Traffic Sign Recognition using CNNs

Pritesh Gurjar[1], Neev Shirke[2], Navin Bubna[3], Malay Khakhar[4], Noshir Tarapore[5]

[1]LY B.Tech Computer Engineering, Science & Technology, Vishwakarma University, Pune, India - 411048
[2]LY B.Tech Computer Engineering, Science & Technology, Vishwakarma University, Pune, India - 411048
[3]LY B.Tech Computer Engineering, Science & Technology, Vishwakarma University, Pune, India - 411048
[4]LY B.Tech Computer Engineering, Science & Technology, Vishwakarma University, Pune, India - 411048
[5]Assistant Professor, Dept. of Computer Engineering, Vishwakarma University, Pune, India - 411048

---***---

Abstract- *This paper aims to summarize the usage of Traffic sign detection and identification and how it could play a massive role in increasing the safety of people while driving. Through this paper, we shall see how a Traffic Sign Recognition system can be implemented using advanced Machine Learning and Convolutional Neural Networks. We believe that implementing such systems could prove beneficial for the evolution of the current safety standards while driving.*

KEYWORDS- TSR (Traffic Sign Recognition), Traffic Sign detection, CNN (Convolutional Neural Network), SSD (Single Shot multibox Detector) , YOLO net (Object detection system).

1. INTRODUCTION

Automobiles are the most efficient way of getting around in this day and age. For many years, the way we transport things has been evolving with improvements at every iteration. From steam and electric vehicles in the 1700s to alternative fuel automobiles and self-driving cars in the 2000s, automobiles have experienced several developments. As humans advance further, so should our work and everything surrounding us. To better the safety standards of the passengers in a vehicle, we designed and proposed a system that uses algorithms like Machine Learning and Convolutional Neural Networks to identify Traffic signs quickly and accurately. We believe our implementation could help in the future of self-driving cars, making the experience safer and more peaceful for all the passengers.

We chose to create such a system to ensure a safe driving experience for everyone. This could even be progressed further and used in self-driving automobiles in the future. By implementing a traffic sign recognition system, we aim to reduce the percentage of on-road accidents by a large percentage. Statistics show that reckless driving and distracted driving cause a large number of accidents and deaths by accidents. Implementing a system that identifies a sign from afar and notifies the driver of the upcoming event would improve the safety standards and reduce the number of accidents.

The aims, scope and objectives of the System are as follows

- Utilising suitable datasets
- Comparing with local signs and modifying datasets if required
- Making use of a two-tier CNN along with YOLO Networks 4
- Getting accurate results with ample time to alert the driver.

2. RELATED WORK

Regular occurrences of terrible accidents result in the loss of life and other valuables. There might be a variety of causes for these accidents, such as bad street maintenance, careless driving, the driver's psychological state, and pedestrians' careless attitude. Another important factor for this might be improper law enforcement and poor conditions of street traffic signs. Signs that are blocked or decaying may confuse the driver. So there are many existing methods for Traffic signs and sign classification.

Zhang et al. [1] proposed a cascaded R-CNN to obtain the multiscale features in pyramids and a multiscale attention method to obtain the weighted multiscale features by dot product and softmax to highlight the traffic sign features and improve the accuracy of the traffic sign detection.

Cao et al. [2] proposed an improved algorithm based on faster region-based CNN for small object detection. An improved loss function based on intersection over union, the multiscale convolutional feature fusion, and the improved non maximum suppression algorithm is introduced to enhance the performance for small object detection.

Since around 2007, traffic sign identification and recognition techniques have relied on color segmentation, shape, and texture data in combination with support vector machine (SVM) classifiers. Later illumination conditions were investigated using a shape-based detector[3]. Slightly later, shapes were utilized to detect arrow traffic lights[4]. More recently Color segmentation has been utilized. Ji et al. [5] introduced a color-based visual selective attention model to build salience maps, which are subsequently identified using an SVM classifier using histogram of oriented gradient (HOG) features. Some prior art used digital maps and GPS data to enhance detection efficiency and accuracy [6]. However, prior knowledge is not always available and is not always required. Because of the widespread use of convolutional neural networks (CNNs), they have been used for traffic sign recognition.

Another method suggests the design and study of the "German Traffic Sign Recognition Benchmark" dataset and competition. The competition results indicate that machine learning algorithms perform well in the difficult task of traffic sign identification. The participants obtained a very high performance of up to 98.98 percent accurate identification rate on this dataset, which is comparable to human performance[7].

Another method involves the use of a new system for the automated detection and identification of traffic signs. Candidate regions are identified as maximally stable extremal regions (MSERs), which provide resistance to variations in lighting conditions. A cascade of SVM classifiers trained on HOG features is used for recognition. This system is accurate at high vehicle speeds, works in a variety of weather situations, and runs at a pace of 20 frames per second on average[8].

3. CURRENT IMPLEMENTATION

While current systems are capable of identifying traffic signs with high accuracy, our system aims to better these results and achieve high accuracy results even in poor weather conditions like severe rain, dust, etc. This system could be useful for the future of self-driving vehicles and may help in improving the current systems used by various organizations/ companies. The system architecture is a two-tier architecture, where the first tier comprises a Convolutional Neural Network. The second tier consists of image processing using various techniques to increase the accuracy of sign identification, even in varying atmospheric conditions. The current system has been created solely by using python, and it's multiple libraries such as numpy, tensorflow, jupyter, YOLOnet4, and cv2 for the neural network and calculations. We have also used libraries such as pygame and tkinter to create a simple user interface.

The hardware components of our current system consist of a camera and a simple graphics processing unit

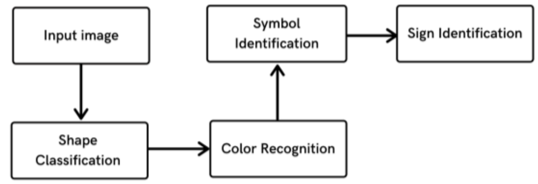

Fig 3.1: Block diagram of current working

4. HIGH LEVEL DESIGN DETAILS

Relevant mathematical model associated with the Project:

1. The model is based on the simulation of traffic signs using recorded videos with various factors affecting quality to make the system optimal.

2. The video is then processed to extract the frames with some delay.

3. Each frame will be processed through convolution neural networks to identify the type and information of traffic signs.

4. According to the 43 types of traffic signs the system will generate appropriate alerts based on picture classification.

Attached below are various diagrams which explain the various steps, methods, functions and libraries implemented in our current working model.

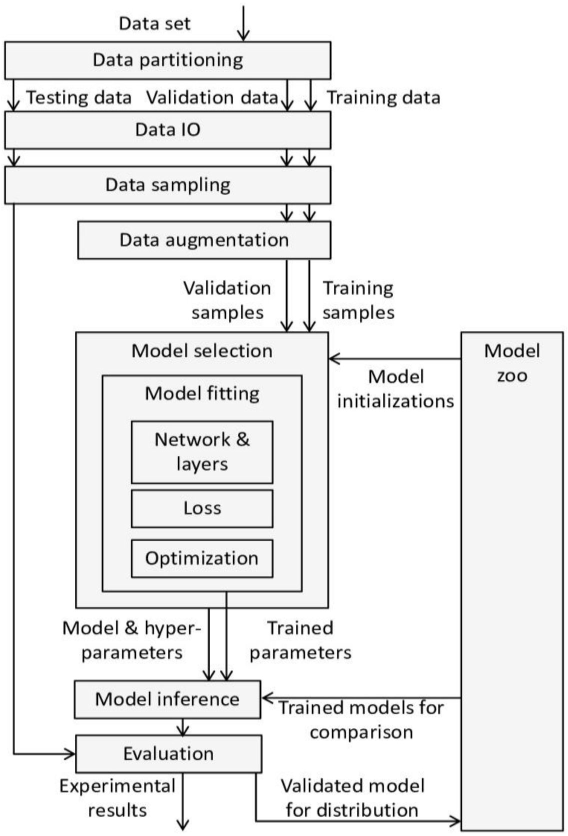

Fig 4.1: Data flow diagram

The above diagram shows the flow of data in the current implementation. The testing and training datasets are used to identify the correct sign with the highest probability.

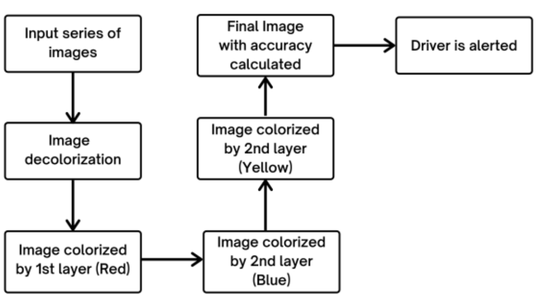

Fig 4.2: Simple flowchart of working of image recognition

The above flowchart explains how our model works in the simplest form. The image / series of images received from the input are passed on for identification. This happens in 3 steps.

- The original image is decolorized to a grayscale version, so that the appropriate sign may be identified on the basis of the shape of the sign.
- Next, the colors are restored one layer at a time (Red, Blue, Yellow) to further identify the type of sign based on its previous result.
- Finally, the now colorized sign is further passed through the dataset to identify any symbols (if present) in the sign itself.
- Once the sign is identified, the system alerts the driver of one of 5 alerts.

V. NEURAL NETWORK ARCHITECTURE

Using a fully connected neural network to make an image classification requires numerous layers and neurons present in a network, which increases the number of parameters leading the network to over-fitting (memorizing the training data only). The input image may also lose its pixels correlation properties since all neurons (carrying pixels values) are connected [9]. Convolutional neural networks have emerged to solve these problems through their kernel filters to extract the main features of the input image and then inject them into a fully connected network to define the class [9]. The chosen architecture in our application is a two tier convolutional neural network (Fig. 5.1) firstly used for handwritten digits recognition [10]. It contains 9 layers: 5 layers of convolution and simplification functions made by 22 5x5 kernel filters and a max-pooling filter of 2x2 to reduce at last the input image of 32x32 into 16 maps of 5x5. The feature images carry the most significant features to define a specified traffic signs class by processing them into a 4 layers fully connected network.

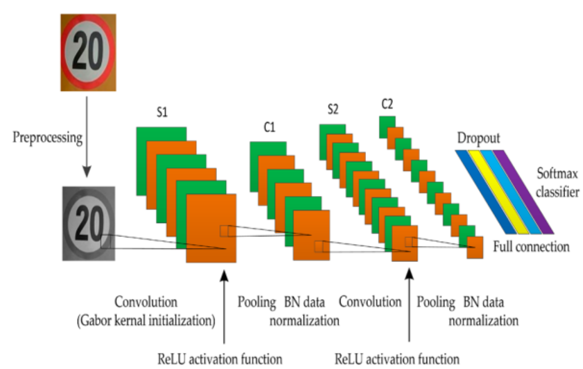

Fig. 5.1: CNN architecture

6. IMPLEMENTATION AND ANALYSIS OF THE CURRENT STRUCTURE

This section contains a descriptive overview of our two tier CNN architecture and is subdivided into 3 main categories.

- Training Data
- Target Detection / Object Detection
- Final results of the implementation

6.1 Training Data

The unbalanced distribution of images in the German Traffic Sign Benchmark privileges some classes over others during the training phase because they are better shown in the form of multiple images. To ensure that the network performs efficiently, some of these classes undergo a data augmentation by applying some geometric transformations (rotation, translation, and shear mapping) on many such images.

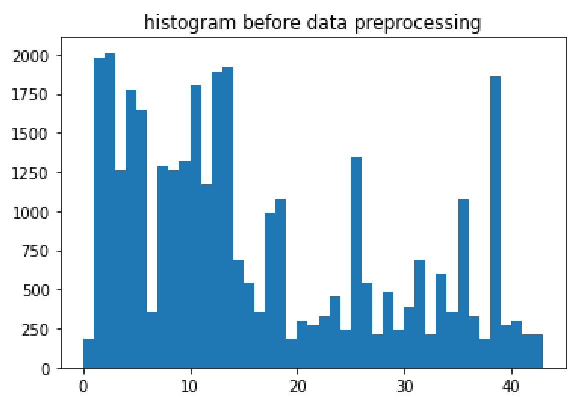

Fig 6.1: Distribution of Dataset

The algorithm takes only classes with less than 1000 images to randomly pick images and makes one of the transformation operations . The resulting images are added to the same class until the number of its elements reaches the bias i.e 1000 images.

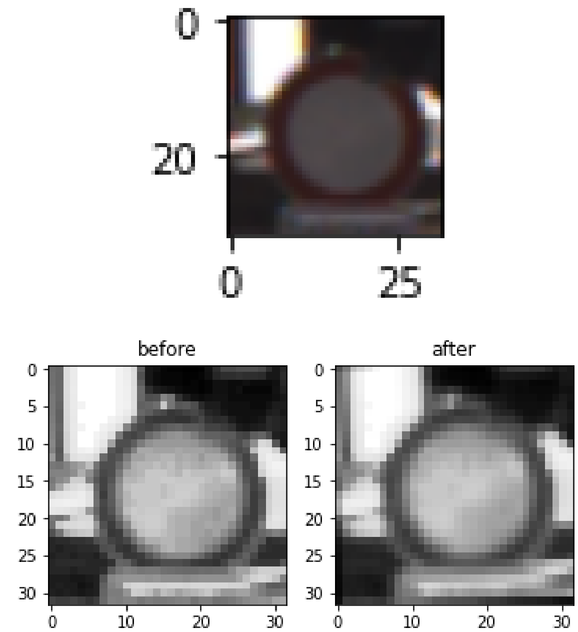

Fig 6.2: Comparison of images

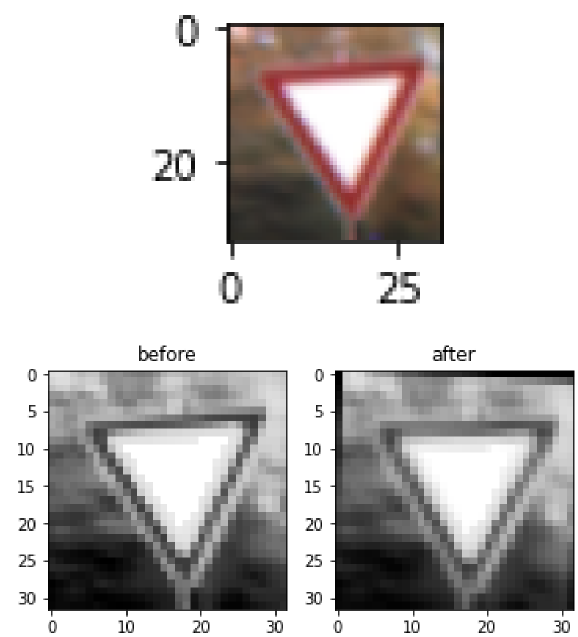

Fig 6.3: Comparison of images

This picture then undergoes multiple transformations like scaling, rotation and translations.

6.2. Target Detection

Our implementation employs YOLO networks to aid the Traffic sign assisted by a series of pre-trained cascaded images, which will provide real-time object detection. Object detection consists of various approaches such as fast R-CNN, Retina-Net, and Single-Shot MultiBox Detector (SSD). Although these approaches have solved the challenges of data limitation and modeling in object detection, they cannot detect objects in a single algorithm run.

Fig 6.4: Normal Image without object detection

Fig 6.5: Image after traffic sign has been detected

Fig 6.6: Traffic Sign detection system

6.3 Implementation Results

To build and train the network, the TensorFlow deep learning library is used. Training and testing were implemented using the dataset and the developed method succeeds in classifying the 43 traffic signs classes. The implementation results of the two tier CNN and its improvement operations show the impact of each changed element. The enrichment of the first layer of CNN fully connected network made the validation accuracy 95.2% after 100 iterations of the learning algorithm. The new given architecture can now combine many more factors to classify traffic signs. After applying data augmentation , an accuracy of 95.6% at the 100th iteration is noticed, making the network performances even better than the last ones. It is also due to the new balanced property of the training data in different classes.

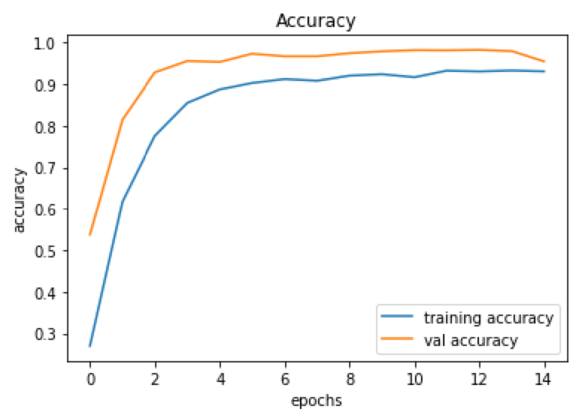

Fig 6.7: Accuracy variation over training time

As seen above, with regular training and testing and constantly learning from our previous iterations, we have come up with the following results, with an increasing rate of accuracy over time. Our implementation had a few minor setbacks and a few improvements, with the above graph showcasing CNN's accuracy over iterations.

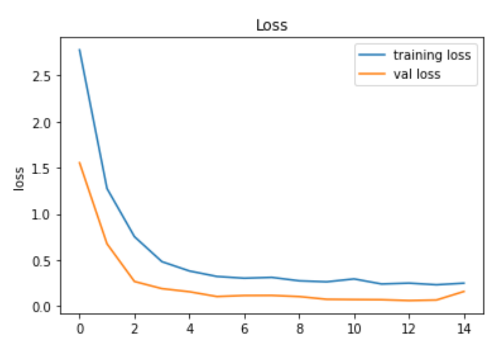

Fig 6.8: Data loss/ accuracy decrease over time taken for training

Over time, we learned from our mistakes, and minimized errors as much as we could. The above graph is a representation of the lows and error points we faced while we were constructing Neural networks.

7. CONCLUSION

Our final results have proved to be very useful with a minimum accuracy of 95.1% which we have extended to 96%. The two tier CNN implementations along with the addition of YOLO Networks have resulted in more accurate results even under versatile weather conditions. While the current implementation has proved to be a bit system heavy, it has also proved that it functions exceptionally well under adverse conditions with higher accuracies.

REFERENCE LINKS

[1] J. Zhang, Z. Xie, J. Sun, X. Zou, and J. Wang, "A cascaded RCNN with multiscale attention and imbalanced samples for traffic sign detection," IEEE Access, vol. 8, 2020.

Fast Traffic Sign Detection Approach Based on Lightweight Network and Multilayer Proposal Network.

[2] C. Cao, B. Wang, W.Zhang et al., "An improved faster R-CNN for small object detection," IEEE Access, vol. 7, 2019. (PDF) An Improved Faster R-CNN for Small Object Detection

[3] De Charette, R.; Nashashibi, F. Real time visual traffic lights recognition based on spot light detection and adaptive traffic lights templates. In: Proceedings of the IEEE Intelligent Vehicles Symposium, 358–363, 2009 (PDF) Real Time Visual Traffic Lights Recognition Based on Spot Light Detection and Adaptive Traffic Lights Templates

[4] Cai, Z.; Gu, M.; Li, Y. Real-time arrow traffic light recognition system for intelligent vehicles. In: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition, 1, 2012. [PDF] Real-time Arrow Traffic Light Recognition System for Intelligent Vehicle

[5] Ji, Y.; Yang, M.; Lu, Z.; Wang, C. Integrating visual selective attention model with HOG features for traffic light detection and recognition. In: Proceedings of the IEEE Intelligent Vehicles Symposium (IV), 280–285, 2015. (PDF) Integrating visual selective attention model with HOG features for traffic light detection and recognition

[6] Fairfield, N.; Urmson, C. Traffic light mapping and detection. In: Proceedings of the IEEE International Conference on Robotics and Automation, 5421–5426, 2011. [PDF] Traffic light mapping and detection

[7] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German traffic sign recognition benchmark: a multi-class classification competition," in Proc. IEEE IJCNN, 2011, pp. 1453–1460. [PDF] The German Traffic Sign Recognition Benchmark: A multi-class classification competition

[8] "Real-Time Detection and Recognition of Road Traffic Signs" by Jack Greenhalgh and Majid Mirmehdi, Senior Member, 2012, IEEE. Real-Time Detection and Recognition of Road Traffic Signs

[9] L. Abdi, "Deep learning traffic sign detection, recognition and augmentation," Proceedings of the Symposium on Applied Computing, Maroc, 2017, p. 131-136

[10] Y. Le Cun, L. Bottou, Y. Bengio and P.Haffner, "Gradient-Based learning applied to document recognition," Proceedings of IEEE, Vol. 86, N°11, p. 2278-2324, 1998.