ORIGINAL ARTICLE

IJCIH
Converging Healthcare & Technology

INTERNATIONAL JOURNAL OF CONVERGENCE IN HEALTHCARE

Published by
IJCIH & Pratyaksh Medicare LLP

www.ijcih.com

An Efficient Traffic Sign Classification and Recognition with Deep Convolutional Neural Networks

Deepika[1], Sharda Vashisth[2], Prabha Sharma[3]

[1]Research Scholar, Department of Multi-Disciplinary Engineering, The NorthCap University, Gurugram, India, [2]Professor, Department of Multi-Disciplinary Engineering, The NorthCap University, Gurugram, India, [3]Professor, Department of Applied Sciences, The NorthCap University, Gurugram, India

Abstract

Driver safety assisted driving, and autonomous driving all depend on automated traffic sign recognition. The most popular deep learning method for recognizing traffic signs is convolutional neural networks. (CNNs). This paper presents a useful technique for automatically detecting images of traffic signs. When employing the two common traffic sign photo datasets GTSRB, our method makes use of our CNN model architecture and performs the best. It aids the driver in safely operating the motor vehicle. The amount of time and effort drivers spend manually evaluating and recognizing traffic signs is excessive. This work provides an autonomous traffic sign identification using a convolutional neural network. Our work here introduces a unique CNN architecture with an Adam optimizer and a batch size of 128 to improve the efficiency of traffic sign recognition. Results based on a complex network were more accurately produced by a convolutional neural network (CNN). With 99.81% precision and a minimum of losses, our system learns from the GTSRB dataset, which contains 43 traffic classes, to identify the correct class of an anonymous traffic sign. The results, however, are better than those of the prior research, which examined this approach's accuracy and efficiency in recognizing traffic signs despite poor weather and hazy image circumstances.

Keywords: *Convolutional neural networks (CNN), Deep learning algorithms, Automatic traffic sign recognition, GTSRB, Standard image dataset.*

Introduction

By using cutting-edge driver assistance systems, traffic sign recognition has assisted in raising driver awareness and, as a result, preventing accidents. A variety of issues, including lighting, rotation invariance, blurry shapes, and color changes in the traffic signs, can sometimes make it difficult to detect them adequately. To detect traffic signs, we have two different sorts of approaches: the first one focuses on conventional image processing methods, while the second one is deep learning-based. Traditional and deep learning-based approaches differ mostly in how well they perform as the size of the data rises. Because deep learning-based approaches require a significant quantity of data to fully grasp them, they perform poorly when our data is little. The constraints of traditional approaches are that we must extract features, that features are not

Corresponding Author:
Deepika
Deptt. of Multi-Disciplinary Engineering, The NorthCap University, Haryana India
email: deepika17ecd001@ncuindia.edu

universal, and that features must be extracted based on the problem domain. The traffic signs are simple for both humans and machines to understand. The influence of interfering variables on the cameras' images, such as the weather, illumination, shooting angle, object occlusion, and others, makes it difficult to practice [1,2,3,4]. Moreover, the traffic signs in the photograph are difficult to see or distinguish because they are too small [4]. TSDR faces many difficulties as a result of this. The ability of autonomous vehicles to recognize traffic signs is crucial for their reactions and safety. On the other hand, it is very difficult to recognize traffic signs due to the signs' small size. Contrarily, recognizing traffic signs is more crucial for accuracy than labeling [5, 6]. Convolutional Neural Networks (CNN) [7] are a type of deep neural network that is comparable to the visual processing of human vision [8, 4] and can learn more discriminative characteristics. In traffic sign identification algorithms, CNN performs better than the current top methods [9]. Traditional research techniques, such as support vector machines and simple machine learning models, are used to classify traffic signs[10].

Recently, several techniques proposed for traffic sign recognition that tackles various challenges like illumination, motion blur, and occlusion. Deepika et al. [11] incorporate focal loss for traffic sign detection and recognition in real-time scenarios. For instance, the technique of histogram of oriented gradients [12] was used as the foundation for a traffic sign identification algorithm. Principal component analysis was used to reduce dimensionality, the Gaussian filter and histogram equalization were combined for efficient image preprocessing, a kernel extreme learning machine (K-ELM) classifier was used to achieve good classification accuracy, and the average processing time was also reduced. Aziz et al. [13] introduced Extreme learning machine techniques that performed remarkably well when enhanced with complimentary mid-level features. Localization of traffic sign instances within an image, however, is still a problem.

The effectiveness of deep learning models is strongly influenced by the dataset size. The performance of the model improves with the size of the dataset. Deep learning model training is rather challenging due to the small sample size because these models struggle with overfitting and generalization errors. The author [14] utilizes the German traffic sign recognition benchmark dataset to classify the traffic signs efficiently. The author uses images with artificial translation and rotations to increase the dataset size. Moreover, the methods employed by Kuros et al. [15] utilize quantum neural networks to classify traffic signs. Thus to lessen the problems with the data the author uses the German traffic sign dataset to cope with such issues. The structure of our article is clearly explained in Fig 1 below.

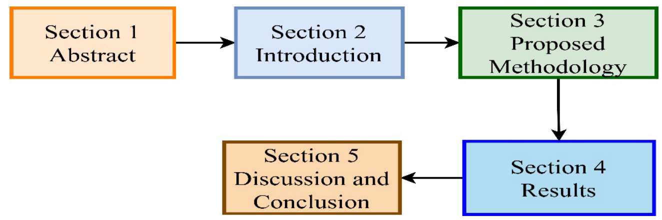

Fig 1. Paper Format

In summary, the major contributions of the papers are structured as follows:

- Using the benchmark dataset, design a CNN-based traffic sign recognition system
- Evaluated the performance of the proposed CNN model on the GTSRB benchmark available dataset and compared their performance with various already existing CNN models.
- Optimized the model using the TensorRT SDK to get real-time results.

The rest of the paper is organized as follows: The introduction to the strategies based on traffic sign recognition is reviewed in Section 2. The proposed methodology for the traffic sign recognition system is described in Section 3. The suggested model's prediction results using the GTSRB dataset are shown in Section 4 along with comparison results. Section 5 brings the report to a close with a discussion and recommendations for the future.

Proposed Methodology: A deep CNN-based traffic sign recognition system is trained using preprocessed images from the traffic sign dataset. It can detect and identify a wide range of signs and symbols. Images are initially transformed into a usable shape from which feature extraction is easier. Then, utilizing the characteristics extracted from CNN's various layers, picture preprocessing techniques identify traffic signs. Our proposed model starts with a convolution layer that imposes 16 filters of size 3x3, which are executed by four lateral pathways: two of them use the sequential CNN flow, and the other two use the hybrid feature approach by including an accretion layer to expand the network in an accretive manner. The accretion layer mixes the hybrid responses produced by the preceding levels. These layers preserve the essential characteristics of the feature maps while improving the neurons' capacity to learn even the smallest details. There are 16, 32, and 64 filters of sizes 1x1, 3x3, and 5x5 on each path, respectively. The feature replies are concatenated and sent to the next convolution layer with a 256 filter of size 3x3 after receiving feature maps from all the paths the feature responses are concatenated and forwarded to the next convolution layer with 256 filters of size 3 × 3 as shown in Fig. 2.

To improve the categorization accuracy of traffic signs, a CNN model was used in this study. On the GTSRB dataset, data preprocessing was first carried out, including Gray Scale conversion, Histogram Equalization, and Normalization. Data augmentation comes after data preparation. To avoid overfitting, data augmentation increased the quantity of training data. The TS CNN model was then put into use and trained using the GTSRB dataset. The following is a detailed description of the architecture as explained below:

Data Preprocessing: The GTSRB dataset [16], which consists of traffic photographs, was employed by the suggested system. 34799 labeled images make up the training set, 4410 labeled images make up the validation set, and 12630 images make up the test set. 43 classes in the dataset have uneven class frequencies [17].

All of the photographs were taken from a series of movies that were shot from moving vehicles at various times of the day and in various lighting conditions, eliminating the problem of motion blur. As a result, the GTSRB dataset has weak contrast and low resolution. Preprocessing data is therefore necessary to improve accuracy. several traffic signs, including those that read "Speed Limit," "No Passing," "Priority Road," "Yield," "Stop," "Bumpy Road," and "Pedestrians,".

The process of normalization involves transforming an input image into a set of pixel values that are familiar to the senses. Since In this approach, grayscale photos are used, and normalization only requires one channel.

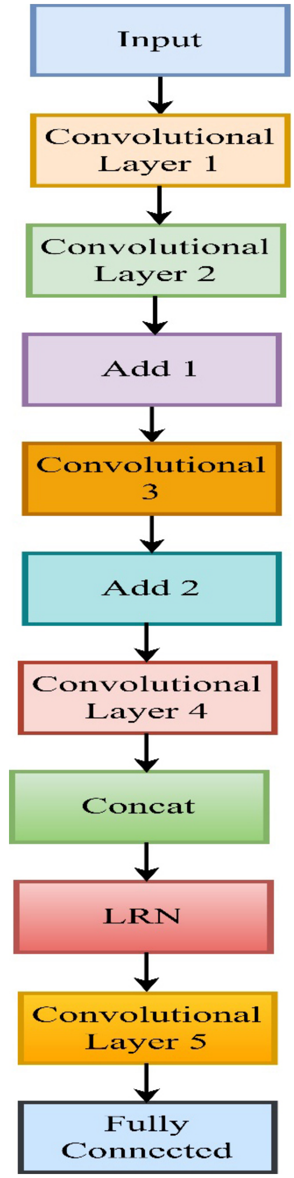

Fig 2. CNN model architecture.

Data Augmentation: Using label-preserved changes, or data augmentation, to artificially increase the dataset is the simplest and most popular technique for reducing overfitting on picture data[18]. By flipping, rotating, and zooming the data, data augmentation is used to increase the number of images in the dataset. The Keras ImageDataGenerator function is used in our model to enhance the data.

Model Summary: A Keras Sequential model is what we have. To construct deep learning networks, Keras, an open-source framework, offers incredibly strong and abstract building pieces and functions as an API on top of TensorFlow. For the majority of issues, you can build models layer-by-layer using Keras' sequential API. Four convolutional layers, two pooling layers, one flatten layer, and four fully connected layers make up the traffic sign CNN model. Table 1 shows the proposed model's network configuration specifications.

Table 1. Detailed configuration of the proposed model

Type	Sub-Layer	Filter	Stride	Output	No of parameters (w+b)
Input	-	-	-	112X11X3	-
Conv-1	-	3X3	2	56X56x16	432+16
Conv-2	2.1 2.2 2.3 2.4	1x1	2	28X28X16	4X (256+16)
Add-1	1.1 1.2	-	-	28X28X16	-
Conv-3	3.1 3.2 3.3 3.4	3X3	2	14X14X32	4X (4608+32)
Add-2	2.1 2.2	-	-	14X14X32	-
Conv-4	4.1 4.2 4.3 4.4	5X5	2	7X7X64	4X (51200+64)
Concat	-	-	-	7X7X256	-
LRN	-	-	-	7X7X256	256+256
Conv-5	-	3X3	2	4X4X256	589824+256
FC	-	-	-	1X1X256	1048576+256

Two prior feature maps of the proposed architecture's paths follow the conventional sequential CNN approach, while the third path adopts a hybrid approach by integrating an accretion layer. To secure more precise information, the accretion layer integrates the key elements of two separate layers. Fig. 4 demonstrates how some neurons at Conv 2.1 and Conv 2.2 were unable to recognize the salient characteristics of the expressive regions. According to the above table, some neurons at Conv 2.1 and Conv 2.2 are unable to recognize the key characteristics of the expressive areas. However, by combining the preferred feature maps of two layers, the accretion layer enriches the features. We use several filter sizes, including 1 1, 3 3, and 5 5, in the proposed network to collect both high-level and micro-level information, as inspired by the Inception layer. We find that a smaller filter size is more effective at recognizing traffic signs. A larger filter size (7 7) could miss the little details. In contrast to other CNN-based models, the LEARNet model does not incorporate a max pooling layer for downsampling. In essence, pooling is the process of applying the max function to obtain high-level edge information. However, because pooling ignores the micro-level data and only extracts high-level features by applying the max function to the input data, it misses some of the important features. However, because facial muscle movement in microexpressions is faint and fleeting, any loss may result in an incorrect classification. In contrast to max pooling, LEARNet uses the convolution layer with stride 2 to downsample the feature maps and capture more data.

Results

Using the Keras Library, traffic sign recognition was accomplished using the TS CNN model. The model employed 50 iterations or 50 times each for a forward pass and backpropagation across the neural network with complete training data. The prediction results on the German traffic sign recognition benchmark dataset as shown in Fig 3. The plot of the model's Accuracy, precision, recall, and F1 score vs the number of epochs is shown in Fig. 4. The epoch has an impact on the model's accuracy. Similarly, the graphic makes it clear that model loss reduces with epoch.

Keras Deep Learning Traffic Sign Classification Results on GTSRB Dataset:

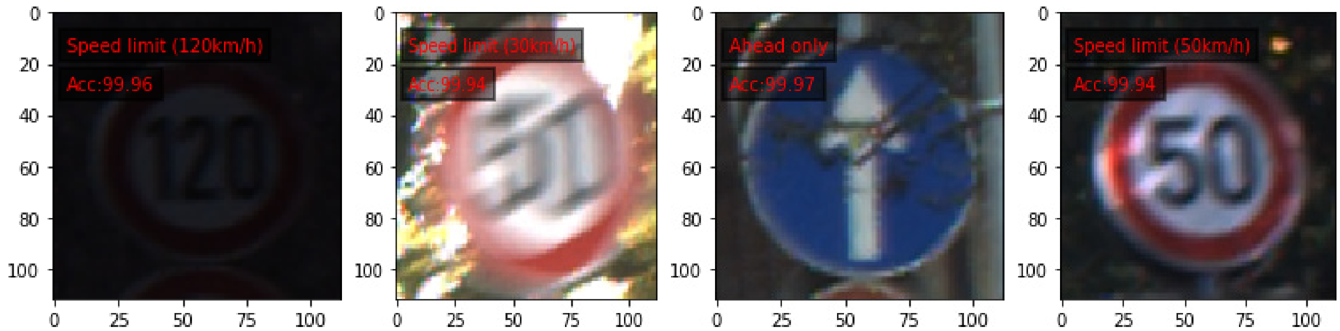

Fig 3. Classification results on the GTSRB dataset.

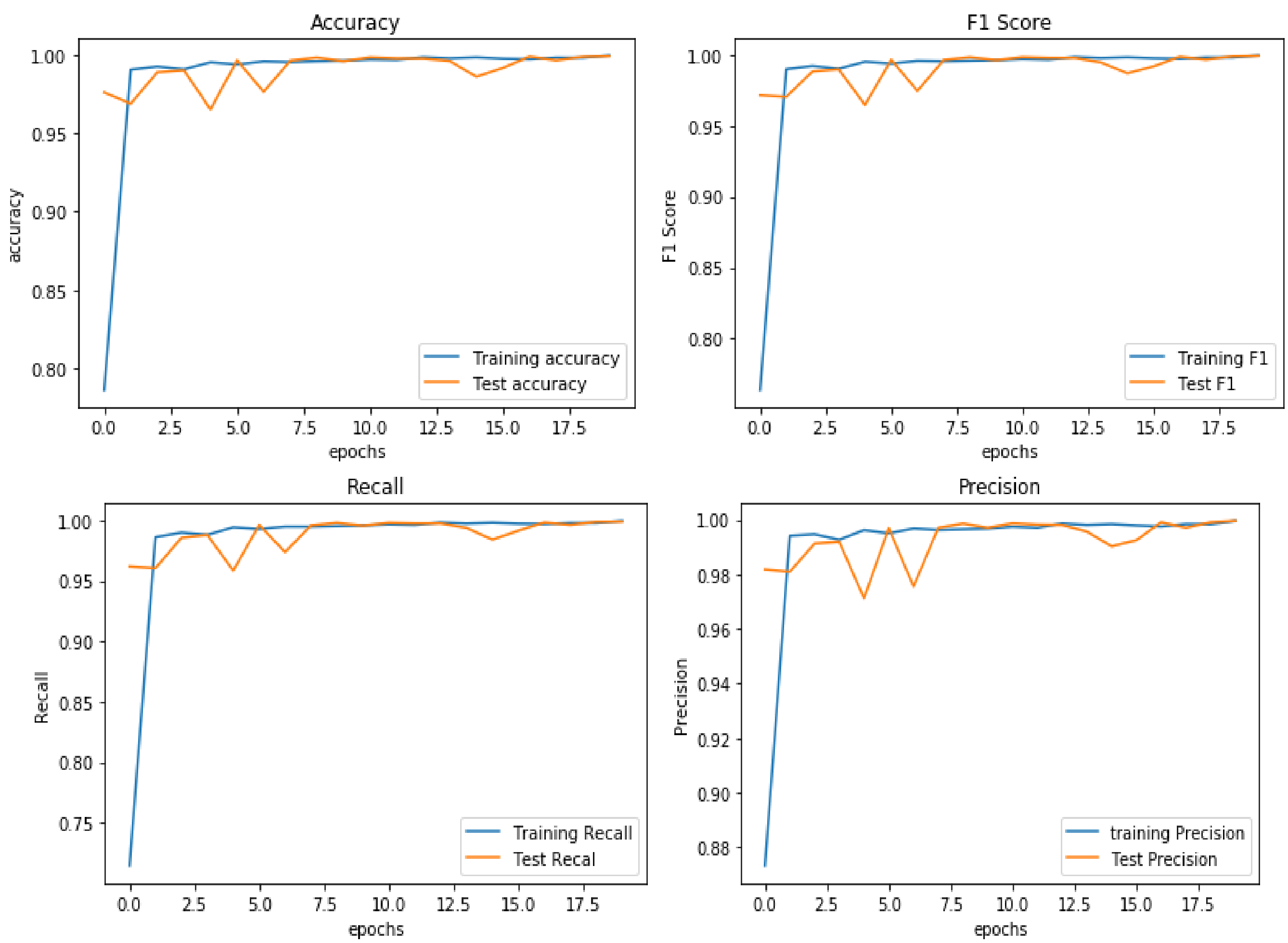

Fig 4. Traffic sign recognition metrics (Accuracy, Recall, Precision, F1 score) vs the number of epochs.

Table 2. Comparative analysis with state-of-the-art results.

Model	Accuracy
CNN [4]	88%
RetinaNet Resnet 50 [11]	96.45%
RetinaNet Resnet 101 [11]	97.45%
GF +HE+HOG [12]	98.54%
CNN [18]	80.50%
Proposed Model	99.81%

The comparison results with the already existing architectures are shown in Table 2 above. Our model was tested using 12630 images in the GTSRB dataset and an accuracy of 99.81% has been achieved. RetinaNet Resnet 50 [11] delivered an accuracy of 96.45% while the same model with Resnet 101 delivers an accuracy of 97.45%. The histogram of oriented gradient-based classification methods in [12] gives an accuracy of 98.54% while the convolutional neural network-based method in [18] provides an accuracy of 80.50%.

Discussion and Future Scope

Using the GTSRB dataset, this study implemented the Traffic Sign Classification utilizing TS CNN Model. The dataset underwent data preparation procedures to enhance the contrast and resolution. Data Augmentation, which adds more photos to the training samples to prevent overfitting, comes after data preprocessing. Our model significantly outperformed the test data with an accuracy of 98.44%. Our model can outperform currently available cutting-edge algorithms.

In the future, we want to improve accuracy by classifying traffic signs with different CNN models.

Conflict of Interest: All authors declare that they have no conflict of interest.

Funding: No funding.

Ethical Clearance: Not Applicable.

References

1. Serna CG, Ruichek Y. Traffic signs detection and classification for European urban environments. IEEE Transactions on Intelligent Transportation Systems. 2019 Sep 23;21(10):4388-99.
2. Wan J, Ding W, Zhu H, Xia M, Huang Z, Tian L, Zhu Y, Wang H. An efficient small traffic sign detection method based on YOLOv3. Journal of Signal Processing Systems. 2021 Aug;93:899-911.
3. Zhang J, Xie Z, Sun J, Zou X, Wang J. A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection. IEEE access. 2020 Feb 7;8:29742-54.
4. Zhu Z, Liang D, Zhang S, Huang X, Li B, Hu S. Traffic-sign detection and classification in the wild. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 2110-2118).
5. Fu J, Zhao C, Xia Y, Liu W. Vehicle and wheel detection: a novel SSD-based approach and associated large-scale benchmark dataset. Multimedia Tools and Applications. 2020 May;79:12615-34.
6. Liu Z, Li D, Ge SS, Tian F. Small traffic sign detection from large image. Applied Intelligence. 2020 Jan;50:1-3.
7. Loukmane A, Graña M, Mestari M. A model for classification of traffic signs using improved convolutional neural network and image enhancement. In2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS) 2020 Oct 21 (pp. 1-8). IEEE.
8. Alghmgham DA, Latif G, Alghazo J, Alzubaidi L. Autonomous traffic sign (ATSR) detection and recognition using deep CNN. Procedia Computer Science. 2019 Jan 1;163:266-74.
9. Sajja TK, Kalluri HK. Image classification using regularized convolutional neural network design with dimensionality reduction modules: RCNN–DRM. Journal of Ambient Intelligence and Humanized Computing. 2021 Oct;12(10):9423-34.
10. Zeng Y, Xu X, Fang Y, Zhao K. Traffic sign recognition using deep convolutional networks and extreme learning machine. InIntelligence Science and Big Data Engineering. Image and Video Data Engineering: 5th International Conference, IScIDE 2015, Suzhou, China, June 14-16, 2015, Revised Selected Papers, Part I 5 2015 (pp. 272-280). Springer International Publishing.
11. Kamboj D, Vashisth S, Saurav S. Evaluation of

focal loss based deep neural networks for traffic sign detection. International Journal of Image and Data Fusion. 2022 Jun 23:1-23.

12. Vashisth S, Saurav S. Histogram of Oriented Gradients based reduced feature for traffic sign recognition. In2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI) 2018 Sep 19 (pp. 2206-2212). IEEE.
13. Aziz S, Youssef F. Traffic sign recognition based on multi-feature fusion and ELM classifier. Procedia Computer Science. 2018 Jan 1;127:146-53.
14. Haque WA, Arefin S, Shihavuddin AS, Hasan MA. DeepThin: A novel lightweight CNN architecture for traffic sign recognition without GPU requirements. Expert Systems with Applications. 2021 Apr 15;168:114481.
15. Kuros S, Kryjak T. Traffic sign classification using deep and quantum neural networks. InComputer Vision and Graphics: Proceedings of the International Conference on Computer Vision and Graphics ICCVG 2022 2023 Feb 11 (pp. 43-55). Cham: Springer Nature Switzerland.
16. https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign.
17. Stallkamp J, Schlipsing M, Salmen J, Igel C. The German traffic sign recognition benchmark: a multi-class classification competition. InThe 2011 international joint conference on neural networks 2011 Jul 31 (pp. 1453-1460). IEEE.
18. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Communications of the ACM. 2017 May 24;60(6):84-90.
19. Persson S. Application of the German Traffic Sign Recognition Benchmark on the VGG16 network using transfer learning and bottleneck features in Keras.
20. Kumar AD. Novel deep learning model for traffic sign detection using capsule networks. arXiv preprint arXiv:1805.04424. 2018 May 11.