A Deep Learning Framework to Traffic Sign Recognition in All Weather Conditions

MSc Research Project
MSc in Data Analytics (Group B)

Aniket Piralkar
Student ID: 20177291

School of Computing
National College of Ireland

Supervisor: Dr. Musfira Jilani, Dr. Paul Stynes, Dr. Pramod Pathak

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Aniket Piralkar
Student ID:	20177291
Programme:	MSc in Data Analytics (Group B)
Year:	2021
Module:	MSc Research Project
Supervisor:	Dr. Musfira Jilani, Dr. Paul Stynes, Dr. Pramod Pathak
Submission Due Date:	15/08/2022
Project Title:	A Deep Learning Framework to Traffic Sign Recognition in All Weather Conditions
Word Count:	8362
Page Count:	20

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:	14th September 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):	

A Deep Learning Framework to Traffic Sign Recognition in All Weather Conditions

Aniket Piralkar
20177291

Abstract

Traffic Sign Recognition (TSR) for Advanced Driver Assistance Systems (ADAS) is crucial for preventing fatal crashes. However, recognizing the traffic signs in adverse weather and lighting conditions is a challenge. The aim of this research is to develop a deep learning framework to recognize the traffic signs in all weather conditions so that drivers can take appropriate action at the right time to avoid a fatal crash. An optimal deep learning model with a classification network will be developed to recognize relevant features of a particular traffic sign and classify it into appropriate categories. A German Traffic Sign Recognition dataset with 43 classes and around 50,000 images is used. This study illustrates a novel approach to classify traffic signs using the ResNet deep learning model and data augmentation. The performance of this model has achieved an accuracy of 98%. In terms of accuracy, this approach successfully classifies traffic signs in all weather conditions. It has also performed well compared to previous work in this field. The performance of the model is evaluated by using various metrics, including accuracy, mean average precision, and classification report. This is the critical step in the development of the deep learning model. However, future work can improve the accuracy further while maintaining current parameters.

Keywords- TSR, ADAS, ResNet, Data Augmentation, Residual Network, Deep Learning.

1 Introduction

The future of the transportation system will rely heavily on autonomous vehicles [1]. The demand for self-driving cars is exponentially increasing, and so is the global autonomous vehicle market. Safety features are an essential prerequisite for consumers across the world. Advanced Driver Assistance System (ADAS) is an integral part of autonomous vehicles. One of the critical steps in the proper functioning of the Advanced Driver Assistance System(ADAS) is the detection and recognition of traffic signs. Several studies have examined the detection and recognition of traffic signs from images. Such as, in (Liu and Li; 2021), a system is implemented for the detection of Chinese traffic signs in a visual direction. In (Arief et al.; 2021), a system is implemented for detecting and recognizing traffic signs from offline-based video in one-way traffic during the day. However, most of these studies do not consider detecting and recognizing traffic signs under various weather conditions.

[1] `https://consult.nationaltransport.ie/ga/node/1390`

Poor weather or lighting conditions affect the visibility of traffic signs leading to accidents [2]. Unfortunately, there are very few studies on the recognition of traffic signs under various weather and lighting conditions. Such as, in (Hasegawa et al.; 2019), the system is implemented for detecting and recognizing traffic signs under three conditions, including clear weather, nighttime and small objects. Nevertheless, the drawback of this research is that it has not considered recognition under all the weather conditions, including rainy or cloudy. Also, the dataset used is not available publicly. To overcome these drawbacks, a dataset containing various weather and lightning conditions images of traffic signs is used in this research.

1.1 Motivation

Due to the high demand for self-driving cars and consumer safety features, the requirement for ADAS has emerged. The primary goal of this research is to design and develop an effective system for accurate traffic sign recognition under any weather and lighting conditions. Using this system will benefit the driver so that the driver can take necessary action. It will assist the driver by announcing or notifying various traffic signs, including pedestrian crossing ahead, sharp bends and curves, no horn zone, various speed limits, and many more. Deep learning techniques are beneficial in improving the efficiency of the TSR system. Traffic signs across the world are very diverse when it comes to color and shape. This system will help the driver by notifying them of upcoming traffic signs so that the driver can plan their controls ahead of time. For example, the TSR system can recognize various speed limit signs and traffic lights to alert the driver to keep the vehicle in the correct lane.

Adverse weather and lighting conditions can affect the driver's visibility on the road. Due to this limited visibility, it becomes difficult for the driver to recognize the traffic signs causing the fatal crashes. Various modern deep learning techniques are proposed for an effective TSR system that can also recognize traffic signs in adverse weather and lighting conditions to solve this issue. Speed of recognition and accuracy is an essential factor in autonomous vehicles. Various deep learning models, including CNN, Yolo, and VGG, are implemented for traffic sign detection and recognition. This research proposes a novel approach to traffic sign recognition using ResNet (Residual Network) and data augmentation, focusing on speed of recognition and accuracy. Data augmentation is used to increase the train dataset and overcome the issue of over-fitting.

The detailed structure of this research paper is as follows:

Section 2: Discussed previous related work related to the TSR system.
Section 3: Explained the research method and its specifications.
Section 4: Explained the design specifications.
Section 5: Explained the implementation of this method.
Section 6: Explained the evaluation of all the experiments.
Section 7: Explained the conclusion for this research paper and future scope.

1.2 Research Question and Objective

To what extent the modern deep learning techniques could contribute to help correctly recognize and classify the traffic signs in any weather conditions to prevent fatal hazards

[2] `https://www.thefloodlawfirm.com/car-accidents/bad-weather-accidents/`

The primary objective of this research study is to design and develop a deep learning framework for recognizing traffic signs from still images. The dataset includes the traffic sign images in normal weather and lighting conditions as well as adverse weather and lighting conditions. Furthermore, data augmentation is used to enhance the train dataset and improve the model's prediction accuracy. Finally, the model's performance will be evaluated using accuracy, mAP (mean average precision), and classification report.

2 Related Work

A brief review of the application of computer vision techniques in this field is discussed, along with an overview of previous TSR findings. Many researchers have worked on the image classification of traffic signs by utilizing various datasets. Therefore, it is challenging to compare studies that have used various models and datasets. However, in this section, every effort is made to comprehend and clarify the various studies conducted so far and to identify the suitable model that would work best for this research.

2.1 Traffic Sign Recognition using deep learning

Traffic sign recognition is one of the famous research areas for researchers in computer vision. Increasing demand for autonomous vehicles has made this area a research hot spot. In research (Islam; 2019) system is proposed for detecting and classifying still traffic sign images. As compared to other studies, this research uses traffic signs that are globally recognized and not limited to a specific region/country. A total of 28 categories of traffic signs are used for classification. Image augmentation is incorporated as a part of the training and validation dataset. Separate machine learning models are implemented for the classification of traffic signs and their shape. CNN is used as a classifier to find the region of interest. Around 40k images were used for traffic sign classification and 3.6k images for shape classification. HSV color space method is used in ROI selection, which separates the intensity of the image from the color information. This is beneficial for accurate color classification regardless of the surrounding environmental conditions. Bounding box techniques are used for ROI selection and converting into 64x64 and 128x128 images, which are given to the classifier as input. The overall accuracy of 90% is achieved in this study, which can be improved by implementing modern deep learning techniques. However, the performance of this model is limited to traffic sign recognition of still images in normal lighting conditions. Further improvement can be made to this research to enable this model to detect the traffic signs in various weather and lighting conditions, as well as accuracy, which can be improved further by using modern deep learning techniques.

As discussed in earlier sections, traffic sign recognition is an integral part of the advanced driver assistance system. Accurate detection and recognition of traffic signs are of utmost importance when it comes to the safety of the driver and their vehicle. However, various environmental factors can damage or obscure the traffic signs, making detection and recognition challenging. This research proposes a novel traffic sign dataset due to the limitation of substantial datasets which is standardized for benchmarking. This dataset consists of the Carla Traffic Signs (CTS) (Siniosoglou et al.; 2021). Data augmentation is used in order to simulate the phenomena such as distorted signs or signs under different weather conditions. A deep learning model with an autoencoder algorithm is proposed in this research. Obtained results show that moderate accuracy of around 90% in all scenarios is achieved. Data augmentation is used in this research to generate traffic sign images in various weather conditions. This can be avoided by obtaining a real-life dataset of traffic signs under various weather conditions. More advanced machine learning models can be used to improve further performance.

Another research which used ROI selection and multi-task CNN for traffic sign classification (Luo et al.; 2018). In this research, a new data-driven approach is proposed to classify all the types of traffic signs, including text-based as well as symbol based. Data is collected by a video camera mounted on a moving vehicle. Traffic sign regions of interest (ROI) are extracted by using the most stable external regions on the grey and adjusted RGB channels. Then refinement, classification, and pre-processing are done. Refined ROIs are assigned to their respective classes via the proposed multi-task convolutional neural network. The neural network in this research is trained with a significant amount of data, including synthetic traffic signs and images taken from street views. Post-processing eventually aggregates the data from all the frames in the classification process. The ultimate result for a video is created by fusing the recognition outputs of all frames. The author claims the excellent efficiency of the proposed system by using the experimental data. This research shows that the CNN model also worked satisfactorily on the video data. Further improvement can be made by using a pre-trained transfer learning model, which can improve the recognition speed and accuracy of the system.

As discussed earlier, the accuracy of traffic sign detection and recognition is affected by the nighttime and bad weather conditions. This research identifies those issues and proposes a deep learning model (Hasegawa et al.; 2019). The accuracy of small objects in images is identified as low. For example, recognition of distant traffic signs is more challenging than nearby traffic signs in still images. A Japanese traffic sign dataset is used for this research which contains 16 classes of traffic signs. Three types of test datasets are created, including clear weather, night time and small object or distant traffic sign datasets. Conventional YOLOv2 and Faster RCNN are used in this research for detection and recognition. The deep learning model is trained using multiple sizes of images in order to make it robust to various scale changes. The author claims the good results for detection and recognition in terms of accuracy. However, the author also faced the issue of training data shortage, which can be fixed in future work. To conclude, training data can be increased in order to improve the accuracy further. Furthermore, more number of classes and weather conditions can be included in the dataset to make a more robust model.

As we have discussed, the convolutional neural network has been used by many researchers to perform the classification. Another research has proposed the CNN method for traffic sign recognition Kapoor et al. (2021). Tensor flow and Keras library are used to train the model in this study. Traffic sign images are divided into sub-regions based on the color of signs using a region proposal based on segmentation. The training dataset is increased in order to solve the over-fitting issue. The layers in CNN are chosen such that data dimensionality reduces. The architecture consists of one softmax layer and multiple convolutional layers. As this is a multi-class classification problem, the softmax layer is a must. The softmax layer works as the activation function in the output layer. The author claims excellent results with 93% accuracy of the proposed model. This research is helpful in understanding the development of a classification network system for traffic sign recognition.

In this research, the author proposed a cascaded CNN model for traffic sign recognition (Kong et al.; 2019). Dataset collected from the traffic signs in South Korea. The author compared the proposed model with YOLO-v2 and claims that the proposed method is hardware friendly and requires much less computational power. The proposed method has two networks, including the detection and recognition network. The detection network is the regression that detects the patch candidates from the input frames. These patch candidates are used as input for the recognition network. The Recognition network works as a classifier to classify these patches. Recognition network design consists of a convolutional layer followed by max-pooling and softmax layer. The author claims a reduction in computational time by around 55% compared to the earlier YOLOv2 tiny model. In addition, the accuracy of around 90% and mAP of 93% is achieved. To conclude, the proposed method is competent in terms of computational complexity, but the dataset used needs to be improved by including more number of traffic sign categories. Also, the recognition rate can be improved further by fine-tuning the model parameters.

Yolo(You Only Look Once) is the well know object detection system. The latest versions of Yolo have shown promising object detection results. The author of this research proposed a system that consists of a Yolov3 detection network and CNN-based classifier for traffic sign recognition (Rajendran et al.; 2019). Yolov3 is a real-time object detection network based on logistic regression. Yolov3 consists of a darknet and residual network and has a total of 53 convolutional layers. Darknet is a feature extractor in this network. It considers the input image as a grid and predicts the bounding boxes and probabilities for various classes. Both the classifier and detection networks are implemented using Keras. In order to improve the speed and reduce the computational complexity, the classifier network is designed with n X 1 convolution followed by a (1 X n) layer. This system has achieved an mAP of 92.2%. Though detector performance was good, few false positive detections have been received. To conclude, the Yolov3 detector with CNN classifier works best for traffic sign recognition. Further accuracy can be improved by using a better classifier model.

Another version of the latest Yolo series, which is Yolov4, is used in this research for traffic sign detection and recognition (Fernando and Sotheeswaran; 2021). The author discussed various causes of collisions due to unrecognized traffic signs along the road. Such as traffic signs located on the far side of the road, traffic signs with significant background, clutter, and changing weather conditions. Bounding box prediction is used as a detection system, and CNN is used as a classifier model. Dataset is manually labeled and used to train the model. Google Colab platform is used to perform this experiment. Feature extraction is performed by CSP Darknet53, which is the backbone of Yolov4. The author claims 84.7% of recognition accuracy and reduction in computation time, which is the benefit of using the Google Colab platform. Along with that, 161 layers of Yolov4 also contributed to higher accuracy. To conclude, Yolov4 outperforms the earlier Yolo versions in terms of accuracy. Also, Google Colab can be used in this research to save computational time.

This research proposed a deep learning model for traffic sign recognition using a convolutional network based on the SSD algorithm (Wang; 2018). First, image normalization is performed before giving the images as input to the model. Then, the pixel value is adjusted in accordance with the proposed model's performance. VGG-16 is used as a top network in the SSD algorithm. It contains convolutional, max-pooling, and softmax layers. The convolutional layer is used as a feature extractor. Finally, output from the fully connected layer is given as input to the softmax layer, which will classify the traffic signs into appropriate categories. The author claims good results as the SSD algorithm has mainly reduced the value of negative samples in training, so that accuracy of the softmax layer is improved to a great extent. To conclude, this experiment is performed on a limited data set, which can be improved by including a diverse dataset of traffic signs. Also, this method is computationally expensive, so it will not be feasible to use in real-time devices.

The author of this research addressed the issue of recognition of a large number of traffic sign categories (Tabernik and Skočaj; 2020). The mask RCNN method is proposed in this study. Around 200 traffic sign categories are included in the dataset used in this research. This dataset is acquired by DFG consulting on Slovenian roads in RGB format. Data was collected in high resolution of 1920 X 1080, but during training, the model was resized due to memory limitations. The regional proposal network is first evaluated, which assesses the region quality and chooses the top regions in quality. After that classification network classifies the traffic signs into appropriate categories. The author claims good results with an mAP of 95.5%. To conclude, mask RCNN proves to be the robust model that can be used to recognize the higher number of traffic sign classes.

Yolov3 is proposed in this research for traffic sign recognition because of its benefits such as high accuracy, robust performance, and quick response (Alhabshee and Shamsudin; 2020). Image pre-processing is used to compensate for the effects of various illuminations and lightning conditions on images. The Malaysian traffic sign dataset is used in this experiment. A bounding box label is used for the annotation in the training dataset. In the execution phase of the model, first, it will resize the images to (448 X 448) then the convolutional network will be evaluated. Depending on the object in the grid lines, class probabilities are predicted, and after that, traffic signs are recognized into categories. A total of 1000 traffic sign images are collected from 10 different categories and used for classification. The research concludes that higher accuracy can be achieved by using various images in each category. The author also concludes that Yolov3 is a robust model with high performance. Recognition time is also improved by using this model. Further accuracy can be improved by using a more extensive and more diverse training dataset and more epoch size.

The author of this research performed the ablation analysis of eight models, including three CNN models and five MLP(Multi-layer Perceptron) (Ng et al.; 2021). This research performed all eight analyses with varfious architecture and batch normalization. Dataset for this experiment is collected from Google Street view functionality of Google maps. Around 5000 images are collected from eight different classes of traffic signs. In data preprocessing, images are cropped and equalized, which nullifies the noise. Data is segmented into the test, train, and validation datasets. Data augmentation is performed on a training dataset to improve the dataset's size. Data augmentation includes cropping, flipping, rotating, and zooming. MLP is a neural network with input, output, and a few hidden layers. All the layers are connected, forming a dense network. Training is performed with 50 epochs and 32 batch sizes. Results for MLP show that batch normalization improves the stability of the model when it is before the activation layer. Accuracy of 92.6% is achieved by using MLP. On the other hand, CNN achieved better accuracy and reduced the execution time. To conclude, CNN proves to be the better model in this experiment. Further accuracy and stability can be improved by adding a dropout layer and batch normalization in CNN.

The author of this research proposed a neural network and trained it for traffic sign recognition. This network has two neural networks, one is for region proposal, and the other is for object classification (Tiron and Poboroniuc; 2019). The custom dataset uses only three traffic sign classes, including pedestrian, stop, and parking signs. RPN (regional proposal network) identifies the object that needs to be recognized, and RCNN classifies the input images faster. Faster RCNN contains the VGG layers, which work as the feature extractor. A feature map is then used for classification. The main task of RCNN is to adjust the bounding box of the detected image region so that it can classify the image into the corresponding category. The traffic sign dataset used in this research is collected using a mobile camera from Romania. Dataset is split into 7:3 ratio for training and validation. Five different meteorological conditions in real life are selected for testing the model. These conditions are day, night, fog, rain, and snow. The author claims 96% accuracy of the proposed model in all five environmental conditions for these three traffic signs. More traffic sign classes can be added in the future to improve the system reliability and be able to use in a real-life autonomous vehicle. In that case, the model needs to be re-trained using new traffic sign image classes.

This research presents an improved CNN model with HOG features for traffic sign recognition. A primary CNN multi-layer network with HOG is used for classifying traffic signs in this experiment (Shangzheng; 2019). The first image is greyed in HOG feature extraction to reduce the illumination effect. A Gamma correction algorithm is used to reduce the impact of local shadows. In image gradient convolution, edge detection algorithms convolute the image. After the cell segmentation, gradient information of the pixels is counted from the histogram. Cell units are grouped to obtain the HOG features, and feature vectors are superimposed. Sixteen classes of traffic signs are incorporated in the video dataset used in this research. The proposed method has achieved high accuracy. To conclude, this approach has less computational complexity and faster operation speed than the CNN approach. Furthermore, this approach can be used on the larger image dataset of traffic signs to test the model's robustness.

This is another research in which the SVM(support vector machine) model is proposed with HOG(histogram of oriented gradient) features (Liu et al.; 2021). A Chinese traffic sign dataset with around 610 prohibition traffic sign images from 15 classes is used in this research. The first step is image pre-processing, in which image size adjustment, image greying, gamma correction, and feature extraction are performed. After that, HOG feature extraction is performed. Finally, SVM is used to perform the classification. The linear transformation is incorporated after the training phase of prohibition traffic signs to train the classifier model. Around 42 binary classifiers are trained, and a prohibition classifier is developed. The custom dataset obtained by the author is then used to evaluate the model performance. The author claims an overall accuracy of 90.2% for this prohibition traffic sign recognition method. To conclude, HOG proves to be an efficient technique for feature extraction by using pixel gradients in image classification. The optimal image size can be considered in future work to improve the classifier model's training efficiency.

A transfer learning approach is proposed in this research paper to detect and recognize the stop traffic sign (Wei et al.; 2018). Various images of traffic signs are used to train this proposed model. The migration training method is incorporated in this research as it is possible to transfer this learning to the new tasks using the exemplary tuned network. Transfer learning significantly reduces the dataset's required training time and size for the model. The RCNN model is pre-trained using the CIFAR-10 dataset and fine-tuned for stop traffic signs by using just 41 images. The RCNN method returns the detection score and the traffic sign category label in the output. This research claims good results for stop sign recognition. In the future, the same approach can be applied to other classes of traffic signs.

The author of this research proposed a transfer learning method for traffic sign recognition (Lin et al.; 2019). Inception v3 model is used in this research, reducing the computational complexity and training data required to train the model. Belgium traffic sign dataset is used in this research. Data augmentation is performed on the dataset in order to enhance the dataset and to analyze the layer-wise convolutional feature presentation. In order to make the robust model, it is trained several times with various learning rate by fine-tuning the parameters. This research claims the higher recognition performance of the model at a 0.05 learning rate. 99% recognition accuracy is achieved by using this model. The author compared the transfer learning method performance with other methods and claimed that transfer learning is more robust and efficient than other models. To conclude, the transfer learning approach is also reliable and efficient for traffic sign recognition.

A CNN model with a low number of parameters based on color space is proposed in this research for traffic sign recognition (Yildiz and Dizdaroğlu; 2021). The input image is subjected to various variants, sizes, and color spaces in the pre-processing stage. Data augmentation is incorporated to balance the classes in the training dataset. The recognition rate varies greatly depending upon the input image size. The recognition rate reduces when training images vary significantly in their size. This research also claims that color spaces significantly influence recognition accuracy. This research includes RGB, CIELab, RIQ, and LGI color spaces. Ten convolutional layers, three max-pooling layers, and three fully linked hidden layers are included in CNN. The softmax layer is coupled to the final layer to create a probability distribution of all the classes. To address the imbalance in the dataset, data augmentation is performed, which includes zooming, rotating, and shifting. Google Colab is used with 50 epoch size with Keras package and TensorFlow platform for training and testing. The model with the 60x60 input image size and RIQ-LGI color spaces achieved the best accuracy of all.

2.2 Use of Data Augmentation

Recognition time and size of the model are essential factors in traffic sign recognition which are discussed in (He et al.; 2020) research. The classic convolutional neural network is used in this study. The advantages of the CNN model are discussed in this paper. The author compared CNN with traditional machine learning models and claims that adding additional convolutional layers is beneficial to extract more features at various levels in the image and exhibits translation invariance to the input image. Additionally, parameter size can be reduced, and the pooling layer can reduce the image dimensions by keeping the critical feature information and speeding up the model training process. Data augmentation is incorporated to overcome the uneven distribution of classes in the dataset. This will enhance the network capacity for generalization and reduce the unnecessary computational load. Data augmentation techniques used in this study include rotation, zoom, flip and random cropping. CNN model used in this study is constructed by using ten layers, primarily including convolutional and pooling. The author claims the good theoretical efficiency of the model. To conclude, data augmentation helped improve the accuracy and over-fitting issue by generalizing the network capacity. Traffic signs under bad weather conditions can be considered a future experiment in this research.

Data augmentation is helpful for enhancing the training dataset and improving the efficiency of the machine learning model. The author of this research proposed an approach of data augmentation of various quality deficits and applied it to the training dataset

(Jöckel et al.; 2019). In this method, training data is augmented with combinations of seven quality deficits, affecting traffic sign recognition. A few of them are rain, darkness, motion blur, and dirt on the lens. It also considers dependency between combined deficits and the effect of other contextual information. This approach is beneficial in providing the augmentation functionality in the training dataset for the model training phase. Images with quality deficits can be generated and included in the training dataset so that the efficiency of the model can be improved, making it a robust model. This approach can be used for various classification problems along with traffic sign classification.

2.3 Literature summary and analysis

All the research mentioned above and studies have significantly benefited this research to a great extent. Many deep learning models have been used in this research to develop the traffic sign recognition system, including CNN, various Yolo versions, and transfer learning methods. Most of the studies have used CNN as their classification model, and CNN proves to be an effective and robust recognition model for traffic signs. Very few studies have worked on recognizing traffic signs in adverse weather and lighting conditions. All the studies and proposed models have performed well, but there is still scope for improvement of traffic sign recognition in various weather and lighting conditions. Transfer learning methods can be used to improve the further efficiency and robustness of the system. Data augmentation is implemented in a few research discussed above to enhance the training dataset and improve the model's efficiency, which can prevent data scarcity. Data augmentation helped in improving the overall accuracy and resolved the overfitting issue by generalizing the network capacity. Data augmentation has caused the training time to increase accordingly. Getting clean and enough data for the research is also a challenging task which is faced by the researchers. In conclusion, this research will evaluate the deep learning model with data augmentation for traffic sign recognition. Data augmentation can benefit this research as the neural network works best with a substantial amount of data.

3 Methodology

As discussed in the literature section, all the deep learning models used for traffic sign recognition have performed very well. However, very few have considered the dataset with the traffic sign images in various weather and lighting conditions. Additionally, very few have implemented the deep learning model with the data augmentation for model training. To address these problems, ResNet deep learning model is used in this research with the dataset that contains traffic signs in various lighting conditions. Furthermore, data augmentation is used to overcome the issue of model overfitting. This research employs the CRISP-DM (Cross Industry Standard Process for data mining) project implementation approach. This approach has six parts, as shown below in figure 1 (Chapman et al.; 1999).

3.1 Business Understanding

Deep learning models such as CNN, RCNN, and Yolo, which are discussed in the earlier section, are quite good at recognizing traffic signs. However, since it is a crucial task, conventional deep learning models can be inefficient sometimes. Employing the latest deep

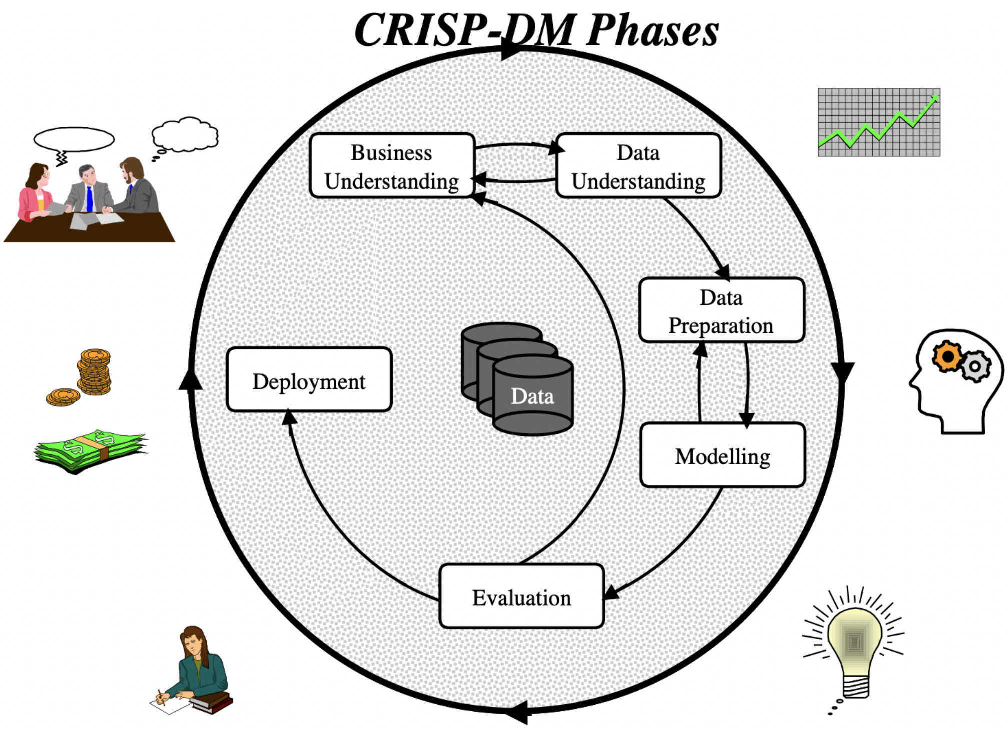

Figure 1: CRISP-DM Phases

learning models can produce a more accurate and robust system. Traffic sign recognition is a part of the advanced driver assistance system (ADAS), which is a crucial element of an autonomous vehicle. Sometimes mistakes caused by individuals result in the deadliest accidents. In order to prevent this, ADAS offers a safe man-machine interface, thereby lowering the risk of accidents. ADAS has several applications, including traffic sign recognition, speed limits, pedestrian detection, no overtaking zone, and many more. The aim of this research is to develop a reliable, efficient, and robust system that will assist the driver in preventing the risk of car accidents.

3.2 Data Understanding

German traffic sign recognition benchmark (GTSRB) dataset is used in this research. This dataset is obtained from the Kaggle website. [3] The dataset includes vast images of traffic signs with 43 classes. The dataset has around 39K training images and 12K test images to evaluate the model. The file format of all the images in the dataset is PNG, which can be used directly for the research. Metadata of the images is also included in the dataset, which can be helpful for the research. Metadata includes dimensions of the images, ROI, class id, and path.

3.3 Data Preparation

The dataset has around 39K training images which belong to 43 classes of traffic signs. Training images are split into train and validation datasets to train the model. The dimension of all the images in the dataset is set to 32 in height and 32 in width at the

[3]https://www.kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign

start of the evaluation. Images are first to read into a list and then converted to an array for further processing. A separate dataset of 12K test images is used for the final model evaluation.

3.3.1 Data splitting

The training dataset of 39K traffic sign images is split into train and validation datasets with the ratio of 80:20, respectively.

3.3.2 Data Augmentation

Data augmentation is used to improve the model's prediction accuracy. Data augmentation techniques include rotation, zoom, width, and height shift. All these techniques are mainly used to improve the prediction accuracy, overall efficiency, and robustness of the model. In addition, it is a crucial method to increase the size of the training dataset. For example, figure 2 shows the data augmentation for a traffic sign.

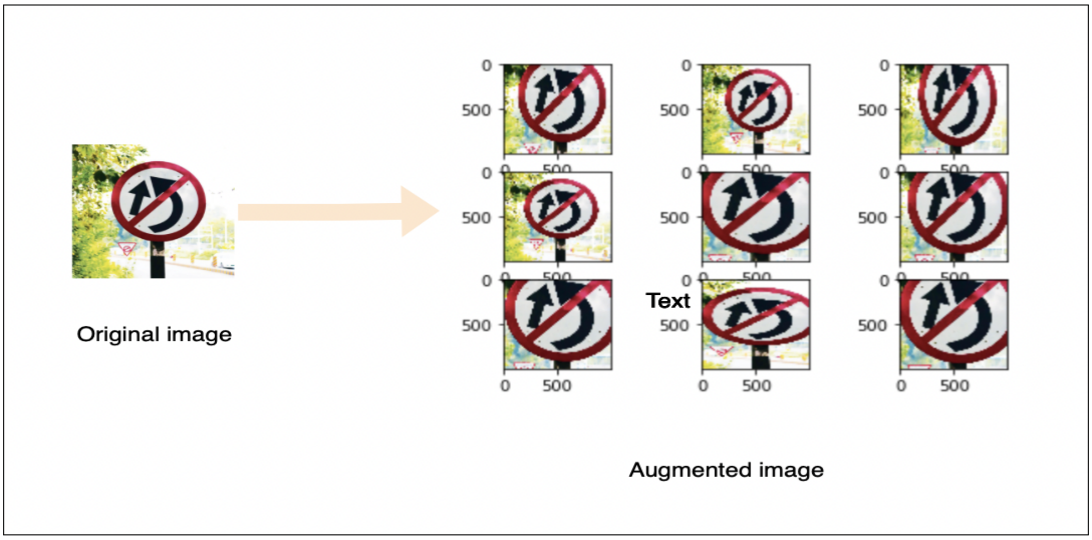

Figure 2: Data Augmentation

3.4 Modeling and Evaluation

Modeling of the proposed ResNet model for traffic sign recognition is described in this section. The workflow of the model is shown in figure 3.

1. The training dataset is split into train and validation data with a ratio of 80:20, respectively.

2. Following that, data augmentation is performed on train data, where train data is expanded using various techniques such as zooming, rotating, and horizontal shift.

3. Augmented train data is used for model training. ResNet model is trained in this phase and validated by using validation data. Validation accuracy is used as a metric for evaluation in this phase. Epoch size and learning rate are configured at this stage.

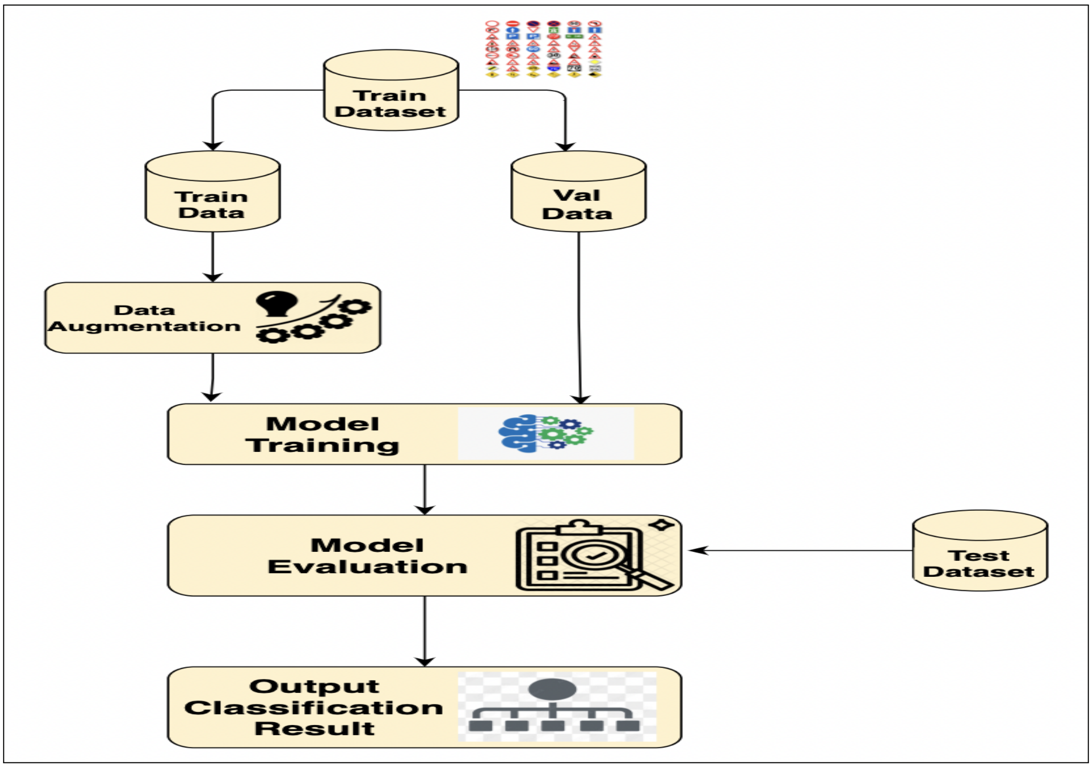

Figure 3: Model workflow

4. A test dataset is used to evaluate this trained model. Prediction on test data is performed at this stage, and test accuracy is calculated.

5. The confusion matrix is visualized, and a classification report is generated at this stage. Random samples from the test dataset are visualized by plotting actual and predicted class values.

4 Design Specification

ResNet model is developed and used in this research for traffic sign recognition. The TensorFlow and the Keras libraries are used in this research to implement the model. The proposed method is inspired by the literature review of He et al. (2016). The ResNet model used in this research is a deep neural network with 30 deep layers, including convolutional and dense layers. Kernel size is (3 x3). ReLu activation layer and batch normalization layer are used after the convolutional block. In addition, an average pooling layer with a pool size of 4 is included. At the output of the model, flatten layer and dense layer with a softmax function of 43 nodes are configured. ResNet model with skip layer functionality is mainly designed to handle the issues of image recognition. However, it is also used for image classification, object detection, and object localization to improve the accuracy (Bouaafia et al.; 2021). In this research, it is used for traffic sign image classification. Figure 4 shows the ResNet model architecture.

- The first layer is the input layer with output size (None, 32, 32, 3). It has a kernel size of (3 x 3) and a stride size of 2.

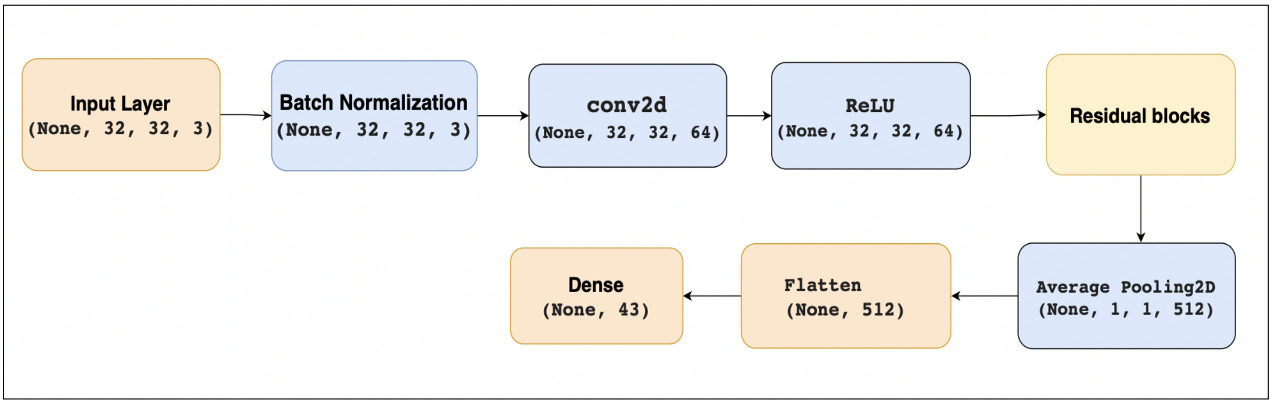

Figure 4: ResNet model Architecture

- Next is the Batch normalization layer with an output size of (None, 32, 32, 3). After that, the Convolutional layer with output size of (None, 32, 32, 64) and the ReLu activation layer with (None, 32, 32, 3) output size.
- Following that, there are four stages of residual blocks with 2, 5, 5, and 2 blocks at each stage, respectively.
- Finally, there is an average pool, flatten and dense layer with softmax function. In total, there are 30 deep convolutional layers.

ResNet (Residual Network) is a deep learning model that adds the skip connection from one network point to the next. The benefit of adding the skip connection is that the model can learn the identity mapping going forward into their deep neural network so that the network can learn extra F(x). Suppose there is no more learning left for a network from a dataset in the training phase. In that case, that network can just learn the identity mapping going forward in their neural network rather than negative learning, which can drop the accuracy. As the dataset in this research is different, for feature extraction purposes, already calibrated weights such as imagenet are not incorporated in this model. Instead, flatten layer is used to connect the dimensions of residual blocks with the output layers. Finally, the average pool and dense layer with the softmax function are used for classification.

5 Implementation

This section will outline the entire process of implementation and execution. This includes the setup of the environment, data processing steps, model execution, and the tools used for this implementation.

5.1 Environment Setup

The range of tools used in this research is based on the implementation stages.

- **System configuration:** Mac OS Monterey, 8GB RAM, 7 core GPU
- **Programming language:** Python 3.6

- **IDE:** Google Colab

5.2 Dataset pre-processing

The German Traffic Sign Recognition Benchmark (GTSRB) is obtained from the Kaggle website. This dataset was used for a multi-class, single image classification challenge at International Joint Conference on Neural Networks (IJCNN) 2011 [4]. The dataset has a train as well as a test data folder. The training dataset has around 39K images from 43 classes of traffic signs. The test dataset has around 12K images. Metadata is also included in the dataset, which has dimensions of the images, ROI, class Id, and path to the images. The dataset is uploaded to Google Drive and then accessed in the Colab notebook to perform the research [5]. Following that, both the dataset are visualized by using the python library "matplotlib.pyplot" [6]. The list dataset is converted to an array using the "numpy" library [7]. Then train dataset is split into train and validation data by using "train_test_split" function with the ratio of 80:20, respectively. The train and validation data are used for model training. Data augmentation is performed on train data before model training. "ImageDataGenerator" from TensorFlow-Keras is used for this process [8].

5.3 Model Training

ResNet (Residual Network) is a deep learning model. The advantage of this model is that it can learn the identity mapping going forward into their deep neural network so that network can learn extra F(x). Input shape is set to (32,32,3), as per the input image dimensions set at the setup stage. 'Softmax' function with the output shape of 43 is used for classification. The learning rate is set to 0.001; as per recommendation from (He et al.; 2016) learning rate should be as low as possible. Batch size set to 32. Multiple residual blocks are added in this model from a pre-defined function, as feature extraction is required for this dataset. 'Adam' optimizer is selected for this model training process. Though the 'Adam' optimizer is computationally costly, it is a fast method and converges rapidly. Data augmentation is used while training the model. Accuracy is used as a metric for model training and validation. One hot encoding is used for the classes in training data; because of that loss function is selected as 'categorical_crossentropy' in model compile. Epoch size is selected after various trial runs of the model. 'val_loss' and 'val_accuracy' is taken into account while selecting the epoch size. It is ensured that the model is not over-fitting.

6 Evaluation

This section will cover several experiments performed in this research to develop a robust and efficient model. Confusion matrix, model accuracy, model loss, and various graphical representations are used to evaluate the performance of the model.

[4] https://www.kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign
[5] https://colab.research.google.com
[6] https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.html
[7] https://numpy.org
[8] https://www.tensorflow.org/api_docs/python/tf/Keras/preprocessing/image/ImageDataGenerator

6.1 Basic CNN

The first experiment is performed by using the basic convolutional neural network, including basic convolutional, pooling, and activation layers. Next, the softmax function is used for classification output with 43 nodes. Next, Tensorflow and Keras are used to define the sequential CNN model in Google Colab. Figure 5 shows the evaluation graphs of the model, including model accuracy and model loss.

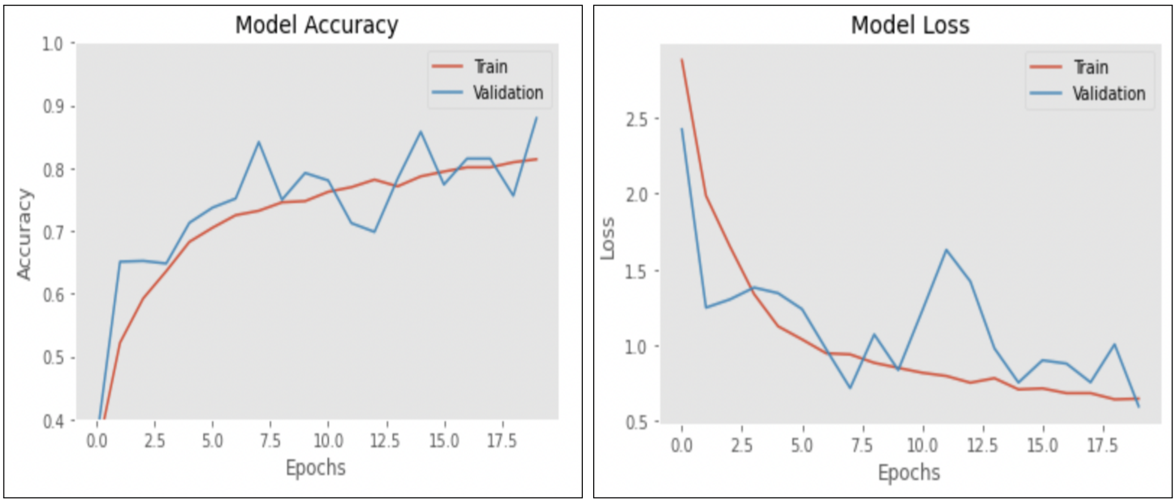

Figure 5: Model accuracy and loss

Figure 6 shows the confusion matrix of the model.

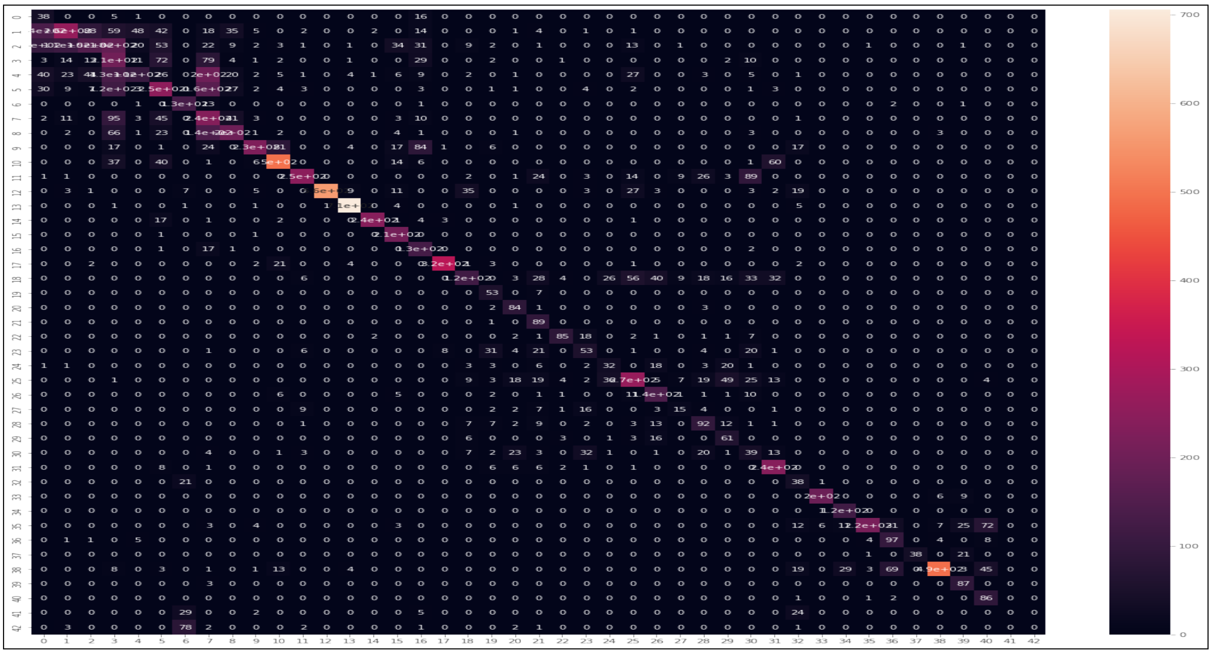

Figure 6: Confusion Matrix

This model is able to classify the traffic signs with a validation accuracy of 87% and test accuracy of 84%. However, as shown in the confusion matrix, there are still some classes with lower accuracy, which can be improved.

6.2 ResNet

The second experiment is performed by using the ResNet model. ResNet is a network of residual blocks that has a skip connection. This model does not use the transfer learning weight such as the imagenet. This model includes an input layer with (32, 32, 3) shape, conv2d with 64 filters, residual blocks, average pool layer, flatten layer, and dense layer with 43 output nodes. Again, Tensorflow and Keras are used to define the model. Figure 7 shows the model evaluation graphs.

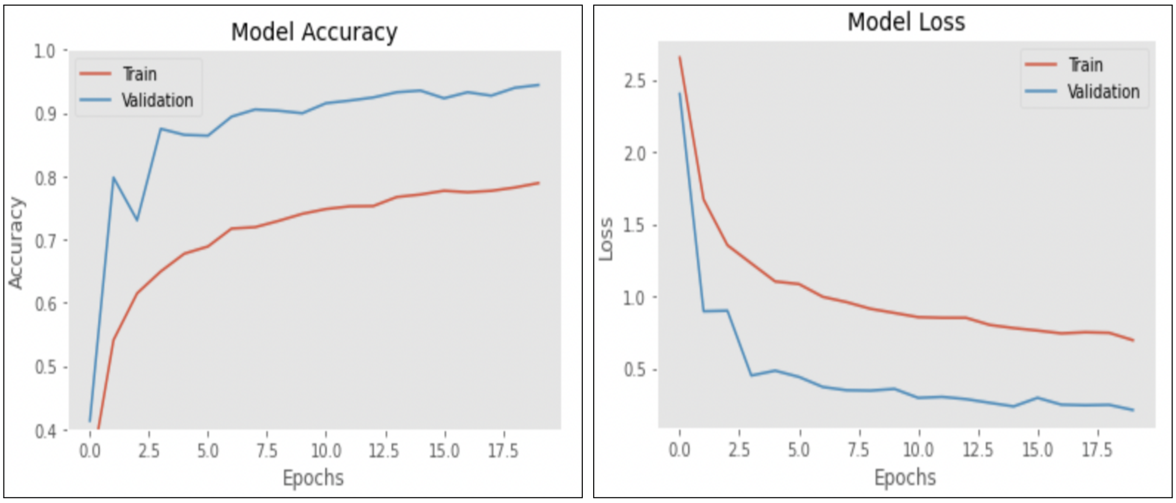

Figure 7: Model accuracy and loss

Figure 8 shows the confusion matrix of the model.

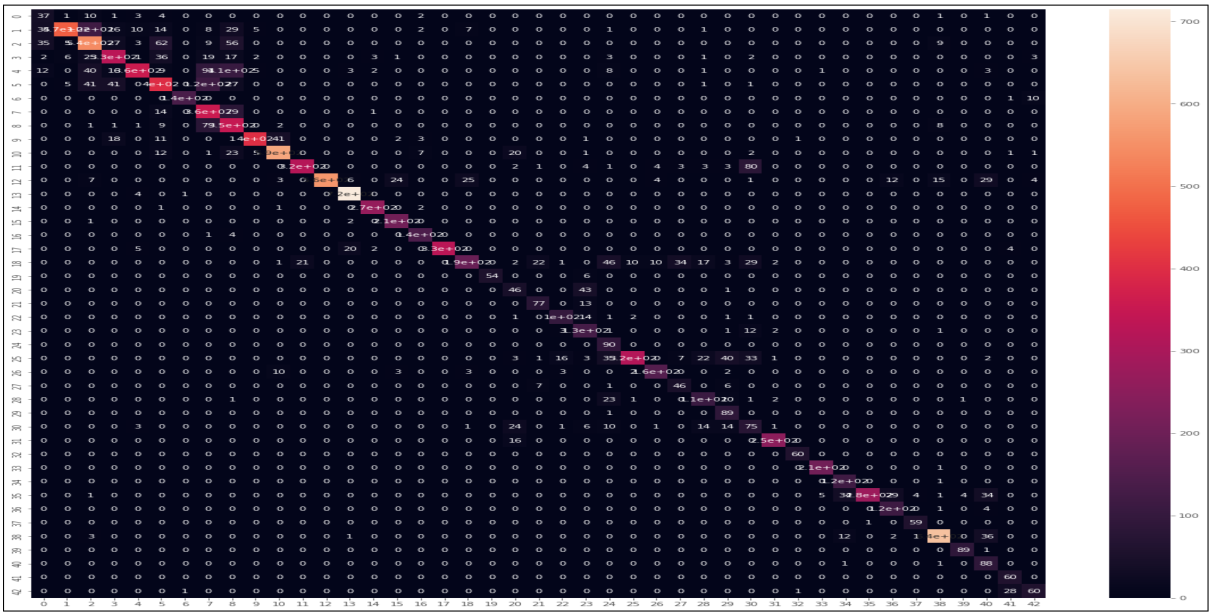

Figure 8: Confusion Matrix

The model is able to classify all the traffic signs with a validation accuracy of 92% and test accuracy of 89%. Data augmentation will be used to improve the model performance further.

6.3 Data Augmentation

In this experiment, data augmentation is used along with the ResNet model. In data augmentation, data is subjected to rotation, zoom, height and width shift, as well as shear range. To improve prediction accuracy, data augmentation is used in this research. Figure 9 shows the model evaluation graphs. As shown in the figure model accuracy graph is improved.

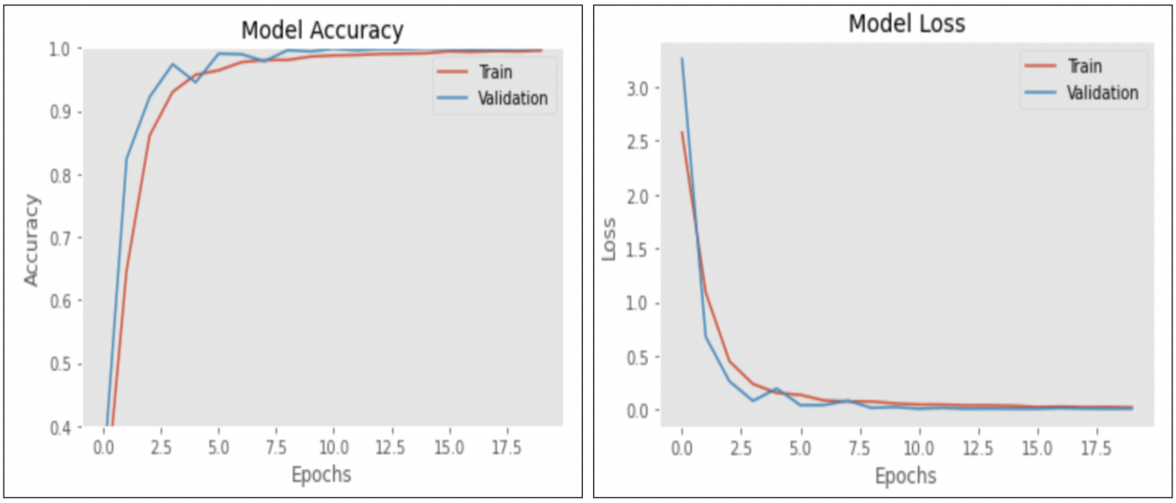

Figure 9: Model accuracy and loss

Figure 10 shows the confusion matrix for this experiment.

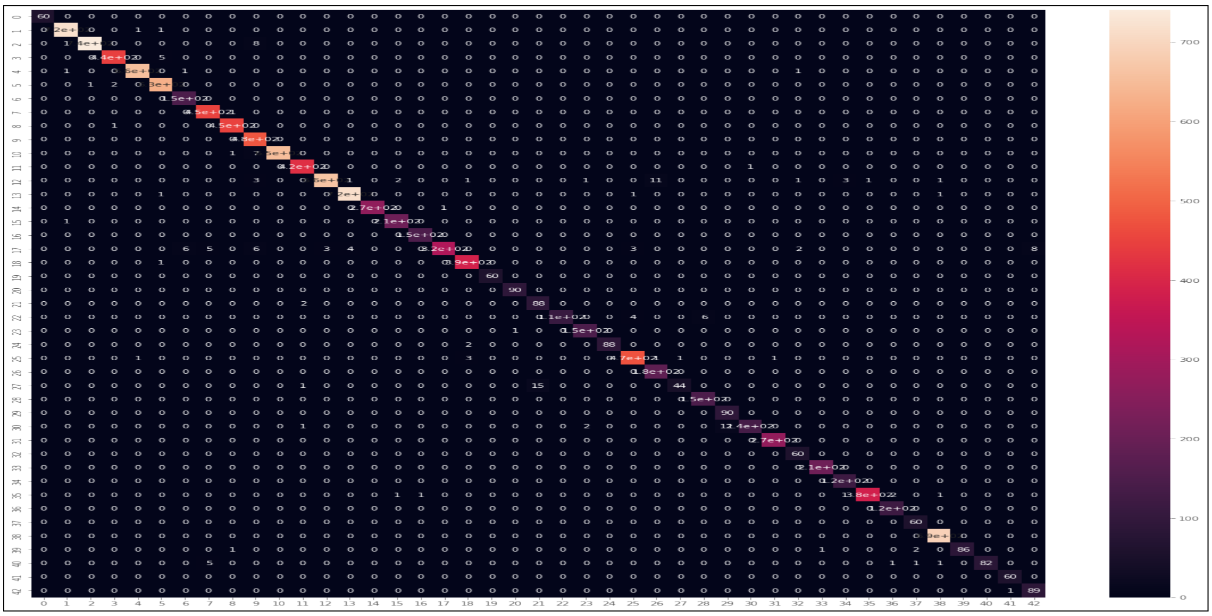

Figure 10: Confusion Matrix

This approach is able to classify the traffic signs with the validation accuracy of 99% and test accuracy of 98%. Data augmentation has helped substantially in improving prediction accuracy.

6.4 Discussion

This research aims to critically evaluate the performance of the proposed ResNet model for traffic sign recognition in various weather and lighting conditions, along with data augmentation. As discussed in the literature review, (Islam; 2019) used CNN for the classification of traffic signs in different environments, and that is globally recognized. The primary difference between the proposed model and the model developed in (Islam; 2019) is that ResNet is used for recognition and data augmentation to enhance the model performance. The dataset used in this research also includes the traffic signs under various weather and lighting conditions. There are 43 classes of traffic signs in this dataset. Each class has traffic sign images under normal and bad lighting and weather conditions. Accuracy is used as a metric for evaluation in this research. System developed in (Islam; 2019) able to classify 28 classes of traffic signs with moderate accuracy of 90% and mAP of 81%. In this research, the proposed system is able to classify 43 classes the traffic signs with the test accuracy of 98% and mAP of 97%, which is higher than (Islam; 2019). In the first few experiments of this research, lower test accuracy was achieved. However, after fine-tuning the model, applying the data augmentation, and configuring the learning rate as well as epoch and batch size, model performance is improved, and accuracy of 98% is achieved. (Kapoor et al.; 2021) also used CNN and German traffic sign recognition benchmark (GTSRB) dataset and was able to achieve 93.58% accuracy. This research outcomes are able to prove that data augmentation can enhance the performance of the deep learning model. Google Colab is used in this research for the implementation of the model, and it offers a limited resource, including RAM, disk, and runtime. Even after subscribing to Google Colab pro, there were a few issues, such as session termination and run time disconnect after utilizing the RAM capacity. To conclude, the overall results of this research are satisfactory, and the proposed system can recognize the traffic signs under various weather and lighting conditions with minimal error. Further changes can be made to this existing system in order to improve the performance of the model.

7 Conclusion and Future Work

This research mainly aims to recognize traffic signs in various weather conditions by employing the ResNet and data augmentation. The proposed model is developed by using the ResNet(residual neural network) model, which is a connection of various residual blocks. A residual block helps the neural network by adding the identity mapping from one layer to the next layer. Data augmentation techniques such as rotation, zoom, shift, and shear range are incorporated to train the model. Various research experiments are evaluated, such as implementing basic CNN and ResNet models with and without data augmentation. The ResNet model with the data augmentation proves to be the best by achieving higher accuracy. Results revealed that data augmentation is of great importance in improving the performance of the model. The developed model is able to classify the traffic signs under various weather and lighting conditions with the test accuracy of 98% and mAP of 97%. The model performed better than (Islam; 2019) and (Kapoor et al.; 2021) in terms of accuracy and mAP.

In future work, the model can be further tuned to improve its performance by modifying the model's structure and configuring different optimizers and loss functions. The model can be upgraded to be used on video datasets captured from moving vehicles.

References

Alhabshee, S. M. and Shamsudin, A. U. b. (2020). Deep learning traffic sign recognition in autonomous vehicle, *2020 IEEE Student Conference on Research and Development (SCOReD)*, pp. 438–442.

Arief, R. W., Nurtanio, I. and Samman, F. A. (2021). Traffic signs detection and recognition system using the yolov4 algorithm, *2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, pp. 1–6.

Bouaafia, S., Messaoud, S., Maraoui, A., Ammari, A. C., Khriji, L. and Machhout, M. (2021). Deep pre-trained models for computer vision applications: Traffic sign recognition, *2021 18th International Multi-Conference on Systems, Signals Devices (SSD)*, pp. 23–28.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. and Wirth, R. (1999). The crisp-dm user guide, *4th CRISP-DM SIG Workshop in Brussels in March*, Vol. 1999, sn.

Fernando, W. H. D. and Sotheeswaran, S. (2021). Automatic road traffic signs detection and recognition using 'you only look once' version 4 (yolov4), *2021 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Vol. 4, pp. 38–43.

Hasegawa, R., Iwamoto, Y. and Chen, Y.-W. (2019). Robust detection and recognition of japanese traffic sign in the complex scenes based on deep learning, *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*, pp. 575–578.

He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.

He, Z., Xiao, Z. and Yan, Z. (2020). Traffic sign recognition based on convolutional neural network model, *2020 Chinese Automation Congress (CAC)*, pp. 155–158.

Islam, M. T. (2019). Traffic sign detection and recognition based on convolutional neural networks, *2019 International Conference on Advances in Computing, Communication and Control (ICAC3)*, pp. 1–6.

Jöckel, L., Kläs, M. and Martínez-Fernández, S. (2019). Safe traffic sign recognition through data augmentation for autonomous vehicles software, *2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, pp. 540–541.

Kapoor, A., Nehra, N. and Deshwal, D. (2021). Traffic signs recognition using cnn, *2021 International Conference on Industrial Electronics Research and Applications (ICIERA)*, pp. 1–4.

Kong, S., Park, J., Lee, S.-S. and Jang, S.-J. (2019). Lightweight traffic sign recognition algorithm based on cascaded cnn, *2019 19th International Conference on Control, Automation and Systems (ICCAS)*, pp. 506–509.

Lin, C., Li, L., Luo, W., Wang, K. C. and Guo, J. (2019). Transfer learning based traffic sign recognition using inception-v3 model, *Periodica Polytechnica Transportation Engineering* **47**(3): 242–250.

Liu, S. and Li, H. (2021). Application of chinese traffic sign detection based on yolov4, *2021 7th International Conference on Computer and Communications (ICCC)*, pp. 804–808.

Liu, Y., Zhong, W., Wang, W., Cao, Q. and Luo, K. (2021). A method for recognizing prohibition traffic sign based on hog-svm, *2021 International Conference on Computer Engineering and Application (ICCEA)*, pp. 486–490.

Luo, H., Yang, Y., Tong, B., Wu, F. and Fan, B. (2018). Traffic sign recognition using a multi-task convolutional neural network, *IEEE Transactions on Intelligent Transportation Systems* **19**(4): 1100–1111.

Ng, Z. B., Ming Lim, K. and Lee, C. P. (2021). Traffic sign recognition with convolutional neural network, *2021 9th International Conference on Information and Communication Technology (ICoICT)*, pp. 48–53.

Rajendran, S. P., Shine, L., Pradeep, R. and Vijayaraghavan, S. (2019). Real-time traffic sign recognition using yolov3 based detector, *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–7.

Shangzheng, L. (2019). A traffic sign image recognition and classification approach based on convolutional neural network, *2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 408–411.

Siniosoglou, I., Sarigiannidis, P., Spyridis, Y., Khadka, A., Efstathopoulos, G. and Lagkas, T. (2021). Synthetic traffic signs dataset for traffic sign detection recognition in distributed smart systems, *2021 17th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 302–308.

Tabernik, D. and Skočaj, D. (2020). Deep learning for large-scale traffic-sign detection and recognition, *IEEE Transactions on Intelligent Transportation Systems* **21**(4): 1427–1440.

Tiron, G.-Z. and Poboroniuc, M.-S. (2019). Neural network based traffic sign recognition for autonomous driving, *2019 International Conference on Electromechanical and Energy Systems (SIELMEN)*, pp. 1–5.

Wang, C. (2018). Research and application of traffic sign detection and recognition based on deep learning, *2018 International Conference on Robots Intelligent System (ICRIS)*, pp. 150–152.

Wei, L., Runge, L. and Xiaolei, L. (2018). Traffic sign detection and recognition via transfer learning, *2018 Chinese Control And Decision Conference (CCDC)*, pp. 5884–5887.

Yildiz, G. and Dizdaroğlu, B. (2021). Convolutional neural network for traffic sign recognition based on color space, *2021 2nd International Informatics and Software Engineering Conference (IISEC)*, pp. 1–5.