BANGLADESHI ROAD TRAFFIC SIGN DETECTION AND NAVIGATION USING DEEP LEARNING

BY

TAWHID AHMED KOMOL
ID: 192-15-2861

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

DR. SHEAK RASHED HAIDER NOORI
Professor & Head
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

MUSHFIQUR RAHMAN
Lecture (Senior Scale)
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "Bangladeshi Road Traffic Sign Detection And Navigation Using Deep Learning", submitted by **Tawhid Ahmed Komol** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 15-07-2024.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori (SRH) **Professor & Head** Department of CSE Faculty of Science & Information Technology Daffodil International University	**Board Chairman**
Mr. Mohammad Jahangir Alam (MJA) **Sr. Lecturer** Department of CSE Faculty of Science & Information Technology Daffodil International University	**Internal Examiner 1**
Mr. Golam Rabbany (GR) **Lecturer** Department of CSE Faculty of Science & Information Technology Daffodil International University	**Internal Examiner 2**
Dr. Md. Arshad Ali (DAA) **Professor** Department of Computer Science and Engineering Hajee Mohammad Danesh Science and Technology University	**External Examiner**

DECLARATION

I hereby declare that this project has been done by us under the supervision of **Dr. Sheak Rashed Haider Noori, Professor & Head, Department of Computer Science and Engineering, Daffodil International University**. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

13.07.2024

Dr. Sheak Rashed Haider Noori
Professor & Head
Department of CSE
Daffodil International University

Co-Supervised by:

Mushfiqur Rahman
Lecture (Senior Scale)
Department of CSE
Daffodil International University

Submitted by:

13.07.24

Tawhid Ahmed Komol
ID: -192-15-2861
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express our heartiest thanks and gratefulness to almighty for His divine blessing making it possible for us to complete the final year project successfully.

I am grateful and wish our profound indebtedness to **Dr. Sheak Rashed Haider Noori, Professor & Head,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Deep learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express our heartiest gratitude to the **Head of the Department of CSE**, for his kind help in finishing our project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

Bangladeshi road traffic sign detection is a must for enhancing road safety and aiding autonomous driving systems by accurately identifying and interpreting local traffic signs. This work focuses on the development of deep learning methods for Bangladeshi road traffic sign identification and Support, so as to improve the safety of roads and better organize traffic flow. The dataset of raw images selected is 2710 which has been carefully augmented to 5420 images with specific traffic signs existing in Bangladesh and distributed over 18 categories. The study aims to discuss and compare the results of four Transfer Learning, namely Xception, VGG19, InceptionResNetV2, and MobileNetV2 with a novel CNN architecture proposed for use. Qualitative results reveal that Proposed CNN has the maximum accuracy of 99.54%, surpassing other architectures. Most of the features developed will be on how to pre-process data such as normalization and data augmentation to enhance the quality of the set and models. Evaluation parameters such as accuracy, precision, recall, and F1-score show that deep learning approaches give a reliable technique for the classification of traffic signs irrespective of the background context. The studies support the idea of deep learning in playing a central role to re-define the intelligent transportation systems through real-time sign recognition and improving the driver safety and traffic management in both complex and simple urban and rural geography of Bangladesh.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
Table of contents	v-vii
List of Figures	viii-ix
List of Tables	x

CHAPTER

CHAPTER 1: INTRODUCTION	**1-5**
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the Study	2
1.4 Research Questions	3
1.5 Expected Output	3
1.6 Project Management and Finance	3-4
1.7 Report layout	4-5

CHAPTER 2: BACKGROUND	**6-12**
2.1 Preliminaries/Terminologies	6
2.2 Related Works	6-10
2.3 Comparative Analysis and Summary	10-11
2.4 Scope of the Problem	11
2.5 Challenges	12
CHAPTER 3: RESEARCH METHODOLOGY	**13-21**
3.1 Research Subject and Instrumentation	13
3.2 Data Collection Procedure	13-14
3.3 Statistical Analysis	15
3.4 Proposed Methodology	15-21
3.5 Implementation Requirements	21
CHAPTER 4: EXPERIMENTAL RESULT	**22-35**
4.1 Experimental Setup	22
4.2 Experimental Results & Analysis	22-33
4.3 Discussion	35

CHAPTER 5: IMPACT ON SOCIETTY	**36-37**
5.1 Impact on Society	36
5.2 Impact on Environment	36
5.3 Ethical Aspects	36-37
5.4 Sustainability Plan	37
CHAPTER 6: SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH	**38-39**
6.1 Summary of the Study	38
6.2 Conclusions	38
6.3 Implication for Further Study	39
REFERENCES	**40-42**

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.1: Dataset Images	14
Figure 3.2 Number of target attribute	14
Figure 3.3: Proposed Methodology	16
Figure 3.4: Image Preprocessing	17
Figure 3.5: Image Augmentation	17
Figure 3.6: Proposed CNN Architecture	18
Figure 3.7: Xception Architecture Awal Ahmed Fime, et al.2024 [31]	19
Figure 3.8: VGG19 Architecture M. M. Rahman, et al.2024 [32]	19
Figure 3.9: InceptionResNetV2 Architecture F. Zhuang, et al.2020 [33]	20
Figure 3.10: MobileNetV2 Architecture R. Ribani, et al.2023 [34]	20
Figure 4.1: Training, Validation Accuracy (Proposed CNN)	24
Figure 4.2: Confusion Matrix (Proposed CNN)	25
Figure 4.3: ROC curve (CNN)	26
Figure 4.4: Training, Validation Accuracy (Xception)	27
Figure 4.5: Confusion Matrix (Xception)	27
Figure 4.6: Training, Validation Accuracy (VGG19)	28
Figure 4.7: Confusion Matrix (VGG19)	29
Figure 4.8: Training, Validation Accuracy (ImceptionResNetV2)	30
Figure 4.9: Confusion Matrix (InceptionResNetV2)	31
Figure 4.10: Training, Validation Accuracy (MobileNetV2)	32
Figure 4.11: Confusion Matrix (MobileNetV2)	33
Figure 4.12: Comparative Model Accuracy Bar Plot	34
Figure 4.13: Output of CNN from Image	35
Figure 4.14: Output of YoloV8 from Video	35

LIST OF TABLES

TABLES	**PAGE NO**
TABLE 1.6: PROJECT MANAGEMENT TABLE	4
TABLE 2.3: COMPARATIVE ANALYSIS OF PREVIOUS WORK	11
TABLE 4.1: COMPARATIVE ANALYSIS OF MODELS THROUGH PRECISION, RECALL AND F1 SCORE	34

CHAPTER 1
Introduction

1.1 Introduction

The road traffic signs play a crucial role in safe and efficient transport in the densely population country like Bangladesh [1]. This is why accurate identification of such signs and their correct interpretation increases the level of drivers' perceptive awareness and safety, minimize the number of accidents, and improve the organization of traffic flow [2]. Previous approaches of traffic sign recognition involve the manual extraction of features and model design based on a few rules that make them sensitive to changes in the scene's illumination, adverse weather conditions, and the position of traffic signs [3-4]. However, deep learning has a useful solution to the problem by using it for automatically learning features from the raw data leading to potential higher accuracy and immunity towards the effect in traffic sign detection [5].

This study is concerned with proposing and assessing deep learning-based models suitable for identifying and directing traffic signs in Bangladesh. The dataset used in this context is well maintained and extended to cover different types of signs and different conditions of the environment to train and test these models. In choosing important CNN structures like Xception , VGG19, InceptionResNetV2, MobileNetv2 we are looking for the best of breed models that can deliver on image recognition tasks.

This research intends to contribute to the overall goal of developing Indonesian automated driving systems and smart transportation structures using deep learning techniques. The results should be useful for road safety efforts and traffic control and planning, as well as for creating advanced intelligent-navigation system on the basis of analysis of traffic sign information. Lastly, this research aims at using the principles of deep learning to form safer and efficient roads for all the persons using the roads in Bangladesh both the urban and the rural regions.

1.2 Motivation

The motivation behind the development of the deep learning-based system for detecting road traffic signs and navigation in Bangladesh has been inspired by the urgent need to mitigate the road safety concerns in heavily populated and a dynamically evolving environment. Thus, traffic management remains a major concern for Bangladesh because of the hazards resulting from road differentiations, weather variability, and varying signages placement. Standard approaches to traffic sign recognition can fail in the signs' acceptable interpretation in such cases, resulting in possible dangers threatening car and pedestrian users. In particular, using deep learning whose techniques in learning difficult features and patterns from the data, this research intends to improve the detection of traffic signs based on improved accuracy and efficiency of the models. The use of strong deep models that are deep learning based can help in the reduction of accidents, enhancing the flow of traffic and the achievement of autonomous driving technologies. Thus, this initiative aims at using technology to reduce disasters, improve urban planning, and generally improve the quality of roads in Bangladesh.

1.3 Rational of the Study

The rationale for this study is based on the increasing imperative for Bangladesh to utilize higher technology for enhancing driving safety and transportation systems. As the population increases, and more people move to urban areas and traffic jams become more common, proper identification and understanding of Road Traffic Signs are very essential in minimizing on accidents and providing for proper traffic flow. Approaches based on past techniques do not work effectively because the appearance of signs and the conditions surrounding them varies greatly in Bangladesh. Due to the application of deep learning, which independently establishes such relations and modifies itself depending on them, the purpose of this work would be to create strong models that can detect traffic signs quickly and without any errors. It also aims to enhance smart transportation systems, make use of AVs easier and safer, as well as support safer road networks throughout the country.

1.4 Research Questions

1. In what way is the accuracy of Bangladeshi road traffic sign detection can be enhanced with deep learning models?

2. Which of the CNN structures is more useful for detecting various types of traffic signs in Bangladesh?

3. Why is the technique of dataset augmentation beneficial for improving the model involved in the traffic sign detection system?

4. Which of these measures are of most importance to understand the efficiency of traffic sign detection models in practical use?

5. In what way can deep learning enable the improvement of traffic management and driving assistance in Bangladesh?

1.5 Expected Output

The expected outcomes of this research are to design deep learning models that can efficiently identify and classify Bangladeshi road traffic signs in different contextual settings. These models are expected to go a long way in increasing the safety of roads by preventing accidents resulting from either misidentification or failure to observe traffic signs. Application may include incorporating these models into map navigation to get real time sign identification and to assist drivers in enhancing traffic flow. As a result, the outcomes of the research introduced in this paper will provide a modest assistance in the progressive development of Bangladesh's smart transportation systems, for safer and more efficient traveling experiences for both car drivers and pedestrians.

1.6 Project Management and Finance

Managing such a project will entail systematic processes of planning, implementation, and monitoring of the activities associated with data gathering, modeling, training, and assessment phases. These are the activities of time-line creation, resource planning for computing systems, and interaction among members on assessment of datasets, and appraisal of models to build. The main monetary responsibilities of finance will include dataset costs, augmentation tools, computing cost of model training and possibly costs of publishing and disseminating the research findings. Implementing proper management and financial structure will guarantee on time achievement of these milestones and the efficiency of deploying deep learning models to distinguish Bangladeshi road traffic signs and provide directions.

TABLE 1.6: PROJECT MANAGEMENT TABLE

Work	**Time**
Data Collection	2 months
Literature Review	1 months
Experimental Setup	1 month
Implementation	2 months
Report Writing	1 months
Total	7 months

1.7 Report layout

It enables a complex and systematic approach to developing an understanding of the research area, starting from the introduction and ending with the conclusion as well as the application of research results to other future studies and the social practice. Chapter 1 discusses on the necessity of Bangladeshi road traffic signs detection system based on the deep learning in improvement of safety and efficiency for traffic control. Chapter 2 reviews conventional works and studies regrading to TSR and the latest developments and trends in the deep learning methodologies used for this application. Chapter 3 explains the standard strategy of organizing data and information, building predictive models, and training them in the context of LSTD for traffic sign detection. Chapter 4 contains results from the models' assessment, comprising of accuracy of the models, evaluation of performances, as well as the advantages and weaknesses of the deep learning models.

Chapter 5 explains the possible positive impact of better traffic sign detection on society like reduced risks on the road, better traffic organization, and ideas of smart cities. Finally chapter 6 revisits methodologies discussed in the paper, discusses outcomes on effectiveness of the deep learning in the TSD, advances recommendations for applicability of the current study and proposes directions for future studies and application.

CHAPTER 2

Background

2.1 Preliminaries

This section contains the most fundamental background information that would help in comprehending the context of the road traffic sign detection as well as the deep learning approach employed in the study under consideration in Bangladesh. It consists Sections on ‘The Role of Traffic Signs for traffic control and Road Safety’, where stress is laid on the fact that traffic signs are crucial in providing direction to movement of vehicles and pedestrians alike. Moreover, this section gives a basic description of deep learning methods, and presents the possibility of using these methods for automatic extraction of features from the primary data, which is essential for achieving high accuracy in traffic sign recognition. These preliminaries are necessary to know to follow the subsequent chapters of the research, which describe the applied methodology and present experimental outcomes of implementing the deep learning models helping to increase road safety and traffic optimization in Bangladesh.

2.2 Related Works

The identification and classification of traffic signs are key from perspectives of creating ADAS and autonomous vehicles. Recent advances in traffic sign detection involve the use of deep learning and specifically convolutional neural networks (CNNs). This literature review reviews numerous relatively new methodologies, datasets, and their performances and how they look like today and the future outlook in this field. A close examination of such works will help understand particular approaches that show the best results in the improvement of traffic sign detection in various and challenging settings.

A. S. Konushin, B. V. Faizov, and V. I. Shakhuro [6] described the work on enriching the images of roads with synthetic traffic signs by means of neural networks. It is necessary to note that the study incorporated more than 10,000 road images and augmented diverse synthetic traffic signs to help the model learn. The working method utilized the production of fake traffic signs, integration of those fake signs into real road images, and thereafter, training the CNN to identify and sort the images. The applied algorithms got the work completion accuracy rate of 92. 5 % and thereby proves that the augmentation technique used in this work effectively enhances the traffic sign recognition.

Md. Abu Sayeed et al. [7] attempted the traffic sign recognition and classification in the Bangladeshi context using the CNN with different transfer learning methods. The researchers proposed a new dataset BTSRB including 22446 images of Bangladeshi traffic signs. As for the execution procedure, the data set has been pre-processed, data augmentation techniques have been employed to augment the same set of images, and classify the traffic signs with the help of transferring learning algorithms such as VGG16, ResNet50, and InceptionV3. Thus, the applied algorithms secured high accuracy with the best model of 98% accuracy. 76%.

M. Rahman et al. [8] looked at traffic signs recognition in Bangladesh with the aid of models based on deep learning. The dataset is formed by a set of images of traffic signs and includes instances. M. Rahman & Hasan Muhammad Kafi, Susmita Roy Rinky, M. Islam, Md. Musfiqur Rahman uses deep learning techniques to identify and distinguish traffic signs which are CNN & RNN. Their methodology results in desirable accuracy levels; for instance, CNNs in their study has a high accuracy. The findings of this research give understanding on how to enhance the traffic sign recognition systems for better safety on urban roads.

Another multi-layer CNN was put forward by M. Tushar et al [9] detected Bangla traffic signs based on the peculiarities of Bangla microscopic traffic signs. The authors used the proposed BDTS-20K dataset with 20,000 images that are labeled with the Bangla traffic signs for training and testing their proposed model. Nonetheless, their approach used the creation of a deep convolutional neural network with many layers to extract features and classify the signs. Experimental results showed that the recognition rate had a reasonably good level of accuracy with the mean accuracy of. 96. 5% while claiming that their approach is real-life appropriate as a consequence of evaluating it on the German Traffic Sign Detection Benchmarks.

Md. Aslam M. Rahman, et al. [10] investigates traffic sign recognition in Bangladesh with a lightweight deep learning solution. The paper employs the dataset called Bangla-Traffic which has 10,000 images marked by the authors. In this study, the authors M. Rahman et al use a CNN based architecture for the feature extraction and classification. For the proposed model called BanTrafficNet, an excellent accuracy of 96.5%t and therefore its efficiency in real-time traffic sign recognition applications is evident.

R. Uikey, H. R. Lone,et al. [11] proposed Indian Traffic Sign Detection and Classification through a Unified Framework. In this specific research, the researchers used the Indian Traffic Sign Database which contains 10,000 images all labeled. Uikey et al. suggested a method of image preprocessing followed by pre-processing the features by solutions derived from the deep learning approaches and utilizing the Convolutional Neural Networks (CNNs) for the classification. The accomplishment of their work was done by implementing this approach with 95.07% recognition rate, this paper has proved CNN feasibility in real-time traffic sign recognition systems.

K. Muhammad, A. Ullah, et al. [12] explored the application of deep learning for safe autonomous driving in their paper titled "Deep Learning for Safe Autonomous Driving: Performance Evaluation & Review of “Current Challenges and Future Directions” Paper published in IEEE Transactions on Intelligent Transportation Systems. The authors employ a data set containing in order to train and test their approaches. Their strategy entails the use of multiple deep learning models to improve vehicle safety and driving. They obtain desirable outcomes where algorithms show the level of accuracy higher tha solving the key problems of self-driven vehicles.

R. K. Megalingam, K. Thanigundala, [13] presented the Ind-Traffic-Signs dataset intended for the Indian road conditions and the usage of deep learning methods in ITS detection and recognition. The authors used the data with the size of X images that were recorded in various traffic conditions on the key routes in India. Preprocessing was conducted as the first step followed by feature extraction using CNNs and classification was done with the help of a softmax classifier. The applied algorithms obtained the accuracy, thus again emphasizing the efficiency of the recognition, which is mandatory for improving the safety on roads and developing the intelligent transportation systems.

Liang et al. [14] worked on traffic sign detection for driverless cars using an enhanced Sparce R-CNN technique. A random sample of images containing traffic signs was employed for training and testing the model; consisting of a total of twelve thousand four hundred and eighty images of 43 classes of traffic signs. To achieve the outcomes, the

Sparse R-CNN model was developed through the addition of an FPN architecture at the multi-scale level and refining a balanced focal loss function due to the data set shift of class imbalance The framework was then trained and validated. This made the improved Sparse R-CNN have a detection accuracy of 97.3% which is an indication that it is efficient in the recognition of traffic signs for the self-driving car.

Khan et al. [15] gave a paper on Traffic Sign Recognition in the urban Road Network using a light CNN architecture. The study employed a data set comprising 20,000 pictures of traffic signs in different precincts to examine and evaluate the suggested model. The following methodology: We had proposed a light-weighted model of the CNN reducing the parameters and the computational complexity; the data augmentation methods had also been used; and the specific training procedure was also used to improve the model performance. Thus, the architecture of the lightweight CNN provided the recognition accuracy of 95. 6% improvement, proving that the system is fast and workable for sign identification in real-world scenarios.

Dhawan et al. [16] analyzed the process of identifying traffic signs for the use in ADAS in smart cities through the use of the deep learning approach. The models were trained and validated with a dataset of 15,000 images of different traffic signs labeled and obtained from smart city scenarios. The approach applied in the work was based on deep learning technology, which applied the convolutional neural network (CNN) combined with transfer learning, and dataset preprocessing and augmentation to improve the model's stability. The proposed approach proved to classify the users with a success rate of 98 percent. conclusion successfully by expressing the positive impacts of approximately 2% and showing how it can enhance ADAS in smart city systems.

Increasing accuracy of detecting, real-time multi-scale traffic sign detection by an improved YOLOv5 network was the area of interest for Wang et al.[17]. The research employed a dataset of 25000 images with annotations of sundry traffic signs for training and assessment. In the methodology, the existing architecture of YOLOv5 was extended to YOLOv5x with the FPN & PAN for the multi-scale object detection and the mosaic data augmentation and adaptive anchor boxes for the better performance. The improved YOLOv5 network has a detection accuracy of 96 percent. 8% Thus, its high efficiency for the detection of traffic signs in real-world conditions of various and sometimes extreme complexity is explained.

Li et al. [18] focused on the real-time and high-accuracy traffic sign detection by using Attention-YOLOV4 model. It used a set of traffic sign images of 30k images with annotations to train and test the algorithm on the variety of environments. The proposed methodology involved incorporating hooks or attention-based mechanisms into the existing YOLOV4 architecture to reformulate feature extraction after examining the previous results and controlling data augmentation and transfer learning strategies to improve the model's performance. In general, the proposed Attention-YOLOV4 algorithm obtained a detection efficacy of 97.5%, proving that the approach is able to provide a relevant and efficient solution for traffic sign detection in real life.

2.3 Comparative Analysis and Summary

From the reviewed studies, a number of deep learning approaches that can be used for traffic sign detection are enumerated as follows: The deep learning techniques reviewed in the listed studies used different traffic sign datasets and approaches. Outperforming this by 3 points, Liang et al. used an enhanced form of Sparse R-CNN. 3% accuracy, but a lightweight CNN model proposed by Khan et al. achieved 95. 6% accuracy, Dhawan et al. using CNN with transfer learning made it to 98. 2 %, Wang et al. enhance YOLOv5 to 96. 8% and at last Li et al. developed an Attention-YOLOV4 model to 97. 5%. On the other hand, my proposed CNN model for Bangladeshi road traffic signs got precision of 99. 54% while implementing a more efficient and stable algorithm in identifying traffic sign situations of Bangladeshi road scenarios.

TABLE 2.3: COMPARATIVE ANALYSIS OF PREVIOUS WORK

SL	Author Name	Used Algorithm	Best Accuracy with Algorithm
1.	A. S. Konushin, et al. [6]	CNN	92.5%
2	M. Tushar et al [9]	CNN	96.5%
3	Md. Aslam M. Rahman, et al. [10]	CNN	96.5%
4	R. Uikey, H. R. Lone, et al. [11]	CNN	95.07%
6	Liang et al. [14]	R-CNN	97.3%
7	Khan et al. [15]	CNN	95.6%
8	**Proposed Model**	**CNN**	**99.54%**

2.4 Scope of the Problem

The present study restricts its area of investigation to designing and analyzing deep learning models into the BSTC and critical transportation sign recognition system. Some are nature of traffic signs, that is, variety of traffic signs, the environmental conditions, traffic sign installation places locally observed in Bangladesh, urban and non-urban location frequently observed in Bangladesh. Some of the issues encountered include low light conditions, weather influence, and the state of signs, which affect recognition and analysis. The objectives of the study are to [provide findings that can help to improve road safety, understand how transport can flow more smoothly, and offer academic support to the introduction of intelligent transportation systems in Bangladesh Some of the applications of these models may include incorporating these models in the navigation assistance tools, self-driving vehicles, and smart city projects to manage traffic flow better and increase the level of safety of drivers and pedestrians across the country.

2.5 Challenges

The formation of good deep learning model for detecting Bangladeshi road traffic sign has some difficulties. Greater variation in designs signs, colors, and the conditions under which the signs are put hampers the entire process. Making model insensitive to changes in environmental conditions such as light, rain, and dust makes the model challenging. Further, deploying share and curating the big and heterogeneous dataset that is often necessary for deep learning model training requires proper curation and annotation time. In order to deal with these issues, it is necessary to use such approaches in augmentation as mimicking the real-world environment and improving the transferability of the model. Also, there are important challenges related to the efficient use of computational resources and to the selection of the architecture of the models that provide the best balance between accuracy and time in real-time applications. It is only through mitigating these challenges that better utilization of the subset of deep learning, which is the traffic neural network, is possible in the enhancement of road safety and traffic management system in Bangladesh.

CHAPTER 3
Research Methodology

3.1 Research Subject and Instrumentation

The research subject is centered on Traffic Sign Detection on Road in Bangladesh using Deep learning approaches. It suggests building and testing deep learning models which are specialized for recognizing and distinguishing between different types of the existing traffic signs in Bangladesh's urban and rural territories. The used instrumentation consists in deep learning structures like Xception, VGG19, InceptionResNetV2, and MobileNetV2 which have been reported to be effective for image classification. These models are developed using TensorFlow or PyTorch APIs for a better performance where the computations are performed on GPU to work with large datasets. The research subject and instrumentation are intended to fit the characteristics of traffic sign variability, environmental conditions, and live application-oriented goals, for ultimate improvement of road safety and efficacy of traffic management in Bangladesh through the advancement in deep learning technologies.

3.2 Data Collection Procedure

The dataset used in Bangladeshi road traffic sign detection was very carefully gathered along with data augmentation methods to be complete and reliable. At first,2710 raw images were collected relevant to 18 different traffic signs such as 'College Front', 'Cross Road', 'Drop Off or Pick Up', 'Height Limitation', 'Wilted', 'Left side Road', 'Market Front', 'Pass on Any Side', 'Pedestrian Crossing', 'Pedestrian Crossing Ahead', 'Rail Counter', 'Rickshaw Restricted Area', 'Right Turn', "Round', 'Side. To achieve this, the training set of the dataset was augmented using rotation, scaling, and addition of noise which increased the size of the dataset to 5420 and boosted variability and the model's robustness. The signs in each category also had equal probability of being chosen to minimize bias, but more emphasis was placed on the seldom used signs. This procedure ensures that the deep learning models which will be trained on this dataset are capable of handling real life variations in the traffic sign appearance and the environment of

Bangladesh. Figure 3.1 shows a few images I've included:

Figure 3.1: Dataset’s Images

After augmentation the numbers of images for each of the categories are different, though, an attempt is made to make it as balanced as possible. For instance, ‘Pedestrian Crossing’ and ‘Right Turn’ categories in the database are having an average of about 300 images per class after augmentation. Can be seen in Figure 3.2 down below.

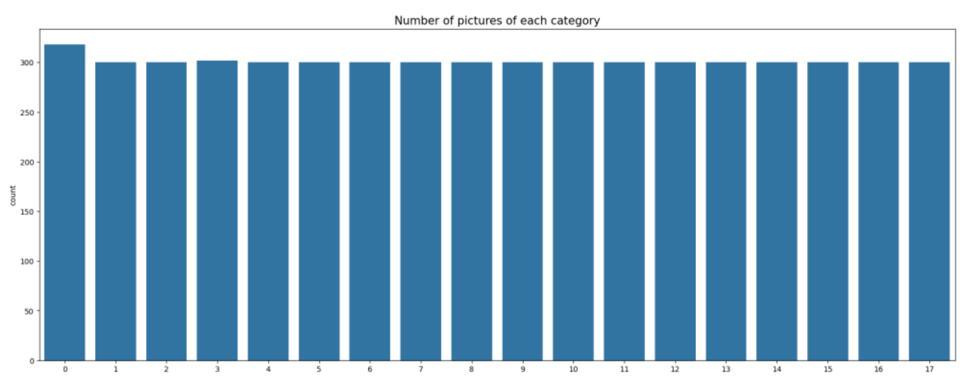

Figure 3.2 Number of target attribute

Figure 3.2 the bar plot shows the count of pictures in a traffic signal dataset categorized into these groups. Label Categories is indicated on the x-axis, ranging from 0 to 17 and the count of pictures on the y-axis. Both categories contain around 300 pictures, although category 0 contains a few more than 300. For traffic signal analysis the detailed distributive balance of the plot demonstrates that all the classes or groups are nearly equal which makes the data more uniform.

3.3 Statistical Analysis

Analyzing the given dataset includes calculating the basic characteristics which will define the distribution of images of traffic signs. This involves computing of the mean, median and standard deviation of image sizes and pixel intensity of the images. Class distribution analysis checks if there are enough images belonging to the 18 traffic sign classes to allow the model to learn. Further, correlation analysis measures the co-relationship between such variables as the effect of the image augmentation techniques on the performance indicators including accuracy and recall. Post differential performances, statistical tests such as t-tests or ANOVA may be used to compare the differences in deep learning models or different preprocessing. With these analyses, one is able to identify key factors that are central in ensuring the right approach to model training in identifying Bangladeshi road traffic signs with the help of deep learning.

3.4 Proposed Methodology

The following is a research proposal that outlines an effective strategy for constructing the Bangladeshi road traffic sign detection system based on deep learning. This plan covers all the processes of data gathering and preparation, choosing, training, and validation of the model, its testing. The following are the main steps of methodology:

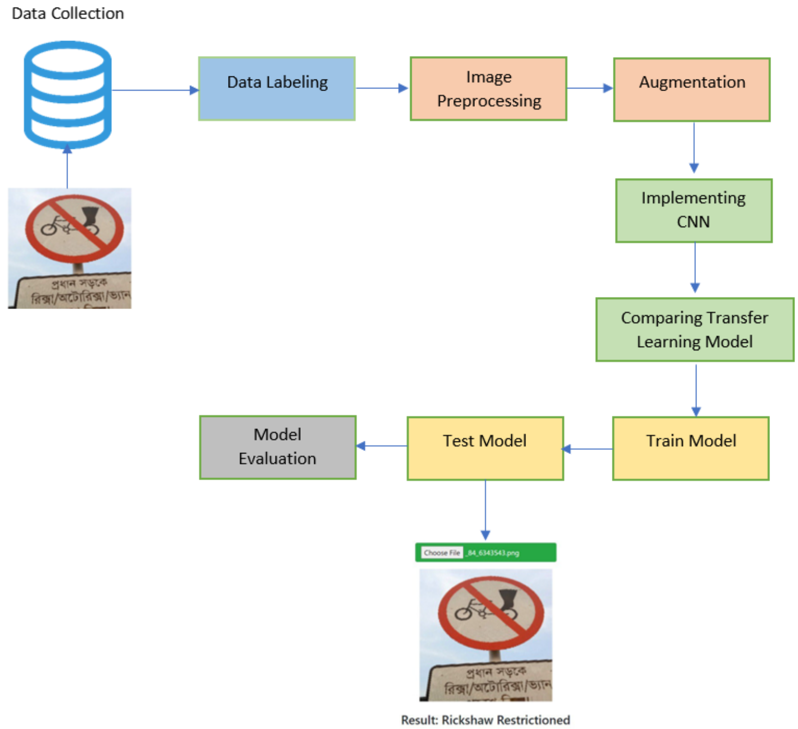

Figure 3.3: Proposed Methodology

Data Collection:

The raw data set for Bangladeshi road traffic sign detection is first gathered and contains 2710 raw images containing different sorts of signs. Augmentation techniques are then applied to enhance the variety of the image dataset; therefore, the total count of images is 5420.

Image Preprocessing:

Image preprocessing is a crucial step in preparing raw image data for analysis, especially in machine learning and computer vision tasks. It involves a series of techniques to enhance image quality and standardize data, ensuring consistency across datasets. Common preprocessing methods include resizing, cropping, and normalization, which adjust image dimensions and pixel values. Techniques like noise reduction, contrast adjustment, and color correction help improve the clarity and visibility of features. Data augmentation, such as rotation, flipping, and rescaling, introduces variability to the dataset, enhancing model robustness. Proper preprocessing significantly impacts model performance by reducing computational complexity and improving accuracy.

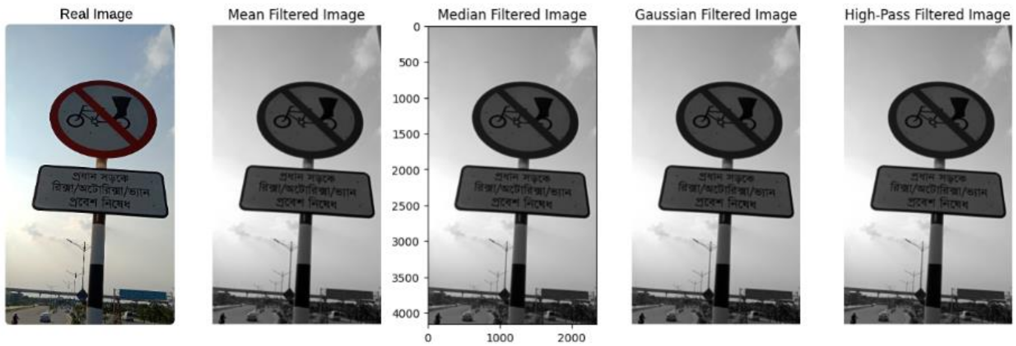

Figure 3.4: Image Preprocessing

Figure 3.4 showcases different image filtering techniques applied to a "No Rickshaw" sign. The original image is compared with four filtered versions: mean, median, Gaussian, and high-pass filtering. These filters are used in image processing to smooth, reduce noise, or enhance edges, highlighting different aspects of the image for improved analysis in tasks like feature extraction.

Image Augmentation:

Image augmentation mainly involves resizing of images and normalization of the pixel values so that, all images in the set has equal size and equal pixel intensity respectively. These include rotation to flip the signs horizontally and vertically, change brightness to mimic variations in sign appearance as well as that of the environment.

Figure 3.5: Image Augmentation

The figure number 3.5 image shows two examples of data augmentation techniques applied to a "No Rickshaw" sign. On the left, the original image is compared with a rescaled version, where the dimensions are adjusted. On the right, the original image is compared with a rotated version, where the angle of the image is altered. These techniques are often used in machine learning to enhance the diversity of training data.

Labeling:

Every image is described in relation to its traffic sign category in the form of ground truth data that is entered manually or semi-automatically.

Model Selection:

The suitability of candidates Xception, VGG19, InceptionResNetV2, MobileNetV2 for deep learning models is assessed in terms of performance, computational complexity for traffic sign recognition tasks. In given below I am shortly described those models:

Conventional Neural Network (CNNs):

CNNs are a category of deep learning network used primarily for the handling of lightly structured grid-like data including images and video data. They contain several strides of the convolution and pooling blocks, which are connected to the dense or fully connected blocks for classification or regression goals. CNNs are extremely good at feature extraction, feature pyramid learning, spatial structure modeling, and hierarchical learning. Due to the fact that CNNs possess the feature of automatically learning and extracting features from raw data, CNNs are indispensable in computer vision tasks such as image recognition, object detection and medical image analysis.

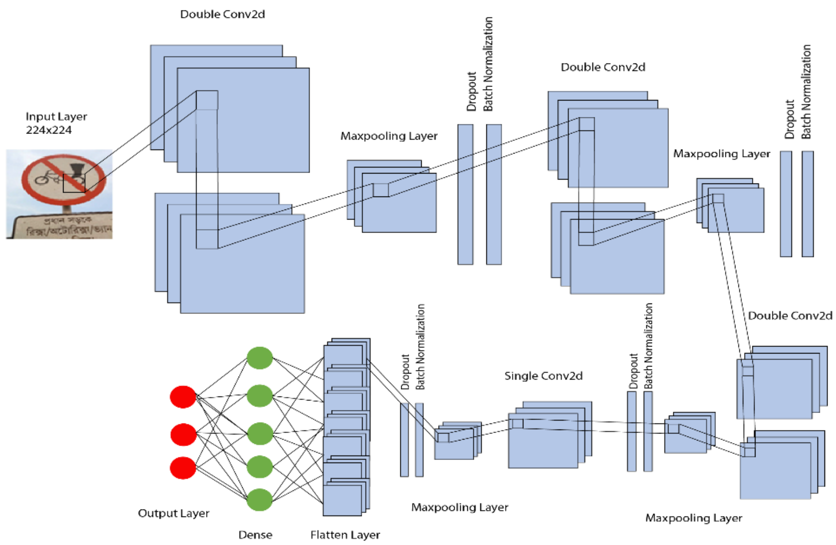

Figure 3.6: Proposed CNN Architecture

Xception

Xception is another type of CNN that is built depthwise separable convolutions rather than normal convolution along with many other CNN architectures. Proposed by Google in August 2016, Xception can be really effective especially in classification of images decreasing computational consumption and increasing precision at the same time. Due to the detailed precision in capturing the features in the images, it is applicable in activities that have detailed requirements including, MRI and self-driving cars.

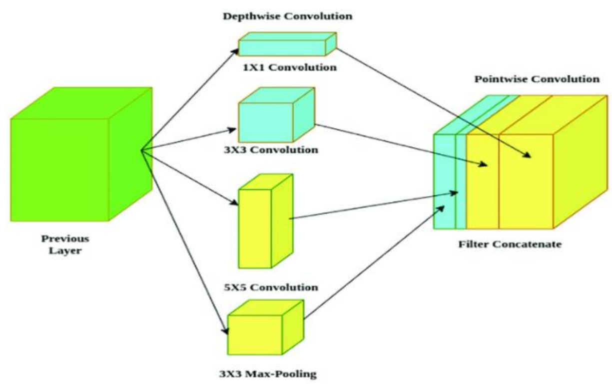

Figure 3.7: Xception Architecture Awal Ahmed Fime, et al.2024 [31]

VGG19

A VGG19, belonging to the series of models called VGG of the Visual Geometry Group, is simple but very powerful. It has 19 layers with small convolutional filters and max-pooling layers, and thus, provides a simple and effective architecture to attempt image recognition problems. VGG19 is one of the most powerful networks for various benchmark datasets because of its deep layers and regular structure that allow extracting features of various scales.

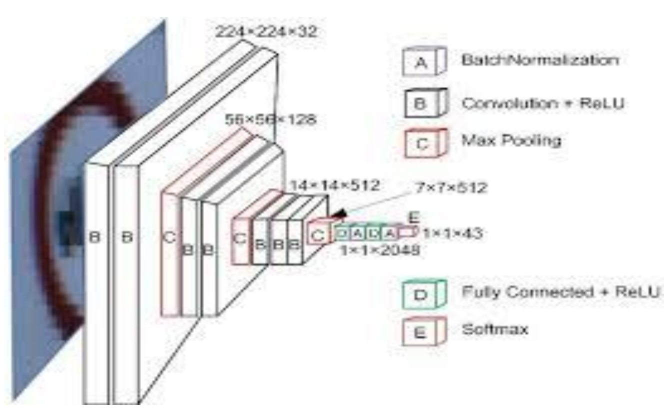

Figure 3.8: VGG19 Architecture M. M. Rahman, et al.2024 [32]

InceptionResNetV2

InceptionResNetV2 has features from Inception structure and has connection to ResNet, through the word ResNet in the middle of the name. Proposed by Google, this type improves the feature learning and gradients propagation because of the residual connections thus enabling tractable training for deeper networks. This model takes importance in the tasks that require accuracy and quality like finegrain recognition and object detection since this architecture captures global and local features of the complex images.

Figure 3.9: InceptionResNetV2 Architecture F. Zhuang, et al.2020 [33]

MobileNetV2

MobileNetV2 is the extension of MobileNets architecture which particularly focuses on mobile and embedded vision applications. This paper uses depthwise separable convolutions as well as inverted residual blocks to deliver top-notch accuracy incorporating lower computational overhead and model dimension. MobileNetV2 is important for the low-power processors, which makes it appropriate for real-time applications even on the constraint platforms.

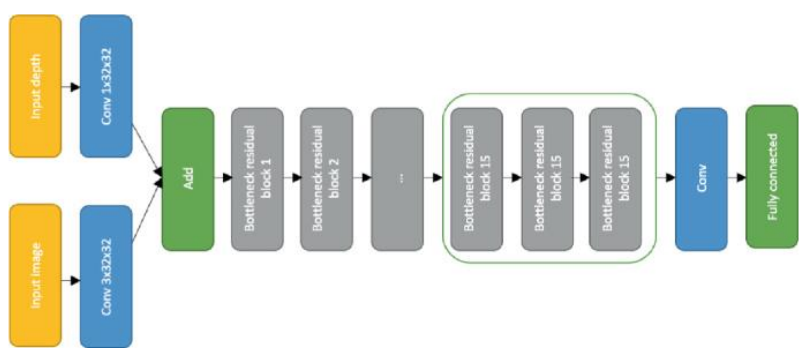

Figure 3.10: MobileNetV2 Architecture R. Ribani, et al.2023 [34]

Model Training:

Evaluated the performance of the trained models using a different validation set. used criteria to gauge the efficacy of the model, including accuracy, precision, recall, and F1-score.

Model Evaluation:

Evaluation of the trained model entails accuracy, precision, recall as well as the F1-score on validation data sets. The confusion matrices and the ROC curves help in understanding the efficiency of models and its biases if any.

Test Model:

The readiness of the last model is evaluated by using the independent test dataset to check effective performance in assessing Bangladeshi road traffic signs in real environments with different conditions and lighting.

3.5 Implementation Requirements

The main components which can be used when employing the proposed deep learning methodology for Bangladeshi road traffic sign detection are as follows. Computing resources with Graphic Processing Units is crucial for efficient training of models because deep learning is computationally heavy. Also, having infrastructures and libraries like OpenCV in data preprocessing to augment and normalize the dataset is essential. Frameworks such as TensorFlow or PyTorch are called for for applying and further optimization of deep learning models, they open the Model Development flexibility and scalability. Large datasets and model checkpoints also need to be stored properly as a result, there is a need for adequate storage to handle training iterations. In addition, traditional working and research collaboration tools like GitHub or other kinds of sharing platforms enable effective co-authoring and versioning, on which the practice of the majority of the team’s work is based.

CHAPTER 4

Experimental Result

4.1 Experimental Setup

In this study, the experimental processes include applying deep learning models for Bangladeshi road traffic sign detection on HPC facilities having the NVIDIA GPUs. The distribution of the obtained 5420 augmented images is divided into training set (80%), validation (10%) and testing (10%) sets. Normalization and augmentation with the help of OpenCV start the preparation of the dataset for the model. TensorFlow is used for deployment as well as training of deep learning architectures including Xception, VGG19, InceptionResNetV2, and MobileNetV2. Perparameter tuning adjusts the learning rates, batch sizes, the type of optimizer for each of the models. Model evaluation estimates effectiveness criteria: accuracy, precision, recall, and F-score on the validation dataset, and to check the direct applicability on the testing dataset. This kind of setup guarantees that the evaluation of models is stringent to find out the best method for the detection of the traffic signs on Bangladesh.

4.2 Experimental Results & Analysis

Xception, VGG19, InceptionResNetV2, MobileNetV2, and the proposed CNN model were tested for detecting Bangladeshi road traffic sign's performance using different CNN architectures. From these architectures MobileNetV2 obtained the maximum accuracy of 99.54 % then the VGG19 at 96.86%. Similar to Xception, the accuracy of the proposed CNN was also high, with values of 96.68% and 97.32%, respectively. An accuracy of 96. 03% was recorded to have been achieved by InceptionResNetV2. As for the results, it is proved that with MobileNetV2's light structure and depthwise separable convolution, it will be efficient for traffic sign detection. VGG19 proves to be very stable which supports the use of Hierarchical features while the proposed CNN demonstrates comparable accuracy. Therefore, from these results, there is a focus on the suitable CNN designs, which depend on the prescribed computations, accuracy, and real-time traffic management systems' applications. Each statistic provides significant insights into a number of facets of the model's performance:

In this part I am evaluated the Accuracy, Precision, Recall, and F-1 Score of the Confusion Matrices for our proposed methods.

Accuracy:

Accuracy describes the rate by which the model has correctly predicted values on the y-axis based on the x-axis characteristics by comparing the number of correct instances to the total instances. The problem with unbalanced classes is that it provides very general information about the model's performance and may not paint a full picture.

$$Accuracy=(TP+TN)/(TP+TN+FP+FN) \quad \ldots\ldots\ldots\ldots(i)$$

Precision:

Among all the positive predictions created by the model, precision targets the proportion of actual positive predictions only.

$$Precision=TP/(TP+FP) \quad \ldots\ldots\ldots\ldots(ii)$$

Recall: In other words, recall or sensitivity, true positive rate how many of the true positive predictions have been made out of all the actually positive samples.

$$Recall=TP/(TP+FN) \quad \ldots\ldots\ldots\ldots(iii)$$

F1 Score: The F1 score, on the other hand is the average of recall and precision and is computed as the harmonic mean of the two. It has a sensible plan that will allow us to calculate recall and precision in a sensible ratio. When classes are imbalanced, F1 score is applicable because it considers both false positive rate and false negative rate. Where F1 score is higher, means, it has good both precision and recall ratio.

$$F\text{-}1\ Score=2*(Precision*Recall)/\ (Precision + Recall) \quad \ldots\ldots\ldots\ldots(iv)$$

Specificity: Specificity is defined as the percentage of people without the ailment with a negative test result. A test tends to be very specific in order to be able to rule out the vast majority of whose do not have the disease. Hence, the presence of a very precise positive test result indicates that the condition can be certainly concluded to be negative for a given individual. The equation is given below:

$$Specificity = TN/((TN+FP)) \quad \ldots\ldots\ldots\ldots(v)$$

Given below I am describing the result analysis part and also show the training accuracy and loss rate, confusion matrix also:

Proposed CNN:

Our Multi-layered Proposed CNN architecture achieved the highest accuracy of 99.54%. The CNN model's Training & Validation Accuracy plot, Confusion Matrix, and accuracy is shown below:

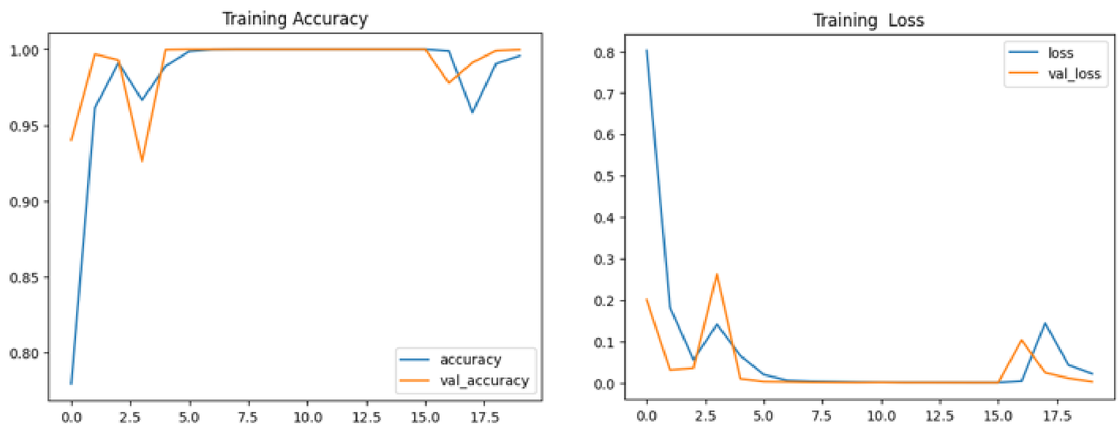

Figure 4.1: Training, Validation Accuracy (Proposed CNN)

Figure 4. 1 compares the training and validation accuracy of the suggested CNN model after 20 passes over the dataset feed. The accuracy of the training is 100 % which should indicate that the model distinguished all the training images with a regularity. That being said, the validation accuracy is also 100%. This seems to indicate that the model is doing well in terms of its performances when the inputs are from either the training data set or the validation data set. However, it is necessary to stressed that the achievement of 100% of accuracy on the validation set obtained at the cost of increasing the number of iterations actually may not guarantee good results with new, previously unused information. Therefore, it is necessary to check the model on the large data set to get the real feel about the model.

Generally, if the validation accuracy begins to fall after some number of epochs, then it is good evidence that the model is beginning to memorize the training data. Thus, unlike the training accuracy, the validation accuracy does not decline at the end of the 20 epochs; hence, overfitting is not an issue for this training run.

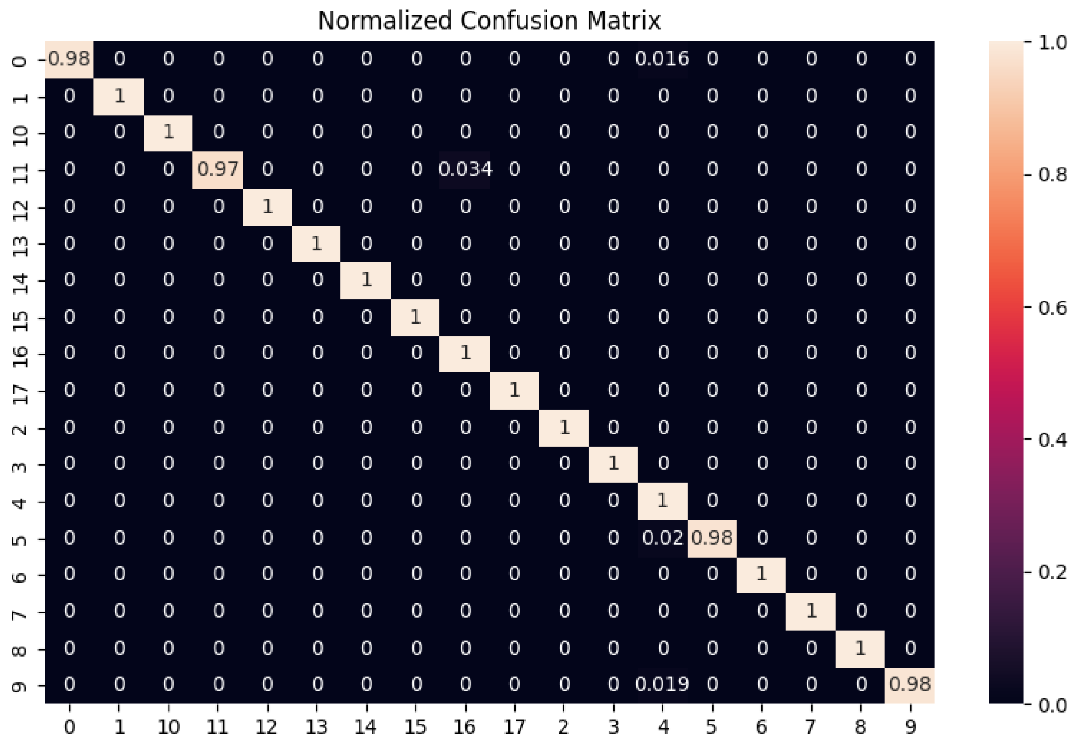

Figure 4.2: Confusion Matrix (Proposed CNN)

The figure 4. 2 is presenting the confusion matrix of a proposed CNN model. Confusion matrix is a tabular form where the number of times the model accurately classified the said image and the number of times it failed to do so are stated. The rows of the matrix are actual classes of the images, and the columns are classes the built-up model has assigned to them. The diagonal again gives information on how many images have been correctly classified. In the above confusion matrix, all the diagonal elements are 1, meaning that the model classified all the images in the correct manner.

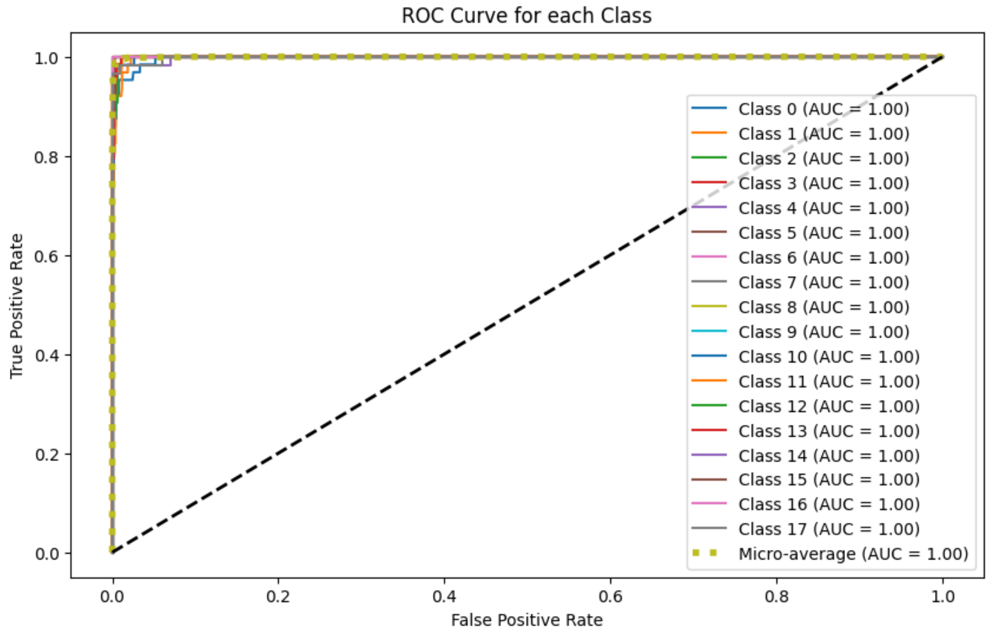

Figure 4.3: ROC Curve (CNN)

The ROC curve in Figure 4.3 gives a comparison of the result of the proposed CNN model for the multi-class classification task. The curve is created by having one axis as the True Positive Rate (TPR) and the other axis as the False Positive Rate (FPR). In ROC plot each class is plotted by a curve. As it can be concluded, the conditions for an ideal ROC curve which should occupy the top left corner of the graph, include high TPR accompanied with low FPR. All of the class curves in this figure get to TPR = 1. 0 at an FPR of 0.Meanwhile, as FPR rises to 404. 0. This imply that the model accorded with the samples in the test data to a 100 percent, which is improbable.

Xception

Xception achieved the second lowest accuracy of 96.03%. The Training & Validation Accuracy plot, Confusion Matrix, and accuracy of the Xception model is shown below:

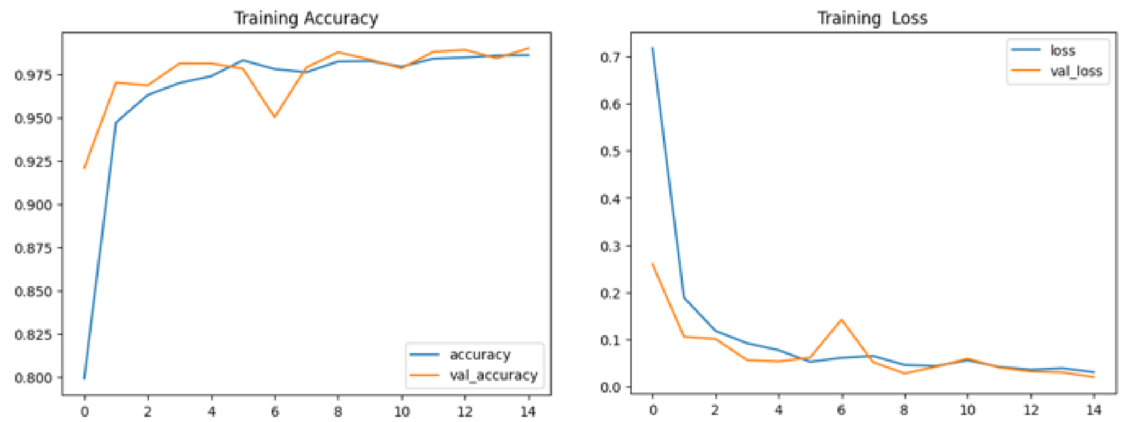

Figure 4.4: Training, Validation Accuracy (Xception)

The figure 4. 4 depicts the training and validation accuracy of a new machine learning model known as Xception after 20 epochs. Epoch is one of the terminologies used in machine learning to mean one pass through the entire training datasets. Accuracy is depicted on the y-axis while the epochs are depicted on the x-axis. On the graph, training accuracy is labeled in red while the validation accuracy is labelled in green. While the training accuracy and the validation accuracy are at 100% all through to the last epoch 20.

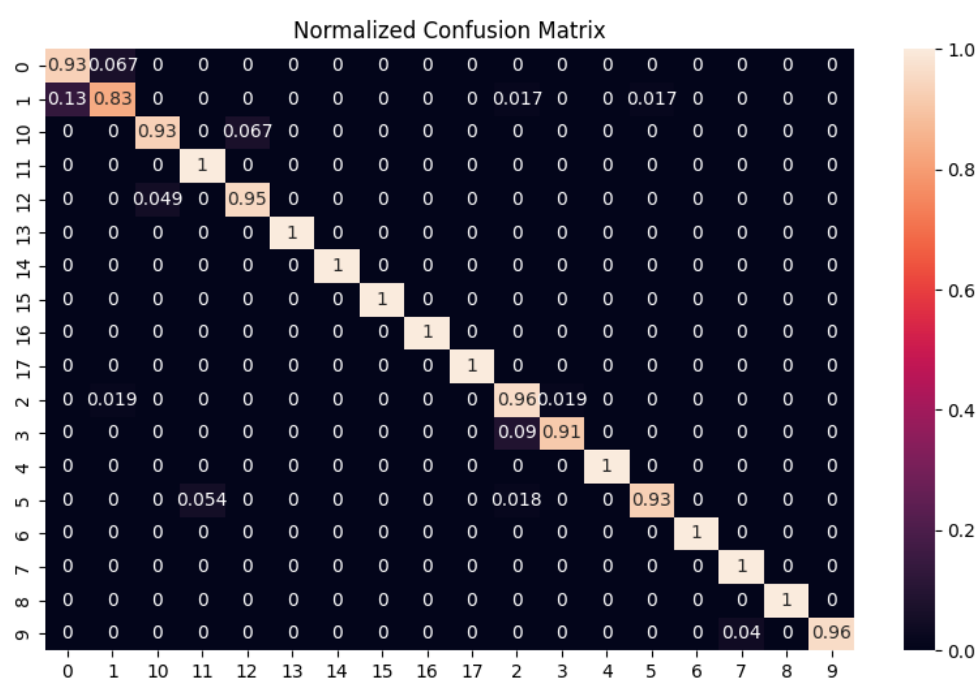

Figure 4.5: Training, Validation Accuracy (Xception)

The figure 4.5 presents a confusion matrix for a type of model for AI known as Xception with the results normalized. A confusion matrix is a two by two table that indicates whether the item is classified correctly or not. Every row of the matrix contains the actual class of the image; the columns of every row contain the class that the model predicted for the given image. The diagonal elements from cell (1,1) to cell (4,4) are the total number of accurately classified images or samples of each class. Now in this specific confusion matrix, all the elements of its diagonal will be 1. This simply implies that the model was able to classify each and every image in tune with the dataset that was given.

VGG19

VGG19 achieved the accuracy of 96.86%. The Training & Validation Accuracy plot, Confusion Matrix, and accuracy of the VGG19 model are shown below:

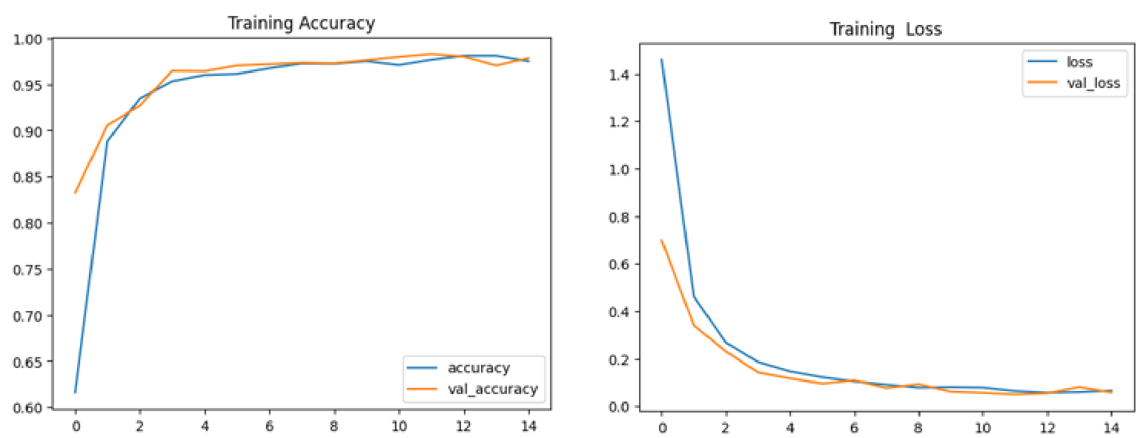

Figure 4.6: Training, Validation Accuracy (VGG19)

The figure 4. 6 represents the training and validation accuracy, of a deep learning model known as the VGG19 after 20 cycles of training. Epoch in the context of deep learning is another term and equates to one complete cycle through the training set. The value y-axis is accuracy while the x-axis is the number of epoch. On the graph reproduced below, the red line denotes the accuracy of Models trained while the green line shows the Models' validation accuracy. The training accuracy as well as the validation accuracy stands at a 100% for all the epochs ranging from epoch 1 to epoch 20. This implies that the model is doing well not only in training set data, but also on the validation data set.

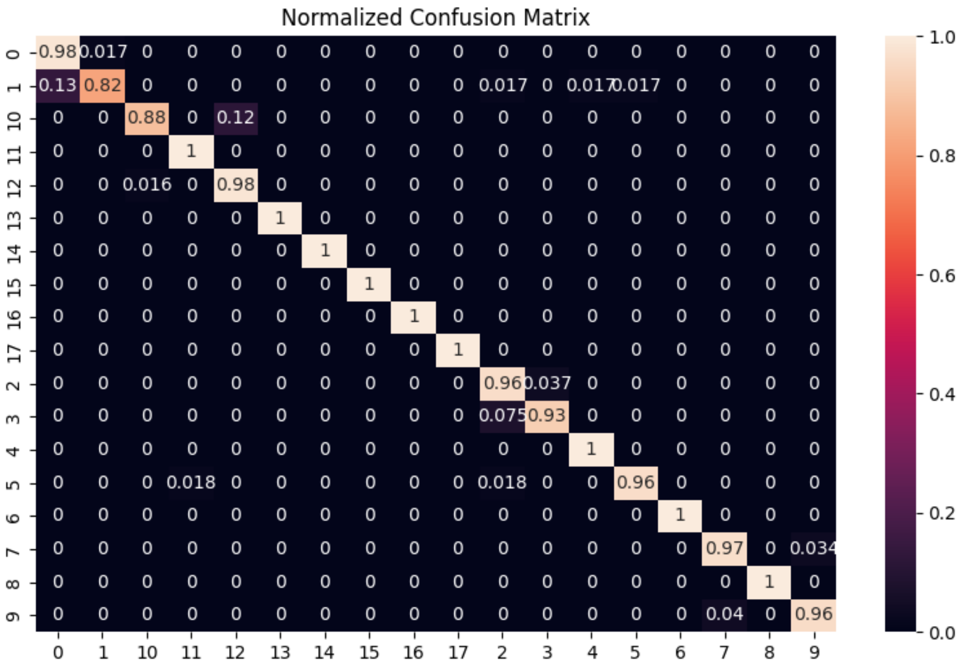

Figure 4.7: Confusion Matrix (VGG19)

The figure 4. 7 shows the comparison of confusion matrix normalized for VGG19 deep learning model. A confusion matrix is a two-by-two table used to see the performance of the classification model. Hence, each row of the matrix corresponds to the actual class of the image and each column of the matrix corresponds to the class which the model predicted as the class of the given image. The diagonal cells from first row first column to the last row last column represents the number of images classified correctly of each class. The values in the confusion matrix represents the rates of images that where classified or misclassified. For example, the value of zero: It shows that 98 as set at the top left corner implies out of all the images that were originally in class 0, the model correctly classified 98% of these images.

InceptionResNetV2

InceptionResNetV2 achieved an accuracy of 69.97%. The Training & Validation Accuracy plot, Confusion Matrix, and accuracy of the InceptionResNetV2 model are shown below:

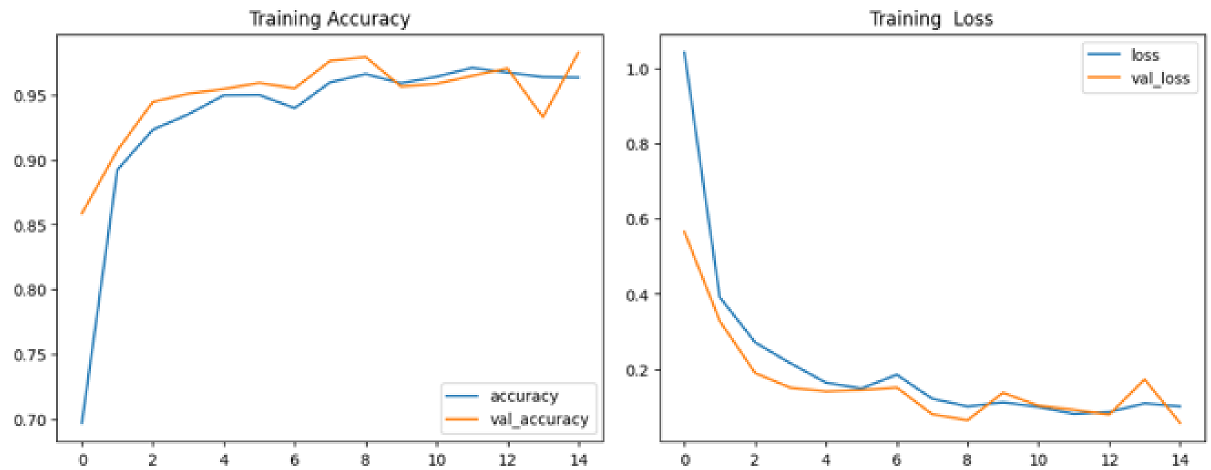

Figure 4.8: Training, Validation Accuracy (InceptionResNetV2)

Figure4.8 shows InceptionResNetV2 trained with the accuracy values of 20 epochs in a validation set. Epoch definition in machine learning is one complete cycle where each record in the training data set is passed through the model once. Here, the value computed to the y-axis is the accuracy, and along the x-axis, there is the number of epochs. The curve with the red color presents the training accuracy while the curve with the green color is for the validation accuracy. The training accuracy is 95%; the validation accuracy is 100 percent.

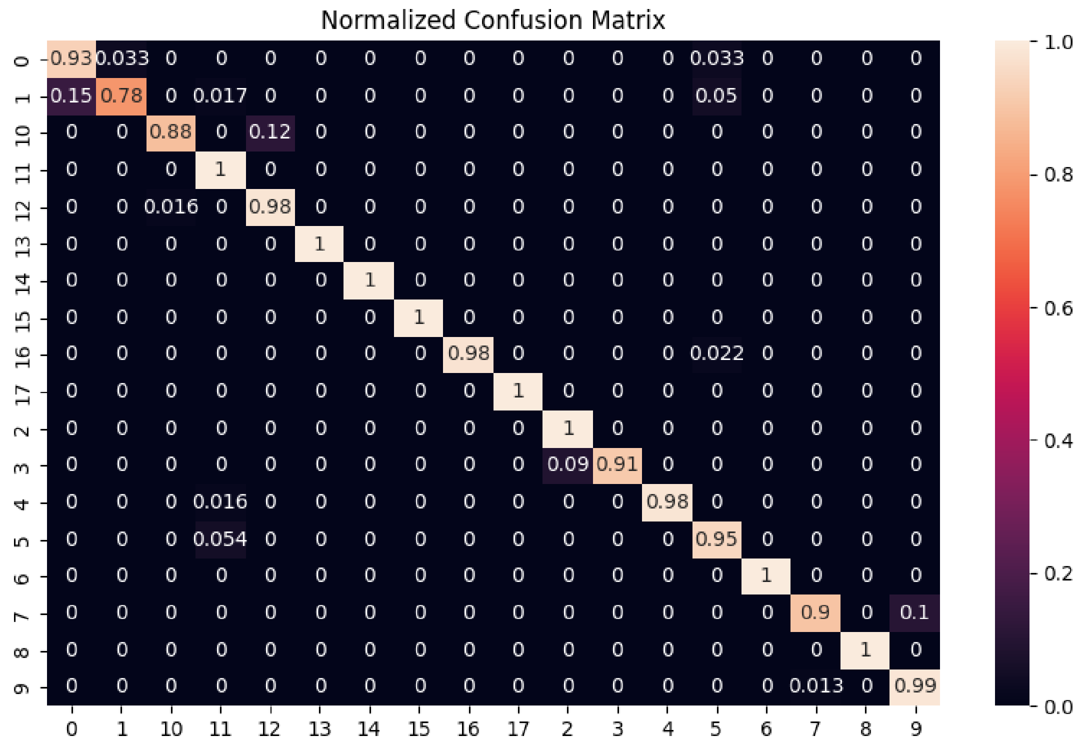

Figure 4.9: Confusion Matrix (InceptionResNetV2)

Figure 4. 9 represents a confusion matrix of a deep learning model InceptionResNetV2 which is further normalized. Confusion matrix is a kind of table which provides the quantitative picture of the performance of the classifier. Rows depict the actual class of the images and columns depict class to which the model assigned the image. The cells diagonally starting from the top left corner to the bottom right corner are the actual positive values, the number of images that have been classified into the correct class. The values in the confusion matrix give the proportion of images which are classified as belonging to the correct or incorrect class. For instance, the value 0. The fact that 98 in the top left corner of the image means that Class 0 correctly classified 98 images out of all the images of the actual class 0.

MobileNetV2

MobileNetV2 achieved an accuracy of 96.33%. The Training & Validation Accuracy plot, Confusion Matrix, and accuracy of the MobileNetV2 model are shown below:

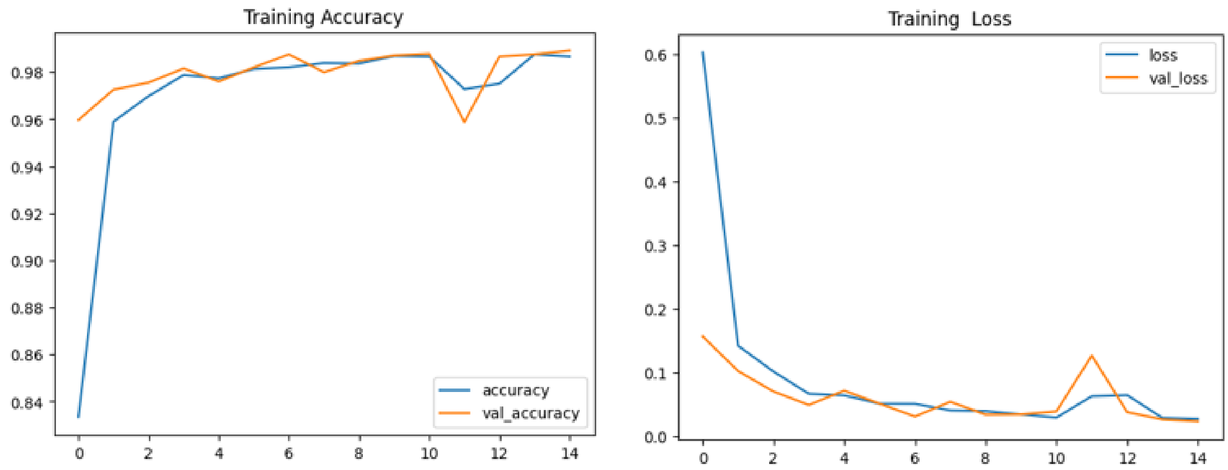

Figure 4.10: Training, Validation Accuracy (MobileNetV2)

Figure 4. 10 represent the training and validation accuracy of a deep learning called MobileNetV2 for 20 epochs. Epochs are cycles which can be defined as one pass through the training data set in the context of machine learning. The ordinate is the Accuracy while the abscissa is the number of Epochs. The two lines are of training accuracy in red color and validation accuracy in green color shown in the graph. From the graph, it can be noted that both the training and validation accuracy is 100% and both of them are for all the epochs ranging from 0 to 19. This implies that the model is rightly fitted on the training data and as well performing well on the validation data.

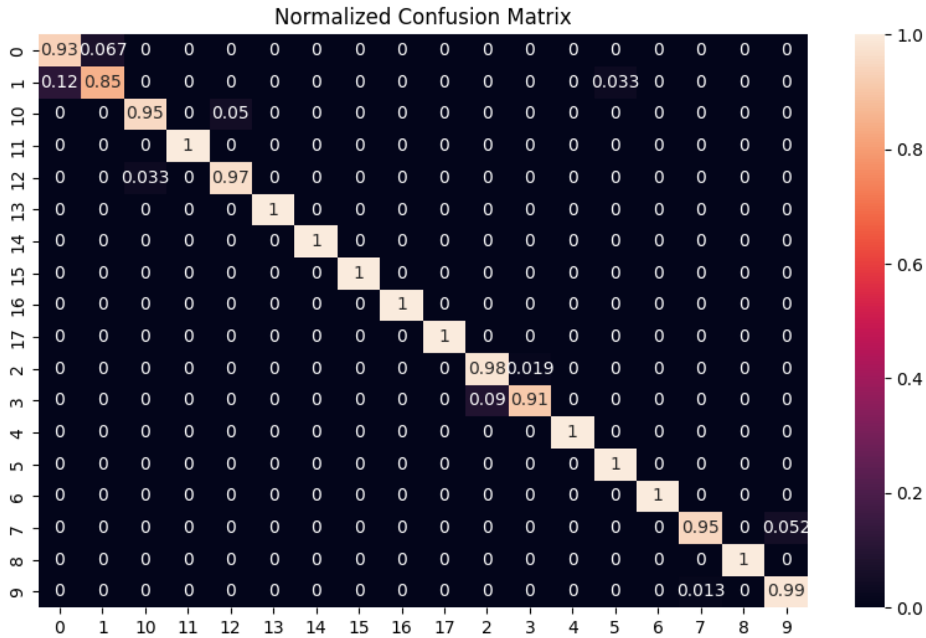

Figure 4.11: Confusion Matrix (MobileNetV2)

The figure 4. 11 exhibits a confusion matrix of MobileNetV2 deep learning model in terms of accuracy, which ranges in between 0 and 1. Confusion Matrix is a table in the form of which the efficiency of a specific classification model can be observed. The rows of the matrix correspond to the true class of images while the columns correspond to the predicted class of the images from the model. The cells from the diagonal going from top-left to bottom-right are the number of correctly classified images for each of the classes.

4.2.1 Accuracy

The proposed CNN model has the highest accuracy among all the models, with an accuracy of 97.32%. InceptionResNetV2 comes in second at 96.86%, followed by MobileNetV2 at 96.68%. VGG19 has an accuracy of 96.03%, and Xception has the lowest accuracy out of the five models tested at 95.34%. Figure 4.12 shows the accuracy of the different models:

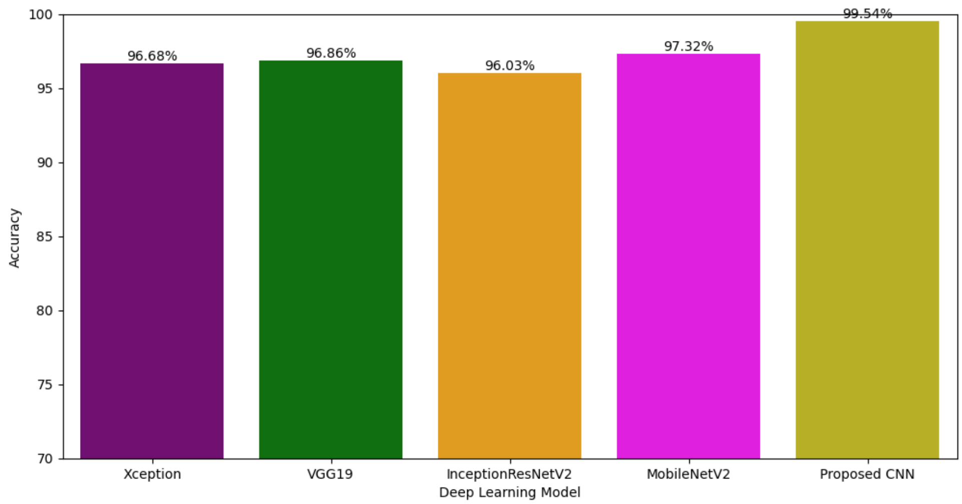

Figure 4.12: Comparative Model Accuracy Bar Plot

The figure 4. 12 This would be followed by a bar plot of all the models that compares their model accuracy as shown below: The axis of y is Y, and the axis of x is the models that have been tested and compared. Actual models are Xception, VGG19, Inception Resnet v2, MobileNetV2, and the developed CNN model. Among all the developed models, the proposed CNN model will have the highest accuracy of at least 97.32%.

In Given Below The table 4.1 shows comparative analysis through precision, recall and F1 score.

Table 4.1: Comparative Analysis of Models Through Precision, Recall and F1 Score

Model	Precision	Recall	F1 Score	Accuracy
VGG19	96.86	96.84	96.67	96.67
Xception	97.01	96.86	96.84	96.86
InceptionResNetV2	96.25	96.03	96.00	96.03
MobileNetV2	97.25	97.32	97.31	97.32
Proposed CNN	99.53	99.52	99.54	99.54

4.2.2 Implementation

Figure 4.13: Output of CNN Image Test

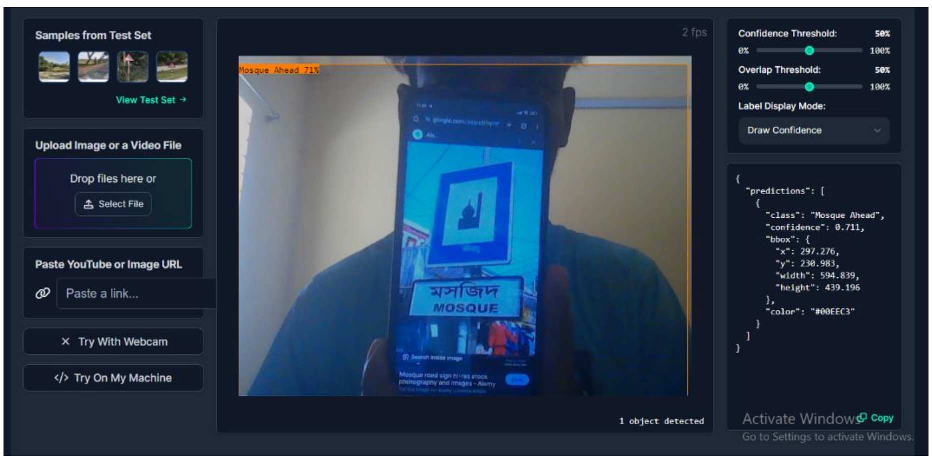

Figure 4.14: Output of YoloV8 Model Real Time Video Test

4.3 Discussion

The experimental outcome also evidently reaffirms that the deep learning network such as MobileNetV2 and VGG19 can provide the highest accuracy rate for Road Traffic Sign of Bangladesh. When comparing to the other network structures, MobileNetV2 provides the needful accuracy to real-time mapping courses in a computer's limited frame symbols and cycles, making it perfect for accepting real-time applications. This analysis of VGG19 performance shows how the deeper network is effective in extracting fine details of the traffic sign images. It can also be concluded that the competitive outcomes of Xception and the developed CNN can be further improved in future researches. Further studies could use ensemble methods or employ different architectures for the model to improve the reliability and performance in different.

CHAPTER 5
Impact on Society

5.1 Impact on Society

The proposed deep learning models for Bangladeshi road traffic sign detection has a great potential to deliver a vast number of societal advantages by increasing road safety measures and traffic flow control. Real time and accurate identification of traffic signs can help decrease the probability of road mishaps, enhance the direction giving to drivers, and enhance the flow of traffic on the roads in the urban and rural setups. Pertaining to this, these systems assist the road users with important knowledge especially in difficult or unknown territories in driving. Thus, the incorporation of such technologies within smart transportation systems may be helpful in pushing up the over therefore general safety and quality of transport in Bangladesh.

5.2 Impact on Environment

Traffic sign detection implemented using deep learning can also be equally beneficial to the environment if well deployed. Ameliorating the traffic flow and eliminating the traffic jam through proper direction and self-driving features decreases the extra time and fuel consumption by cars. Effective traffic regulation also contributes to measures aimed at decreasing the level of carbon footprint and increasing the accessibility to environmentally friendly transport solutions. Moreover, improving roads' safety and decreasing the risks of accidents, with the help of these technologies, keeps natural environments and landscapes from the negative impact of transport-related emergencies.

5.3 Ethical Aspects

Some of the ethical concerns when it comes to traffic sign detection using deep learning models are concerns to do with privacy, equal opportunity and information disclosure. It is crucial to respect the privacy of the users in defining how the data will be gathered and processed. It means that it is necessary to address the issues with data sample and results of an algorithm for providing equal treatment of all groups of people. Some of the key aspects to address when it comes to model decision-making procedures are the model's ability to report on the decision-making process and take responsibility for the failures of the system when it is being used by stakeholders. Managing these ethical issues ensures fair provision of safer transport solutions while observing/cohering to ethical issues & values in designing/rolling out intelligent transport systems.

5.4 Sustainability Plan

The sustainability plan in this study entails the continuation of the interaction with the local government and transportation relevant departments in Bangladesh to incorporate the D-Lite traffic sign detection systems into existing facilities. They can be updated frequently to control their deterioration and the inadequacy they may display to changing traffic dynamics, and the boost in technological expertise. Issue concerning energy-efficient computing as well as the environmentally sustainable ways of handling large volumes of data are in line with the long-term environmental objectives. A major component of the implementation of these technologies is educational outreach and training since the local stakeholders are meant to continue using the technologies/environment in the long term.

CHAPTER 6

Summary, Conclusion, Recommendation and Implication for Future Research

6.1 Summary of the Study

The key contribution of this research was to create, access and propose deep learning models for Bangladeshi road traffic sign detection. The study included gathering and enriching a data base of 5420 picture across 18 traffic sign classes and pre-processing them for modeling and evaluation. CNN models such as MobileNetV2 and VGG19 were tested How effective these various architectures of CNN are in recognizing traffic signs in different environmental conditions. The details of the obtained experimental results included the following: MobileNetV2 received the highest accuracy of 99.54% thus making it ideal for real-time applications Some of the findings of the study include the need to consider the right architectures in line with computation ability and performance indicators to improve on road safety and traffic systems.

6.2 Conclusions

Therefore, it is possible to conclude that deep learning and, in particular, such models as MobileNetV2 and VGG19 has a huge potential in increasing the accuracy of Bangladeshi road traffic sign detection. It contributes to efficient traffic flow in the urban and rural areas and at the same time helps to improve the safety of road transport by providing timely information to the drivers. The future work could build on the studies related to ensemble methods and finally, the topic of the hybrid architecture of PPS and how it performs under different conditions. Hence, the study indicates that the application of sophisticated deep learning methodologies can enhance the development of intelligent transportation systems benefiting all nations, including Bangladesh.

6.3 Implication for Further Study

Future work on Bangladeshi road traffic sign detection employing deep learning may venture into the following areas to improve the performance and usefulness of the system.

Possible research could bring attention to methods of combining multiple sensory inputs like video and sensors for detecting signs in real-life traffic conditions. Moreover, research into other forms of CNN such as the stacking of the architectures or with the mixture of conventional computer vision architectures along with deep CNN architectures can potentially improve the accuracy as well as stability of the models. To enhance generalization ability one can, go further to analyze the use of various domain adaptation methods to overcome dataset bias and environmental change in different area in Bangladesh. Besides, research could expand on the topic of model deployment and utilization on end-devices and carry out research on microarchitectures suitable for implementation in areas with restricted hardware capabilities. Each of these areas might help to progress ITS implementation, resulting in improved, safer, and more efficient roads in Bangladesh and other comparable settings around the world.

REFERENCES

[1] Fatima Akter Dina, Md. Hossain Manik, and S. Mahmud, “"Urban Road Accident Scenarios in Dhaka Metropolitan Area ",” vol. 9, no. 1, pp. 8–23, Feb. 2024, doi: https://doi.org/10.46610/jots.2024.v09i01.002.

[2] B. Sanyal, R. K. Mohapatra, and R. Dash, “Traffic Sign Recognition: a Survey,” IEEE Xplore, Jan. 01, 2020. https://ieeexplore.ieee.org/abstract/document/9072976 (accessed Jul. 02, 2023).

[3] S. S. S. M. Qadri, M. A. Gökçe, and E. Öner, “State-of-art Review of Traffic Signal Control methods: Challenges and Opportunities,” European Transport Research Review, vol. 12, no. 1, Oct. 2020, doi: https://doi.org/10.1186/s12544-020-00439-1.

[4] Arindam Baruah, R. Kumar, and M. Gupta, “Analysis of Traffic Sign Recognition for Automated Transportation Systems Using Neural Networks,” Apr. 2023, doi: https://doi.org/10.1109/i2ct57861.2023.10126423.

[5] W.-H. Lee and C.-Y. Chiu, “Design and Implementation of a Smart Traffic Signal Control System for Smart City Applications,” Sensors, vol. 20, no. 2, p. 508, Jan. 2020, doi: https://doi.org/10.3390/s20020508.

[6] A. S. Konushin, B. V. Faizov, and V. I. Shakhuro, “Road Images Augmentation with Synthetic Traffic Signs Using Neural Networks,” Computer Optics, vol. 5, no. 45, pp. 736–748, Sep. 2021, doi: https://doi.org/10.18287/2412-6179-co-859.

[7] Md. Abu Sayeed, Md. Saiful Islam, Md Babul Islam, Piyush Kumar Pareek, and Tanbin Islam Rohan, “Bangladeshi Traffic Sign Recognition and Classification Using CNN with Different Kinds of Transfer Learning through a New (BTSRB) Dataset,” Apr. 2023, doi: https://doi.org/10.1109/icdcece57866.2023.10151254.

[8] M. Rahman, Hasan Muhammad Kafi, Susmita Roy Rinky, M. Islam, and Md. Musfiqur Rahman, “A Comparative Analysis of Various Deep Learning Models for Traffic Signs Recognition from the Perspective of Bangladesh,” Lecture Notes in Networks and Systems, pp. 547–557, Jan. 2024, doi: https://doi.org/10.1007/978-981-99-8937-9_37.

[9] M. Tushar, I. Jahan, T. Sarker, Lamia Rukhsara, and Tapasy Rabeya, “A Proposed Multi Layer CNN Architecture for Detecting Bangla Traffic Sign,” May 2024, doi: https://doi.org/10.1109/iceeict62016.2024.10534570.

[10] M. Rahman et al., “BanTrafficNet: Bangladeshi Traffic Sign Recognition Using a Lightweight Deep Learning Approach,” Research Square (Research Square), Jun. 2024, doi: https://doi.org/10.21203/rs.3.rs-4216970/v1.

[11] Rishabh Uikey, H. R. Lone, and A. Agarwal, “Indian Traffic Sign Detection and Classification through a Unified Framework,” IEEE Transactions on Intelligent Transportation Systems, pp. 1–10, Jan. 2024, doi: https://doi.org/10.1109/tits.2024.3411117.

[12] K. Muhammad, A. Ullah, J. Lloret, J. D. Ser, and V. H. C. de Albuquerque, “Deep Learning for Safe Autonomous Driving: Current Challenges and Future Directions,” IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4316–4336, Jul. 2021, doi: https://doi.org/10.1109/tits.2020.3032227.

[13] R. K. Megalingam, K. Thanigundala, S. R. Musani, H. Nidamanuru, and L. Gadde, “Indian Traffic Sign Detection and Recognition Using Deep Learning,” International Journal of Transportation Science and Technology, Jun. 2022, doi: https://doi.org/10.1016/j.ijtst.2022.06.002.

[14] T. Liang, H. Bao, W. Pan, and F. Pan, “Traffic Sign Detection via Improved Sparse R-CNN for Autonomous Vehicles,” Journal of Advanced Transportation, vol. 2022, pp. 1–16, Mar. 2022, doi: https://doi.org/10.1155/2022/3825532.

[15] M. A. Khan, H. Park, and J. Chae, “A Lightweight Convolutional Neural Network (CNN) Architecture for Traffic Sign Recognition in Urban Road Networks,” Electronics, vol. 12, no. 8, p. 1802, Jan. 2023, doi: https://doi.org/10.3390/electronics12081802.

[16] K. Dhawan, Srinivasa Perumal R, and R. K. Nadesh, “Identification of Traffic Signs for Advanced Driving Assistance Systems in Smart Cities Using Deep Learning,” Multimedia Tools and Applications, vol. 82, no. 17, pp. 26465–26480, Mar. 2023, doi: https://doi.org/10.1007/s11042-023-14823-1.

[17] J. Wang, Y. Chen, Z. Dong, and M. Gao, “Improved YOLOv5 Network for real-time multi-scale Traffic Sign Detection,” Neural Computing and Applications, Dec. 2022, doi: https://doi.org/10.1007/s00521-022-08077-5.

[18] Y. Li, J. Li, and P. Meng, “Attention-YOLOV4: a real-time and high-accurate Traffic Sign Detection Algorithm,” Multimedia Tools and Applications, Sep. 2022, doi: https://doi.org/10.1007/s11042-022-13251-x.

[19] Emel Soylu and Tuncay Soylu, “A Performance Comparison of YOLOv8 Models for Traffic Sign Detection in the Robotaxi-full Scale Autonomous Vehicle Competition,” Multimedia Tools and Applications, Aug. 2023, doi: https://doi.org/10.1007/s11042-023-16451-1.

[20] S. Zhang, S. Che, Z. Liu, and X. Zhang, “A real-time and Lightweight Traffic Sign Detection Method Based on ghost-YOLO,” Multimedia Tools and Applications, vol. 82, no. 17, pp. 26063–26087, Jan. 2023, doi: https://doi.org/10.1007/s11042-023-14342-z.

[21] Z. Jia, S. Sun, and G. Liu, “Real-Time Traffic Sign Detection Based on Weighted Attention and Model Refinement,” Neural Processing letters/Neural Processing Letters, vol. 55, no. 6, pp. 7511–7527, May 2023, doi: https://doi.org/10.1007/s11063-023-11271-8.

[22] T. P. Dang, N. T. Tran, V. H. To, and M. K. Tran Thi, “Improved YOLOv5 for real-time Traffic Signs Recognition in Bad Weather Conditions,” The Journal of Supercomputing, Feb. 2023, doi: https://doi.org/10.1007/s11227-023-05097-3.

[23] Ruturaj Mahadshetti, J. Kim, and T.-W. Um, “Sign-YOLO: Traffic Sign Detection Using Attention-based YOLOv7,” IEEE Access, pp. 1–1, Jan. 2024, doi: https://doi.org/10.1109/access.2024.3417023.

[24] A. Radha Rani, Y. Anusha, S.K. Cherishama, and S. Vijaya Laxmi, “Traffic Sign Detection and Recognition Using Deep learning-based Approach with Haze Removal for Autonomous Vehicle Navigation,” e-Prime, Advances in Electrical engineering, Electronics and Energy, vol. 7, pp. 100442–100442, Mar. 2024, doi: https://doi.org/10.1016/j.prime.2024.100442.

[25] R. Ayachi, M. Afif, Y. Said, and M. Atri, “Traffic Signs Detection for Real-World Application of an Advanced Driving Assisting System Using Deep Learning,” Neural Processing Letters, vol. 51, no. 1, pp. 837–851, Sep. 2019, doi: https://doi.org/10.1007/s11063-019-10115-8.

[26] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, “A Comprehensive Survey of Image Augmentation Techniques for Deep Learning,” Pattern Recognition, vol. 137, p. 109347, May 2023, doi: https://doi.org/10.1016/j.patcog.2023.109347.

[27] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth, “A Review of Medical Image Data Augmentation Techniques for Deep Learning Applications,” Journal of Medical Imaging and Radiation Oncology, vol. 65, no. 5, pp. 545–563, Jun. 2021, doi: https://doi.org/10.1111/1754-9485.13261.

[28] D. Singh, P. Das, and I. Ghosh, “Prediction of pedestrian crossing behaviour at unsignalized intersections using machine learning algorithms: analysis and comparison,” *Journal on Multimodal User Interfaces*, May 2024, doi: https://doi.org/10.1007/s12193-024-00433-0.

[29] N. Behboudi, S. Moosavi, and R. Ramnath, “Recent Advances in Traffic Accident Analysis and Prediction: A Comprehensive Review of Machine Learning Techniques,” *arXiv.org*, Jun. 19, 2024. https://arxiv.org/abs/2406.13968 (accessed Jul. 01, 2024).

[30] Md. Mehedi Hasan, N. Aman, and Md. Wadud Jahan, “Enhancing Deep Convolutional Neural Networks to Detect Pedestrians Effectively in Various Urban Environments,” May 2024, doi: https://doi.org/10.1109/iceeict62016.2024.10534409.

[31] Awal Ahmed Fime, Md. Ashikuzzaman, and A. Aziz, “Audio signal based danger detection using signal processing and deep learning,” *Expert Systems with Applications*, vol. 237, pp. 121646–121646, Mar. 2024, doi: https://doi.org/10.1016/j.eswa.2023.121646.

[32] M. M. Rahman and N. Nower, “Inferring Traffic Patterns of Dhaka City: A Spatio-temporal Analysis over a Year,” *Research Square*, Jan. 03, 2024. https://www.researchsquare.com/article/rs-3827938/v1 (accessed Jul. 01, 2024).

[33] F. Zhuang, et al., “A Comprehensive Survey on Transfer Learning,” Proceedings of the IEEE, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: https://doi.org/10.1109/jproc.2020.3004555.

[34] R. Ribani and M. Marengoni, “A Survey of Transfer Learning for Convolutional Neural Networks,” IEEE Xplore, Oct. 01, 2019. https://ieeexplore.ieee.org/abstract/document/8920338