An Efficient Deep Neural Network For Traffic Sign Classification in Autonomous Vehicles

MSc Research Project
Data Analytics

Aiswarian Sebastian Varghese
Student ID: 20207816

School of Computing
National College of Ireland

Supervisor: Mr. Hicham Rifai

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Aiswarian Sebastian Varghese
Student ID:	20207816
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Mr. Hicham Rifai
Submission Due Date:	15/08/2022
Project Title:	An Efficient Deep Neural Network For Traffic Sign Classification in Autonomous Vehicles
Word Count:	6691
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	15th August 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

An Efficient Deep Neural Network For Traffic Sign Classification in Autonomous Vehicles

Aiswarian Sebastian Varghese
20207816

Abstract

Intelligent transportation system is an emerging leap in the automobile sector, Classifying and recognizing the road and traffic signs are the supreme aspect for the companies for developing their own prototype in this industry. Convolutioal neural networks in deep learning is an fortunate method to attain accurate results in these image processing tasks.

Although Numerous approach by different researchers have achieved promising results in classifying and predicting traffic signs using traditional hand crafted methods, The proposed work focuses on developing a convolutional neural network classifier where this deep learning technique classifies all the traffic signs images with improved accuracy and precision. the performance of the model is further improved by effective fine tuning in the hyper parameters. The data points in the considered dataset contained the traffic sign images and markings in different environmental conditions and distinct lighting conditions. These signs are first preprocessed to identify its extract clinical information from it and then preprocessed images are transformed to identify the area of interest thus signs are accurately recognized using deep learning. We use an end to end open source platform tensor flow to implement CNN. The hyperparameter tuning of the model is performed using the Grid search CV technique and optimized value for the parameters are chosen for improved efficiency of the model. The performance of the model is evaluated using the performance metrics. The presented model gaid the overall accuracy of 98.44% using the deep neural network model.

1 Introduction

The efficiency of the image processing tasks relies when the model correctly identifies the region of interest from the given data point. The domains like image processing and speech recognition are advancing with a phenomenal potential with deep learning techniques. For image processing tasks the neural networks especially, convolutional neural networks show high robustness for the recognition and classification of image data. As we said significant role of image processing applications are to identify the area of interest from the data and the major challenge is to identify this significant information from the given content if it is very small. So, to be a model to be credible it is important to identify and classify the area of interest if it is not exposed in the content. Traffic signs plays a prominent role in attaining safety and discipline in roads for the vehicles as well as the pedestrians. Traffic signs plays a pivotal role in ensuring safety for vehicles especially during heavy traffic conditions, bad environmental weather conditions.

Many researchers approached this problem of classifying traffic signs with different traditional and handcrafted methods like histogram-oriented gradient (HOG), speeded up robust features and invariant scale feature transform. But still this domain remains challenging for them. With the advancement of deep learning and machine learning algorithms the models are trained by transfer learning processes where the accurate and precise information is gained from self and recursive study from the data through hidden layers in the neural network architecture.

The corpus of road traffic signs is categorized depending on their function and what they imply. Also, different traffic signs have particular shape in appearance and distinctive nature. The entire architecture of the presented work is carried out as shown in the figure below.

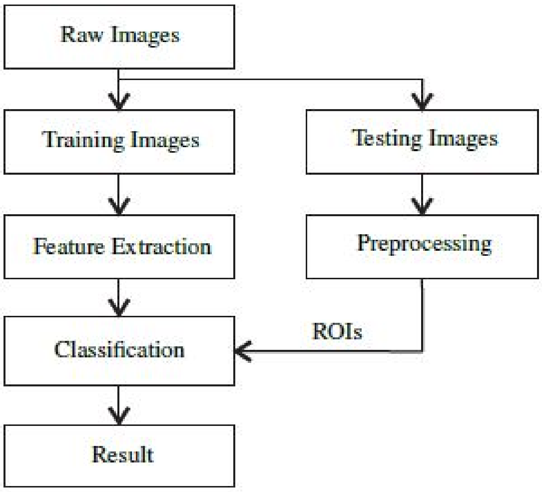

Figure 1: Block Diagram

The Presented work focuses on addressing the research question Given below.

"How an optimized deep convolutional neural network helps in classifying and recognizing Traffic sign boards irrespective of thier shape and color efficienly with higher accuracy than existing traditional models"

The main objectives of the proposed work is given below:

- Identifying the existing technologies in the proposed problem and exploring it for the development of the proposed work.
- Pre-processing the the GTSRB dataset which is taken from Kaggle repository using the python language.
- implementation of deep convolutional neural network for classifying the traffic sign.
- The performance of the model is improved by hyperparameter tuning using Grid Search CV.
- Evaluation of the dataset and why it is selected for building the traffic sign classification system.
- Identified the value of the presented problem and the performance is evaluated using performance metrics.

2 Related Work

Intelligent driving system domain is advancing with phenomenal potential every day. But there are still some concerns resides in the phase of classification and recognition of traffic signs in autonomous vehicles. Many researchers have presented different solutions to overcome this problem. The existing works depicts that the accuracy in relied in the discriminatory features in the image and the use of generic classifier for building the model. There are also different traditional and hand crafted techniques used to counter this problem like histogram oriented gradient (HOG), SURF, SIFT techniques. In the presented work we build a deep learning model using convolutional neural network for accurately classify and recognize the target data..

2.1 Color Segmentation Methods in Traffic Light Classification

One of the common traditional method in the domain of traffic sign classification is color segmentation method. (Chen and Lu; June 2016) developed a model where the salience of image is calculated with itti Koch method here the image is passed and converted to single vector through gradient transformation and then the vector is passed to series of decision trees for the image classification. The performance is evaluated by using SVR.

One of the major issues researchers faced during the traffic sig image classification is how they can accurately estimate the boundary of the signboard and extract information and (Lee and Kim; May 2018) stated that this is one of the major challenge in the intelligent driving system. (Lee and Kim; May 2018) he suggested that performing more preprocessing like contour estimation and image segmentation can improve the efficiency of the models. His work focused on predicting the traffic signs from a 2-d pose traffic sign data, even though it achieved 68 % accuracy the model failed to classify or predict images in poor light condition.

Taking this drawback into consideration (Saini and Nikhil S; 2017) developed model in 2018. According to his point of view the performance of the model in the bad environment conditions as an critical issue that may lead to safety concerns in the autonomous vehicles. He mentioned that this as a major challenge in traffic sign classification system. For this (Saini and Nikhil S; 2017) vision based three stage traffic sign recognition system. In the initial phase the input image is preprocessed and the extraction of the information is done by HSV image segment method. After that the image is clipped to sufficient aspect ratio. This clipping of image is dine to avoid the noises that may lead to false results and high computational time. Now for the challenge he mentioned he used a clustering technique here the maximum external region technique to identify the boundary and the structure of the traffic signals. The MSER technique used in this model actually convert the frames to binary values and then the point interest is filtered out using these binary levels.

After the area of interest is identified then it is passed to Histogram Oriented Gradient technique for the giving color gradient to the data, this is done to accurately identify the boundary and neglect the false positive that will be predicted by the system. In the detection phase Support vector machines are used to recognize the traffic signs. (Saini and Nikhil S; 2017) The model achieved 95 % performance which is pretty good the model only recognize only specific traffic signals and also the angle of the data when it is captured showed up some difficulties for the proposed technique.

A traffic sign differ each other in its specific shape and the color of the sign board. In

2016 Ma Xing and Wang Yan proposed the problem where the recognition of the traffic sign using the Zernike moment and image standardization (Ma Xing and Yinchuan; 2016). Here in this work the input images are segmented into its components, by doing this they achieved a good computational speed then the HOG descriptors are used to identify the design and contour of the traffic sign then the classification I done using SVM. Now the model is tested and the images are recognized using the Zernike moment. The model performed for the data that as captured in different environment circumstances but the system requirements were very high due to high computational time.

2.2 Optimizing color and shape Segmentation

High computational speed and detection speed in traffic sign system are major concern in the intelligent driving industry. Because if the autonomous cars fails for the quick response it may lead to accidents on the roads. Taking this situation into consideration (Handoko; February 2021) developed model to improve the processing speed using the shape and color segmentation. In the proposed model the images that are passed to the models are images that having low frame rate, and this frame rate is further uplifted by passing to the AI system. HSV threshold technique is used for the segmentation of the frame filter. By using this technique the saturation of the filters are neglected. The performance of the model was really good with 88% is achieved with the training set and also 97% of the frames were filtered out. This is done using Gaussian Blur matrix and open kernel techniques

(Rajesh; August 2011) developed a model using Coherence vector of oriented gradients (CVOG) technique for recognizing the traffic signs efficiently. The basic idea behind CVOG technique is to generate a matrix from the input features. That is as we give the input to the model the gradient of the data is extracted from transpose of the horizontal filters over the y axis. By transforming the input data the new gradient matrix is derived for the input data. From this derived matrix the area of interest is identified by examining the pixels and particular gradient. The coherence of the image is calculated by the variable tilde. After identifying the area of interest the matrix is then passed to the SVM classifier for building the model. 80 % of accuracy is achieved and it can be fine-tuned with by using different combination in features.

Due to huge advancement in the deep neural networks the image processing models uses these neural network for the classification tasks. Training the model in the deep learning is challenging. (Kaiming He; 2016) developed a system for image processing systems in deep neural networks where the images are trained faster than the existing techniques. The input data to the model is first transformed into its residual functions. This transformation brings more efficiency and the training of the images will be easier. For the development (Kaiming He; 2016) uses ImageNet dataset. So for the processing the residual matrix is passed instead of raw data.

for the data points that is captured in wide areas the depth of the neural network has a prominent role in determining the area of interest from the huge canvas.

2.3 Role of Artificial Intelligence

For any development of the model a good literature review should be carried out to identify the latest technology in the domain. (Saritha and Kumar; October 2020) in her paper where it talks about the role of the artificial intelligence in image processing system. The thesis consists of three parts where the image processing systems should improve for attaining efficient and better recognition. The paper also talks about how the recognition systems can be improved for visually impaired people. Autobay of dash cam , RINGO and UVG Driver assistance are the tools that are presented in the paper. The paper further discuss the relevance and the challenges that are faced by the traffic sign detection system in the modern world.

Recognition and classifying the input images are done by transforming the images through different preprocessing stages like Segmentation, scaling etc also failure in any of the processing stage may lead to severe outcomes.

(Yeh and Lin; 2020) developed a traffic sign recognition system for the Taiwan roads. The model was built using the convolutional neural networks and the model tried to classify and recognize the shape and color state from the input image. The model was built to implement in the urban areas of the city. For the development of the model the LISA dataset as used. In the detection stage two cameras was used to capture images from different distances. The features was shared in this end to end neural network. The implemented model performed with an 85 % accuracy but the problem was the model failed to identify traffic signals arrows in it.

Due to high training and processing time the researchers dive into deep neural networks for building the models for classification and recognition problems as those models possess to have less learning rate and low computational time. (Krizhevsky et al.; 2012) finds that Image Net dataset as one of the most complicated dataset and he decided to build a CNN model with the dataset. This research got an wide appreciation for his work. (Krizhevsky et al.; 2012) wants to develop a model for image processing tasks with least error rates. He used top 5 error rate grading rubric to validate the model for the five predictions. The framework showcases the advantages of building model with by implementing data preprocessing like data augmentation and adding dropout layers. He proved that that efficiency of the model drops whenever a layer is removed from the model.

In 2018 (Perrson; 2018) developed a VGG16 system here it is an pertained CNN for traffic sign classification problem the major issue the (Perrson; 2018) faced was the lack in the training data points hile building the model the accuracy of the model as low due to this reason.

2.4 Inception Module Method

The growth of the automobile mobile in this age demands the need for the intelligent driving systems. So it is vital to develop efficient models that helps in autonomous driving systems. (Zhao Dongfang; November 2019) presented a model that uses the ImageNet data for the classification and recognition of the traffic lights. The proposed model is built with CNN architecture and adding normalizing layer. The normalizing layer is added to avoid the overfitting of the model. The ROI of the data is extracted using Hebbain Principle. The entire system is built with a sparse structure for improving the performance of the model. The proposed VNN model was efficient enough for the classification and recognition in the image processing tasks.

Embedded systems are developed by different manufacturers from early 2015. (Manato Hirabayashia; 2018) developed system which is built on the top of robot operating systems and autoware for recognizing th state of color in the traffic light systems. The autoware technology in the proposed study is responsible for the perception and sensing and also managing the modules in the autonomous driving cars. For the work (Manato Hirabayashia; 2018) used digital camera and LiDAR sensor where it emits the UV rays to recognize and identify the obstacle. The data for the model is acquired from this mapping technique.

The automobile companies are in a verge of competition to implement latest technology in their products. (Rao and Desaib; July 2021) researched about this problem in his thesis how these advance technologies helps in the development of the industry. From his study he states that the the openCv module in the deep learning along with ML embedded sensors will be big leap in the classification and the recognition of the traffic sign systems. Also he added in the scope of attaching IR proximity sensors embedded cameras along with Arduino module for the real time implication in the autonomous cars. The IR proximity sensors can be acquire real time data of the traffic signs by emitting IR waves and this captured images can be given to the model for classification and recognize these models through openCV model and the vehicle itself can take necessary action on roads. (Rao and Desaib; July 2021) in the presented work uses Hough Gradient Technology for the segmentation of the frames in the input data for finding out the region of interest.

2.5 Stereo vision method

Sterio vision technology was implemented by (Andreas Fregin and Dietmayer; June 2017) in 2017 for classifying and recognizing the traffic signs. The author proposed multiple ways to implement this technique. The first method was to identify and filter out the ROI by evaluating the disparity value of the input data. Then the second method was to identify the relational position of the object. This is achieved by the computing the extrinsic and intrinsic value of the image calibration to know the distance between the vehicle and the obstacle. This method made a large impact by filtering 30% to 70% false positive of the geographic locations. The efficiency of the three methods was calculated by evaluating the standard deviation and outliers of these disparity data.

2.6 Deep Learning Methods in Classification

(Islam and Raj; April 2017) presented a model that classifies the road traffic signs for the Malaysian road networks. For the developed work (Islam and Raj; April 2017) has taken information or taken the data dynamically by placing the digital camera over the dash of the moving vehicle. The paper states that this as t the first time an artificial neural network as implemented for the traffic classification domain. The extracted features are identified using color segmentation and then it is parsed through ANN network. The testing of the model has also done with Malaysian traffic sign data (LISA). The model possessed an accuracy of 93% the model failed to classify and recognize the different shaped traffic signs and poor performance in recognizing road signs.

2.7 Using Attention Model

The updation in the traffic sign detection system are developing day by day. In 2018 (Yifan Lu1 and Hall; September 2018) proposed for implementing real time traffic sign recognition model here the ROI of the images are extracted from the street vie images. Using the attention model. Attention models are used to identify the region of interests from small contexts. The model is implemented and tested using TT100K dataset. The model possess 90% of accuracy but the major drawback was it only classified and recognized traffic light images.

A strategic change has been made in the traffic signal detection system by the rise of deep learning techniques which was a challenging domain for years. According to (Yuan Yuan and Wang; February 2019) The major challenges of this system was mainly categorized into two. First one mentioned in the (Yuan Yuan and Wang; February 2019) was the identification of the smaller traffic signs from a panoramic view of a particular street and the second challenge that Yuan Yuan mentioned was the resemblance of the traffic signs from road view where exact context information is not given. In 2019 they proposed two methods to overcome these challenges. First method was the they created highly connected deep convolution layers was implemented with multi resolution feature network. Secondly the proposed a vertical spatial sequence model to retrieve the context information for the efficient detection of the system. The evaluation of the proposed model was tested on 3 different data sets.

Capturing the ROI from the panoramic street views are the challenging part in the traffic sign recognition system. (Jianming zhang; January 2020) approached this problem by identifying the ROI in illumination variation method and how can we classify and recognize the traffic signs in bad weather conditions. The work as mainly focused on pyramidical shaped traffic signs and the presented work as implemented using cascaded RCNN. Along with this model (Jianming zhang; January 2020) implemented a multiscale attention model was also implemented to improve the accuracy. The model achieved 83.68% with the CCTSDb benchmark.

Efficient version need to be solid enough to locate register day and night situations. it'll be hard to discover the place of the visitors mild in night time conditions due to the smaller size.(Phuc Manh Nguyen; October 2020) finds this as challenging and proposed a model for effectively classifying the traffic sign in different daylight conditions by incorporating traditional deep learning and hand crafted techniques. The RCNN model was used for the detection of the signs and they implemented and tested the model with CCD dataset. The model achieved 80% performance which as better than color based model.

2.8 Conclusion

By going through the above related work from different researchers a meaningful literature survey was done before successful implementation of the presented work. From the survey we came to know that there are many publicly and privately available datasets that was used for the problem solving in this particular domain so identifying the correct dataset and with large samples of the traffic sign images are the most important phase in performance of the model. Also the dataset should contain the data points with different angles and also there should be images captured in different environmental conditions to improve the efficiency of the presented model. The presented work chooses to implement the problem with deep convolutional neural network as this method is more suitable for the road traffic signs.

3 Methodology

For the development of the proposed work the overall processes like data mining and knowledge discovery from the data is done by following the steps in the KDD methodology. An outline of these KDD process is shown in the figure below.

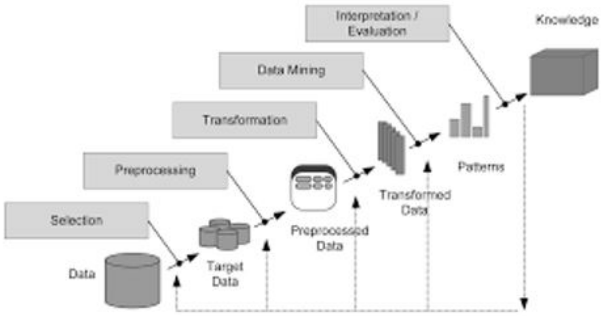

Figure 2: KDD Process

For improving the precision and accuracy in classifying and recognizing traffic signs we implemented Convolutional neural network for the development of our model. Essentially the images of the traffic signs are preprocessed and trained to classify and recognize these images accurately and precisely by giving these images as input to the model. After the training we check the efficiency of the model by passing the test dataset and evaluating whether the model is capable of identifying the traffic signs which in return helps the autonomous cars can understand. During the preprocessing stage the every data point in the GTSRB dataset is converted into gray scale following the Histogram equalization and then normalizing the dataset. Also to avoid overfitting of the model data augmentation is also performed in the after the preprocessing stage by expanding the training data. After that we implemented and trained the CNN model with the GTSRB dataset.

3.1 Data Acqusiation

Several road traffic signs datasets are available publically which is published by different countries like USA, Belgium, Germany, Croatia and china. The presented work is developed by considering German Traffic Sign Recognition Benchmark (GTSRB) which is taken from public Kaggle repository.

3.1.1 German Traffic Sign Recognition Benchmark

GTSRB Dataset is a dynamically maintained globally accepted dataset for the traffic sign classification. After the meaningful literature survey lots of researchers have considered this dataset for comparing different traffic recognition systems around the world. The dataset is contributed and tested by different scientists of several domain and it is approved in the GTSRB conference conducted by its authors. The dataset is updated with all the latest updations with new data points. The captured images in this data set is extracted from 1 second video segments of the traffic conditions from a moving vehicle that is recorded in different environmental and lighting conditions. The dataset consists of a total of 43 sample traffic sign classes with different distortion and resolution of images in each class with different class frequency. The entire dataset is divided into training and test data with 34,799 and 12,630 samples respectively. This division in dataset is given to check the efficiency of the model. The resolutions in the sample ranges from 15x15 to 250x250 pixels so we normalize this resolution into 48x48 pixels so as to determine the region of interest and processing of image.

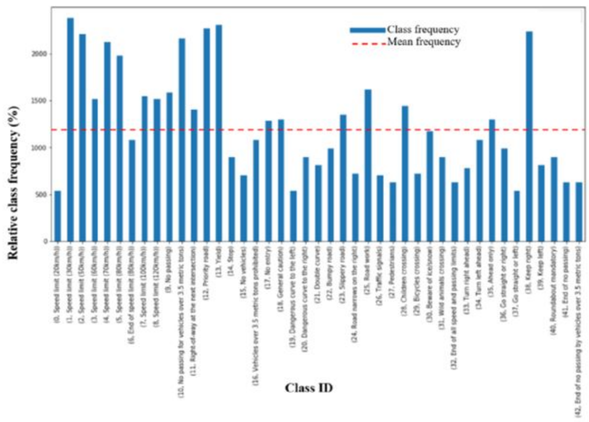

Figure 3: Dataset Description

3.2 Data Preprocessing

The presented model classifies traffic signs where the model is trained using the GTSRB dataset. This dataset consists of different classes of traffic sign images. There are in total of 47,234 samples of images and the entire dataset is divided into training test and validation set. The test set and validation set consists of 12630 and 4410 images respectively. As we said there are 43 classes and the in each classes there are sample image with an imbalance frequency. As mentioned above the images are captured from 1 min video sequences captured by digital camera by placing it on car dash board under different lighting conditions. The images in the dataset possess low resolution and low contrast. So the dataset should be passed through the preprocessing stage before implementing the model. There are different signs like stop the vehicle, yield, railway line ahead etc. signs in this benchmark. The preprocessing stage of the benchmark consists of three phases.

- Conversion of image to gray scale
- Histogram equalization
- Normalization of image

In the image processing tasks the preprocessing plays a vital role. That is the images that are captured in the dataset will be in a colored or in the RGB format which increase the computational time and speed so for better processing of the images the data should be converted into gray scale images as the response time increases there is high chances of occurring accidents on the road. The conversion is done without losing any details of the image.

After the conversion to grayscale we mentioned that because images are captured while moving the vehicle the pictures may have low contrast. So to regain its shape and details the histogram gradient is used by applying histogram equalization. The basic idea of the histogram equalization is to enhance the contrast by decreasing number of gray scales in the image. This equalizing technique joins two neighboring gray levels into high intense single gray level hence the image possess equal gray distribution property.

The datapoints in the dataset will have different pixel rates so it is inevitable to standardize these pixel values so normalization is a technique where the image pixels are normalized.

```
img = equalize(img)
plt.imshow(img, cmap = 'gray')
plt.axis('off')
print(img.shape)
```

Figure 4: Histogram equalization

3.2.1 Data Augmentation

The 43 classes in the dataset has different class frequencies of data in it, so this situation may lead to overfitting of the model due to less datapoints available. So the dataset can be enlarged artificially by applying data augmentation technique. This technique is applied to increase the nmber of the sample size of the image in the dataset by transformations like rotation, flipping and zooming the images. We can use data augmentation by importing ImageDataGenerator function in Keras

4 Design Specification

The entire data points in the GTSRB dataset consist of different classes of traffic signs in RGB format. The entire architecture of the presented work can be divided into these phases. Preprocessing stage, implementation stage and classifying and recognizing the traffic signs. Preprocessing stage is an important phase in development where we clean data for improving the efficiency of the model. During this stage the images in the dataset is converted into grayscale then we normalize the data, also we adjust contrast of the data using image's histogram. The model complexity increases if we process raw RGB images so to avoid this complexity and time consumption we are converting the images into grayscale. The images possess having high bit size is lowered to 8 bits without losing any information from the image. The data points consists of images with different contrasts, this situation is dealt where the contrast is increased by diminishing gray levels with uniformly distributing it by the property of gray distribution of the image by performing histogram equalization.

The images may have different pixel values which may cause problem during the processing. To overcome this situation, we normalize the data points by applying statistical distribution. Data augmentation is also done as the samples in each class in the dataset is of different frequencies. So as to avoid overfitting of the model data augmentation is done in the presented work using Image Data Generator function which is available in Keras library.

The processed data is now trained and used for classification of traffic sign using convolutional neural network. The model is built in google collaboratory using python language and the output and results are examined and evaluated.

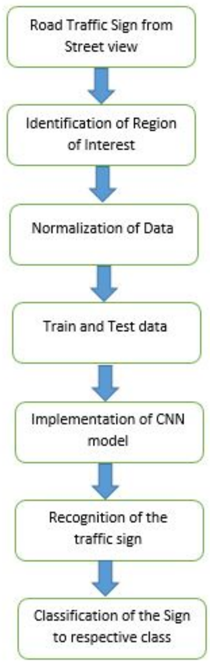

Figure 5: System Architecture

5 Implementation

After the preprocessing phase, now the 47,234 traffic sign images are normalized and ready for the model application. As we mentioned in the initial sages the data is divided into both training and validation set for identifying the efficiency and the model accuracy that is 34,799 images are used for training the model and 12,630 traffic sign images are used for validate the model. To avoid any biases the split in the dataset is done using the function random sampling. After the meaningful literature survey the presented work is a sequential model that is developed using Keras. Where this open source platform is highly prone for building blocks and creating powerful deep neural networks.

5.1 Convolutional neural networks

As the proposed work deals with the images and process the pixel data we are implementing convolutional neural network which is an artificial neural network for this image recognition work. Convolutional neural network architecture have different layers, the major layers in CNN are listed below.

- Input Layer
- Hidden Layer
- Output Layer

The data points are traversed through each layers and the parameters in each layers helps in reducing the error in classification of the image to its appropriate class. The images are either down sampled or up sampled for the easiness in the computation of the data. The neural network actually mimics the functionalities of the neurons in the human brain. These layers presented in the convolutional neural networks are placed in such a way that it can grab every visual field in the image by avoiding piecemeal image processing problem which is caused by other artificial neural networks. The hidden layer in the CNN network consists of different convolutional layers, Fully Connected layers, pooling layers and normalization layer. The convolutional layer is the core part of the neural network where all the computational works are accurately done by using different kernel functions. Also extraction of the information from the data is done at this layer and it is passed to tensors by raw pixels and then the fully connected layers classify the particular data into its class. The architecture of the workflow in the convolutional neural network is illustrated in the figure below.

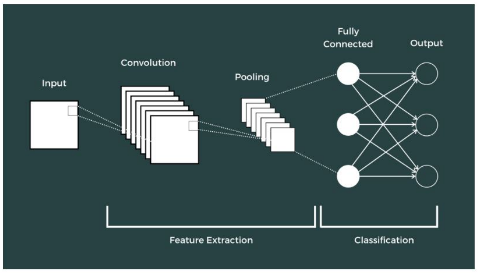

Figure 6: CNN Architecture

5.2 Activation Function

Activation functions are the mathematical gateway where it defines hat weighted sum of neuron should be transformed from one layer to another. There are multiple activation function choices and choosing appropriate one will improve the performance and capability of the neural network

5.2.1 ReLu Activation Function

This is the rectified linear activation function and this is the most common activation function used in the hidden layers. Calculation of ReLu function is as follows: Max(0.0,x), That is 0.0 is returned if the value is negative else the value is returned.

5.2.2 Sigmoid Function

This is a logistic function where it takes any real input values and outputs within the range of 0 to 1 It is calculated as 1.0(1.0 +e ˆ-x)

5.2.3 Tanh Activation Function

This is hyperbolic tangent activation function that takes any real number as input and the outputs the value in the range of -1 to 1. It is calculated as tanh = (exˆ – eˆ-x) / (exˆ + eˆ-x)

After we apply the activation functions to the data then we move to max pooling layers where it is responsible for identifying the region of interest and extracting the information from the data points by eliminating number of parameters from the data points. Max pooling layers enhances the generalization of the image by identifying the invariant features hence improve the performance of the model. Drop out layers are added to avoid the overfitting of the model that is drop out layers are added to the final max pooling layer and the neurons that are dropped by this layer will not propagate through the final layers. Following the max pooling layer the architecture places the fully connected layers where the exact classification of the image takes place.

5.3 Hyperparameter Tuning

Hyper parameter tuning is very important in the machine learning models to improve the performance and the efficiency of the models. by optimizing the the weights of the parameters we get the best evatuation metrics during the computation. Grid Search cross validation can be used for optimizing values in any machine learning model.Grid search cross validation is a library which is available in scikit learn to optimize the parameters and compare the results. Also for implementing Grid Search CV scikit learn uses rappers API for supporting optimizations for the sequential model.By performing this hyper parameter tuning we gets the best performing combinations among the test and train set.

EPOCH	50
BATCH SIZE	128
LOSS FUNCTION	categorical_crossentropy
OPTIMIZER	Adam
ACTIVATION FUNCTIONS	RELU & SOFTMAX
DROPOUT	20% & 50%

Figure 7: Model Hyperparameters

5.4 Evaluation

In this phase we discuss how our model is evaluated by the performance of the model. This metrics helps us to identify the dynamics of our model with metrcis like accuracy, precision, confusion matrix, f1 score etc. The most important instrument is the confusion matrix where we can actually find how much our system is performing with true or false predictions. all other performance metrics are derived from the confusion matrix. (Ashwitha A; October 2020)

- Accuracy :This implies the total true or correct prediction from total number of predictions made by system.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision: Precision is defined as, it quantifies the number of true predictions that belongs to the total true positives.

$$Precision = \frac{TP}{TP + FP}$$

- Recall: Precision is defined as, it quantifies the number of true predictions from the positive class.

$$Recall = \frac{TP}{TP + FN}$$

5.5 Convolutional Neural Network Model

The presented model is a sequential model that is built on Keras to classify and recognize the traffic signs for autonomous vehicles to ensure safety and integrity. Keras is a high end API which is developed by google for building deep neural networks. It is an open-source library that works on top of Tensor flow. In the sequential model the data points are reliant each other because of their sequential order and they are not identically distributed. In the implemented work the developed traffic sign convolutional neural network comprises of four convolutional layers, two maxpooling layers, four fully connected layers and one flatten layer. The summary of the model is shown in the figure below.

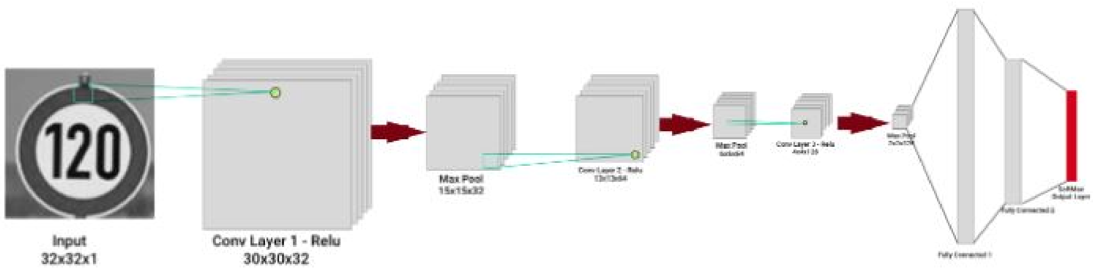

Figure 8: Traffic Sign Classification using CNN

The images of different traffic signs are given to the model after preprocessing and reshaping the data into 32*32*1. So this reshaped images are given as the input to our model. The entire model comprises of four convolutional layers and the kernel size for the first two convolutional layers is 5*5. Where these filters are used to extract the information from the image data. And the batch size, the number of samples processed id is given 128. After the relevant literature study the activation function used for the model building is ReLu activation function, where it trains faster and is suitable for the classification of images problem. ReLu activation function is applied to the model in each convolutional layer. After the data is passed to the max pooling layer where it is used to identify most important features of the data. Max pooling clips the image to lower size without losing its information and generating pixels that contains the information about data. From the figure 1 we can see that the max pooling the parameter value is 0. The ReLu activation function and maxpooling layers helps to discover patterns and or it selects superior features from the data. Now e passes to the third and fourth convolutional layer in the model, the kernel size is given as 3*3 and the batch size is reduced to 15. Also the second max-pooling layer is added with the kernel size of 2*2. We added a dropout layer after the max pooling layers where the neuros are randomly drops during the training of the model. The major function of adding the dropout layer is to avoid overfitting of the model. The fully connected layers are implemented for the classification of the images. Also, there are hidden nodes in these fully connected and are 256, 128, 64 respectively, as there are 43 classes of traffic signs in total the number of hidden layers in the last output will contain will be 43. Dense layers are normal layers of neurons and it is deeply interconnected to its last layer that is it will accept input from all the previous layer. Along the dense layer 50% of drop out is also added so as to make the pixels more informative. As our work is a multi nominal classifier, we add a SoftMax layer to last dense layer basically software layer is an activation function for a system that classifies into more than two labels. So, in this work SoftMax layer will have 43 neurons as the dataset contains 43 different classes. So, by this deep learning approach the class which have higher probability will be the output for the input data. Hyper parameter tuning is an important phase in building deep neural networks as we have chosen Adam optimizer for optimization of the model. Adam optimizer is an alternative for Gradient descent in training deep neural networks. As our dataset contains large number of data points Adam optimizer is efficient for this problem, also memory requirement of this optimizes is very low. The loss function is handled by categorical cross entropy as our output contains more than two classes.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 28, 28, 60)	1560
conv2d_1 (Conv2D)	(None, 24, 24, 60)	90060
max_pooling2d (MaxPooling2D)	(None, 12, 12, 60)	0
conv2d_2 (Conv2D)	(None, 10, 10, 30)	16230
conv2d_3 (Conv2D)	(None, 8, 8, 30)	8130
max_pooling2d_1 (MaxPooling2D)	(None, 4, 4, 30)	0
flatten (Flatten)	(None, 480)	0
dense (Dense)	(None, 500)	240500
dropout (Dropout)	(None, 500)	0
dense_1 (Dense)	(None, 43)	21543

Total params: 378,023
Trainable params: 378,023
Non-trainable params: 0

Figure 9: CNN Layers

6 Discussion

Traffic sign classification system is a multiclass classification system that accurately classify and recognize upon the given traffic sign. The entire model is built in open source Keras library. The number of epochs for the model is set as 50 after efficient hyper parameter tuning. The number of epochs defines how many times the training dataset is propagate both forward and backward through the neural networks. For the presented work we tried different number of epochs. Best results are achieved when we give 15 and 50 epochs for the neural network. The figure below shows the model accuracy as we increase the number of epochs. We could see that as we increase the number of epochs the model accuracy also increases. But as the number of epochs changes the weights are changed and it have a huge prone to go into an overfitting curve. Also, the batch size given for the work is optimized to 15 that is total of 15 raining samples are run in a single batch. Also, the model loss vs epoch plot is also illustrated and we can clearly see the loss of model is decreased with epoch.

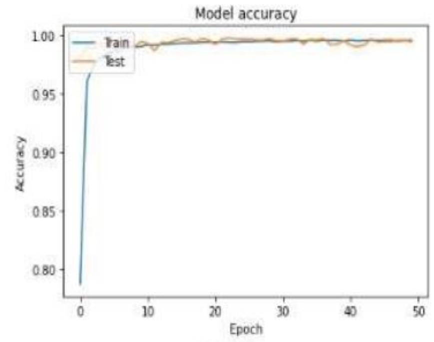

Figure 10: CNN Layers

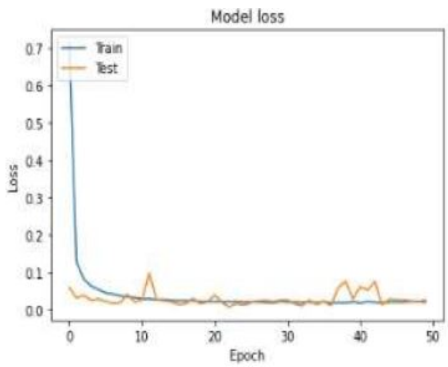

Figure 11: CNN Layers

7 Conclusion and Future Work

There is a phenomenal advancement in the area of artificial intelligence day by day, slowly the deep learning systems are taking control over the almost every domains. The presented work focuses on how we can classify the traffic signs in the road accurately and precisely than other existing technologies for the safety and integrity of the autonomous driving cars. All the competitors in the automobile industry are in a trying to embed latest technology for their products. The main challenge in implementing deep learning models in the sector is very risk that any minute problem in any of these machine learning can cause severe cause for the passengers and also for the pedestrians and animals on the road.

For efficient working of the model the preliminary thing that we need to have is the efficient dataset. For the presented work e chose the German Traffic sign Recognition benchmark, where the dataset is globally proven by many scientists all over the domain. The dataset includes more than 47,234 traffic image samples and also the set itself have divided into three for testing the effectiveness of the system. Every companies in automobile sector are in a clinch of implementing newest technology. The system has a very deep neural network architecture with four convolution layers, to max pooling layers, four fully connected layers so the system can effectively classify all the images of almost all the traffic sign images and road markings that is given as the input to the model. By implementing this model e try to ensure a good driving experience and safety in the autonomous cars. The presented model has achieved 98.44 % of accuracy with the convolutional neural network. That is this accuracy of the model is very good compared to the models with traditional and hand crafted features.

Considering the future scope as our system just classifies and recognize the traffic images statically, considering the huge development and future scope in Artificial intelligence, we are planning to develop an AI embedded Arduino system fixing in the cars that classifies and recognize the traffic signs in real time scenarios. The future scope of the work is also expected to be efficient using the new updating in the convolutional neural networks.

8 Acknowledgment

The Author shows a huge gratitude towards Mr.Hicham Rifai for his remarkable help and assistance for the completion of the presented work. His proficiency in the field helped in refining in both technical and documentation front.

References

Andreas Fregin, J. M. and Dietmayer, K. (June 2017). Three ways of using stereo vision for traffic light recognition, *2017 IEEE Intelligent Vehicles Symposium (IV),Redondo Beach, CA, USA* (17066180).

Ashwitha A, Shruthi G, S. H. R. M. U. A. P. R. M. T. C. (October 2020). Sarcasm detection in natural language processing, *Materials Today: Proceedings 37* .

Chen, T. and Lu, S. (June 2016). Accurate and efficient traffic sign detection using discriminative adaboost and support vector regression, *IEEE Transactions on Vehicular Technology* **65**(6): 4006 – 4015.

Handoko, Jehoshua Hanky Pratama, B. Y. (February 2021). Traffic sign detection optimization using color and shape segmentation as pre-processing system, *TELKOMNIKA Telecommunication, Computing, Electronics and Control* **19**(1): 173 – 181.

Islam, K. T. and Raj, R. G. (April 2017). Real-time (vision-based) road sign recognition using an artificial neural network, *www.mdpi.com/journal/sensors* (28406471).

Jianming zhang, Zhipeng xie, J. s. X. z. J. w. (January 2020). A cascaded r-cnn with multi-scale attention and imbalanced samples for traffic sign detection, *IEEE Transactions on Image processing* **8**: 29742 – 29754.

Kaiming He, Xiangyu Zhang, S. R. J. S. (2016). Deep residual learning for image recognition, *IEEE Conference on Computer Vision and Pattern Recognition* pp. 770 – 778.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems* **25**: 1097– 1105.

Lee, H. S. and Kim, K. (May 2018). Simultaneous traffic sign detection and boundary estimation using convolutional neural network, *IEEE Transactions on intelligent transportation systems* **19**(5): 1652 – 1663.

Ma Xing, Mu Chunyang, W. Y. and Yinchuan (2016). Traffic sign detection and recognition using color standardization and zernike moments, *Chinese Control and Decision Conference (CCDC)* .

Manato Hirabayashia, Adi Sujiwoa, A. M. S. K. M. E. (2018). Traffic light recognition using high-definition map features, *Robotics and Autonomous Systems* .

Perrson, S. (2018). Application of the german traffic sign recognition benchmark on the vgg16 network using transfer learning and bottleneck features in keras, *Teknisk-naturvetenskaplig fakultet UTH-enheten* .

Phuc Manh Nguyen, Vu Cong Nguyen, S. N. N. L. M. T. D. H. X. N. V. D. N. (October 2020). Robust traffic light detection and classification under day and night conditions, *20th International Conference on Control, Automation and Systems (ICCAS 2020)* pp. 565 – 578.

Rajesh, Rajeev, S. L. G. R. (August 2011). Coherence vector of oriented gradients for traffic sign recognition using neural networks, *The 2011 International Joint Conference on Neural Networks* pp. 907 – 911.

Rao, S. S. and Desaib, S. R. (July 2021). Machine learning based traffic light detection and ir sensor based proximity sensing for autonomous cars, *2nd International Conference on IoT Based Control Networks and Intelligent Systems* .

Saini, S. and Nikhil S, Krishna Reddy Konda, H. S. B. G. N. (2017). An efficient vision-based traffic light detection and state recognition for autonomous vehicles, *2017 IEEE Intelligent Vehicles Symposium (IV)* **17**: 11–15.

Saritha and Kumar, D. A. (October 2020). The role of artificial intelligence in automatic traffic light detection system, *International Journal of Scientific Technology research* **9**(10): 4006 – 4015.

Yeh, T.-W. and Lin, H.-Y. (2020). Detection and recognition of arrow traffic signals using a two-stage neural network structure, *6th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2020)* pp. 322–330.

Yifan Lu1, Jiaming Lu, S. Z. and Hall, P. (September 2018). Traffic signal detection and classification in street views using an attention model, *Computational Visual Media* **4**(3): 253–266.

Yuan Yuan, Z. X. and Wang, Q. (February 2019). Vssa-net: Vertical spatial sequence attention network for traffic sign detection, *IEEE Transactions on Image processing* **28**(7): 3423 – 3434.

Zhao Dongfang, Kang Wenjing, L. T. L. G. (November 2019). Traffic sign classification network using inception module, *14th IEEE International Conference on Electronic Measurement Instruments* .