteristics of color rather than getting diverted by brightness changes. This conversion makes color-based segmentation more robust to variations in different lighting conditions.

Binarization: This produces a binary image through binarization-a process that takes grayscale photographs and transforms them into black-and-white images with each pixel having pure black or white color. This results in the removal of in-between shades of gray, and it makes the method differentiate the significant features or characters from the signs, making the system quite helpful in the detecting the traffic signs. TSR uses binarization to make images even simpler for analysis to be made through emphasizing edges and unique features of signs. This allows the differentiation between the sign forms and the rest of the noise, much more so in complex environments.

Morphological operations:These binary images are then processed by morphological techniques, such as dilation, erosion, and closure, in order to further narrow down the shapes by eliminating all sorts of noise. Erosion reduces the borders of an object so that small, insignificant artifacts lose existence, whereas dilation amplifies them and makes features such as edges of a traffic sign much more visible. Finally, closing ensures that characteristics preserved of a traffic sign are actually prepared for analysis by filling gaps and reducing edges. These changes also make the forms of traffic signs much more recognizable and uniform.

ROI (Region of Interest) Extraction: This would reduce the background information with the isolation of the part of the image containing the traffic sign. It would pay more attention to the area of interest and block the other distracting areas of information within the image; hence, the network will learn to categorize the indicators more effectively. The procedure improves the accuracy of classification and gives better redictions for different crowded and heterogeneous conditions because it helps the model to focus only on the road traffic sign, which is most important part of image. only on the road traffic sign, which is the most important part of the image.

Fig. 4. Sample Images After Preprocessing from each Class

C. Model Building

The architecture used for traffic sign categorization uses convolutional neural networks (CNNs). The second layer from the top is a 2D convolutional layer named Conv2D that applies 16 filters and then ReLU activation, followed by another one of 32 filters. Dropout layer and MaxPooling are applied to avoid overfitting and reduce spatial dimensions. For further regularization, the architecture consists of deeper layers of 64 and 128 filters. In addition, there is a MaxPooling layer and Dropout. The feature maps are flattened into the 1D vector. A Dropout layer follows, along with dense layer containing 512 units,the ReLU activation. The output layer classifies the 43 types of traffic sign types, with softmax activation.

- Creating numpy arrays: After processing, numpy arrays are created for processed images as well as processed labels. This is an important step because the data have to be in the numerical array format so that deep learning.Frameworks like TensorFlow or PyTorch can process the input data efficiently.
- Normalization: The values of the pixel of the images are divided by 255.0 for normalization. This normalization brings the values within the range [0, 1] because pixel values normally vary between 0 and 255. Since normalization ensures that the input values are minimal and within a regular range, this in turn helps stabilise the CNN while training and leads to convergence.

$$\text{processed\_images} = \frac{\text{processed\_images}}{255.0}$$

- Label Encoding to_categorical: We need to convert the labels to one-hot encoded vectors with To_categorical. A label of 6, for instance, would translate into a vector of size 43, something like [0, 0, 0, 0, 0, 0, 1, 0,.]. This is necessary because one-hot encoding provides the model with an obvious way of contrasting a predicted output with the true class during training, and classification models such as CNNs return a list of probabilities for each class.

TABLE II
CNN MODEL ARCHITECTURE USED FOR TRAFFIC SIGN CLASSIFICATION

Layer Type	Description	Parameters
Input Layer	Accepts images of size 30x30 with 3 color channels (RGB).	Input shape = (30, 30, 3)
Conv2D Layer 1	2D Convolutional layer with 16 filters, kernel size 3x3, and ReLU activation.	Filters = 16, Kernel size = (3,3), Activation = ReLU
Conv2D Layer 2	Conv2D with 32 filters, kernel size 3x3, and ReLU	Filters = 32, Kernel size = (3,3), Activation = ReLU
MaxPool2D Layer 1	MaxPool2D layer to down sample feature maps.	Pool size = (2, 2)
Dropout Layer 1	Dropout layer to reduce overfitting.	Dropout rate = 0.25
Conv2D Layer 3	Conv2D layer having 64 filters, kernel size 3x3, and ReLU activation.	Filters = 64, Kernel size = (3,3), Activation = ReLU
Conv2D Layer 4	Conv2D layer having 128 filters, kernel size 3x3, and ReLU activation.	Filters = 128, Kernel size = (3,3), Activation = ReLU
MaxPool2D Layer 2	MaxPool2D layer to down sample feature maps.	Pool size = (2, 2)
Dropout Layer 2	Dropout layer to reduce overfitting.	Dropout rate = 0.25
Flatten Layer	Flattens the 2D matrices into a 1D vector for the fully connected layer.	-
Dense Layer 1	Fully connected layer with 512 neurons and ReLU activation.	Units = 512, Activation = ReLU
Dropout Layer 3	Dropout layer to reduce overfitting.	Dropout rate = 0.5
Dense Output Layer	Fully connected layer with 43 neurons (for 43 classes) and softmax activation.	Units = 43, Activation = Softmax

D. Model Enhancement

- Hyperparameter tuning: To tune hyperparameters, grid search or random search is applied in order to reduce the amount of calculation time and therefore improve accuracy.

- Early Stopping: This is a kind of regularizer against the overfitting problem; training will stop when the model's performance on the validation set begins to degrade.
- Dropout: This is another technique applied in layers. At each training step, it randomly removes units so that the model becomes more broadly applicable and less sensitive to any given single neuron.

IV. Result

The CNN's first neural network model The data from the train set is expressed in a set of graphs shows the accuracy of the train set, which was 70% used to train model, and that of test set, which was 30%. Apart from demonstrating the accuracy of the performance of the model, some test were carried out on the test datasets to confirm the correctness of the model. The accuracy of the model was 0.9969, with the test loss of 0.0102 and this is a fair assessment.

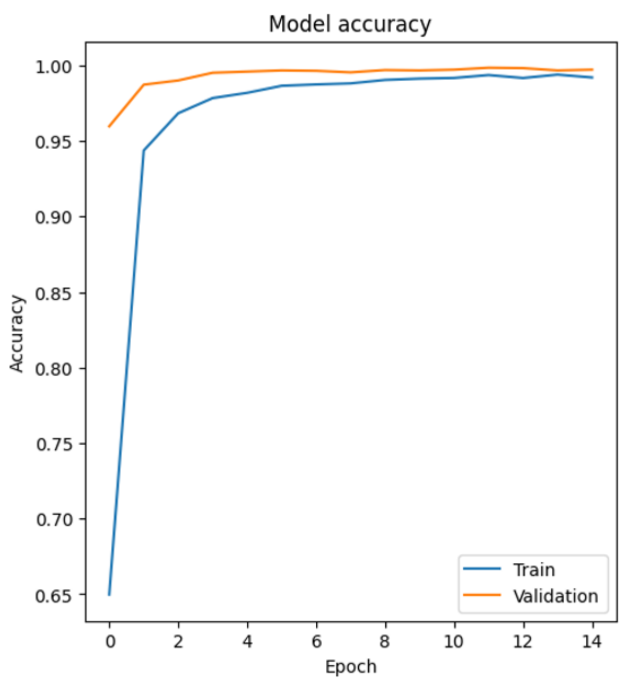

Fig. 5. Model Accuracy

V. Execution Time

The computation time spent during testing and training of the traffic sign recognition model was supported by the time that had to complete the project, which includes importing data, scaling and standardizing photos, training the model with the CNN architecture, and performing assessment processes. However, it would still depend on many factors such as size of dataset, epochs, the depth of the CNN model, and the amount of preprocessing done on images before entry into the CNN. For example, in my experiment, I have used morphological operations, HSV conversion, and a Gaussian blur on all images before feeding them into the CNN. Optimization of these operations and judicious use of the GPU resources directly

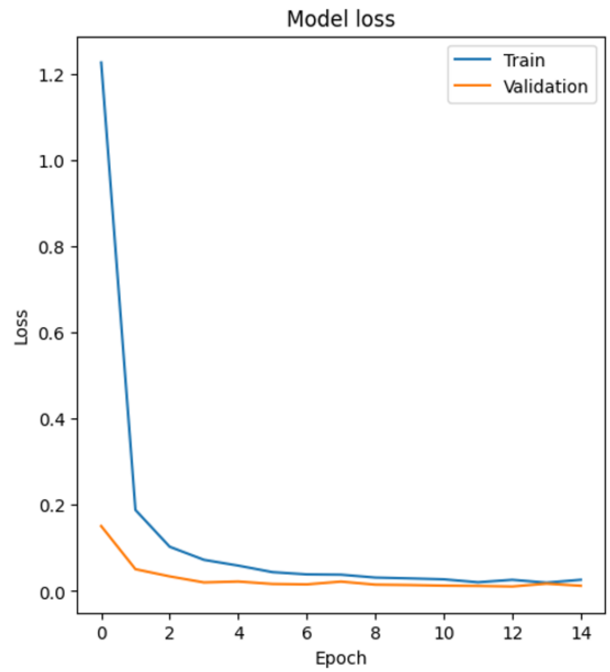

Fig. 6. Model Loss

reflects increased speed with relatively reduced execution times.

Table below displays the testing results of a model with a dropout layer in validation test. It indicates that the model performs equally well for seven groups, and evaluation of each group is performed in terms of time required in each iteration, recognition accuracy, and the validation loss. The time usage in iterations ranges between 2 and 18 seconds, where identification accuracy ranges from 95.97% to 99.64% with an increase in the number of iterations. The validation loss drops drastically, from 0.1509 to 0.0156, showing generalization improvement and less overfitting for each group.

TABLE III
VALIDATION TEST RESULTS WITH DROPOUT LAYERS

Group	Time/Iteration	Recognition accuracy	Validation Loss
1	18s	0.9597	0.1509
2	3s	0.9872	0.0508
3	2s	0.9901	0.0341
4	2s	0.9952	0.0200
5	3s	0.9959	0.0221
6	3s	0.9967	0.0165
7	3s	0.9964	0.0156

VI. Conclusion

Detection and recognition of traffic signs play a very important role in sustaining road safety, considering that most of the drivers will respect the laws of traffic only. In this article, the automation for the detection process has been

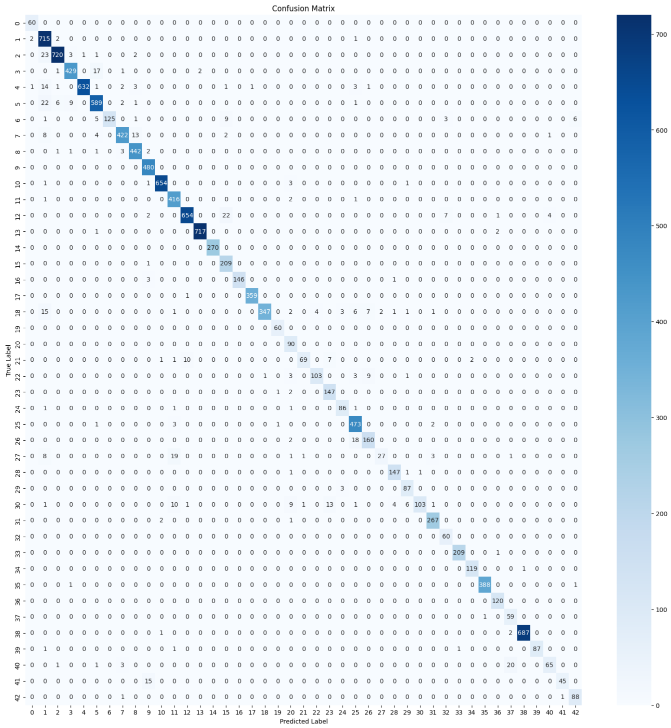

Fig. 7. Confusion Matrix

Fig. 8. Random image predictions from the traffic sign classification model

successfully achieved, along with the use of deep learning-based approaches, including Convolutional Neural Networks (CNNs). This technology is definitely a great advantage because it holds tremendous benefits for the distracted or visually defective drivers who may be quite unable to notice the signs or recognize them in realtime. It is important to note that the model's input is obtained through advanced image preprocessing techniques like morphological operations, transformation to HSV space, and Gaussian blur. Such augmentations help the model better identify signs in varied scenarios. This use of CNNs was very effective in learning the specific features of traffic signs,resulting in high classification accuracy. Adaptability and flexibility of CNNs combined with improvements in the training process data-driven make this a strong and reliable real-world Traffic Sign Recognition system. It is an essential contribution to the development of intelligent transportation systems that improve road safety and driving experience for everyone, and it offers good opportunities for further increasing accuracy and for its real-time deployment.

References

[1] C. Liu, S. Li, F. Chang and Y. Wang, "Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives," *IEEE Access*, vol. 7, pp. 86578-86596, 2019, doi: 10.1109/ACCESS.2019.2924947.

[2] X. Peng, X. Chen and C. Liu, "Real-time traffic sign text detection based on deep learning," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 768, Mar. 2020.

[3] Y. Du, X. Zhang, G. Zhang and W. Sun, "Detection and recognition of text traffic signs above the road," *Int. J. Sensor Netw.*, vol. 35, no. 2, pp. 69-78, 2021.

[4] William, Marco Magdy, Pavly Salah Zaki, Bolis Karam Soliman, Kerolos Gamal Alexsan, Maher Mansour, Magdy El-Moursy, and Kerolos Khalil. "Traffic signs detection and recognition system using deep learning." In *2019 Ninth international conference on intelligent computing and information systems (ICICIS)*, pp. 160-166. IEEE, 2019.

[5] X. Qiao, "Research on Traffic sign recognition based on CNN Deep Learning Network," *Procedia Computer Science*, vol. 228, pp. 826-837, 2023.

[6] A. Alam and Z. A. Jaffery, "Indian traffic sign detection and recognition," *International Journal of Intelligent Transportation Systems Research*, vol. 18, pp. 98-112, 2020.

[7] Y. Liu and W. Zhong, "A novel SVM network using HOG feature for prohibition traffic sign recognition," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.

[8] Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, and W. Liu, "Traffic sign detection and recognition using fully convolutional network guided proposals," *Neurocomputing*, vol. 214, pp. 758-766, 2016.

[9] S. Mehta, C. Paunwala, and B. Vaidya, "CNN based traffic sign classification using adam optimizer," In *2019 international conference on intelligent computing and control systems (ICCS)*, pp. 1293-1298. IEEE, 2019.

[10] L. Jia, X. Shi, and L. Wei, "Design of Traffic Sign Detection and Recognition Algorithm Based on Template Matching," In *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, pp. 237-240. IEEE, 2020.

[11] Z. Fazekas, G. Balázs, A. Boulmakoul, and P. Gáspár, "Visualization of traffic sign related rules used in road environment-type detection," In *2019 Modern Safety Technologies in Transportation (MOSATT)*, pp. 53-57. IEEE, 2019.

[12] A. Aksjonov and V. Kyrki, "Rule-based decision-making system for autonomous vehicles at intersections with mixed traffic environment," In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 660-666. IEEE, 2021.

[13] M. Anandhalli and V. P. Baligar, "An approach to detect vehicles in multiple climatic conditions using the corner point approach," *J. Intell. Syst.*, vol. 27, no. 3, pp. 363-376, Jul. 2018.

[14] L. L. Li, "Image edge detection algorithm and its application in traffic video segmentation," 2015.

[15] Y. Jin, Y. Fu, W. Wang, J. Guo, C. Ren and X. Xiang, "Multi-feature fusion and enhancement single shot detector for traffic sign recognition," *IEEE Access*, vol. 8, pp. 38931-38940, 2020.

[16] S. He et al., "Automatic Recognition of Traffic Signs Based on Visual Inspection," *IEEE Access*, vol. 9, pp. 43253-43261, 2021, doi: 10.1109/ACCESS.2021.3059052.