Vehicle-NN- A CNN Based Local Vehicle Detection Classifier

BY

Umme Fariha Kabir
ID: 171-15-8753
AND

Abida Ali
ID: 171-15-8754

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

Supervised By

Ms. Nishat Sultana
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Ms. Zerin Nasrin Tumpa

Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JUNE 2021

APPROVAL

This Project titled “**Vehicle-NN- A CNN Based Local Vehicle Detection Classifier** ”, submitted by *Umme Fariha Kabir* and *Abida Ali* 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 *1 June 2021*.

BOARD OF EXAMINERS

____

Dr. Touhid Bhuiyan **Chairman**
Professor and Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

____

Md. Tarek Habib **Internal Examiner**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

____

Saiful Islam **Internal Examiner**
Seenior Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Md Arshad Ali **External Examiner**
Associate Professor
Department of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Ms. Nishat Sultana, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

____

Ms. Nishat Sultana
Lceturer
Department of CSE
Daffodil International University

Co-Supervised by:

____

Ms. Zerin Nasrin Tumpa
Lceturer
Department of CSE
Daffodil International University

Submitted by:

____

Umme Fariha kabir
ID: -171-15-8753
Department of CSE
Daffodil International University

____

Abida Ali
ID: -171-15-8754
Department of CSE

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Ms. Nishat Sultana**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine learning*" to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

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

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

In this technological era, computer vision accomplishes superior performance and it introduced to the Convolutional neural network (CNN) which is acquainted for recognize an object. In addition to object detection and classification are considered difficult tasks in computer vision. From the ancient times vehicle is the only communication media to move from one place to another place. In this automation period local vehicles are decreasing day by day. Because these type of vehicles motion are slow. We want to preserve our culture with the help of vehicle-NN model.We presented our own CNN model name Vehicle-NN which will be a convenient effect on the vehicle sector of Bangladesh. We compared our CNN model with other pre-trained models like MobileNet, VGG16, InceptionV3. This proposed model plays an effective role in the future automation vehicle identification sector. Besides we want to hold on to our traditional roots and preserving our local culture. Our work approach is novel for the detection of local vehicles in Bangladesh. In this experiment, we have used 7 classes (Bus, CNG, Leguna, Pickup, Rickshaw, Thelagari, Van). We trained our model with our own dataset and our CNN model acquired accuracy is 96%.

TABLE OF CONTENTS

Chapter 3 : Research Methodology

Chapter 4 : Experimental Results and Discussion

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

In the last few decades advancement at a tremendous rate. Revolution of technology our life becomes dynamic. The vehicle is essential for faster communication from one phase to another phase. In both the past and the new world, we can't conjecture our single movement without the help of vehicles. The term(vehicle) is coined from a Latin word. Extensively increasing the techniques of automatic vehicles for time-consuming. [9] As a consequence our local vehicle at the end of the extinct door. The advancement of technology widely increases the new electric car from tesla or else the autonomous vehicles by Mercedes. [14] In this paper, we consider how CNN can be executed in image classification approaches.The field of computer science is computer vision that is a process of creating a digital system. The computer vision concept is based on basically instructing computers to a procedure to visualize data(videos and images) at the pixel level and understand it. The algorithm of computer vision depends on pattern recognition. Computer vision relies on Deep learning. In modern years deep learning has been creating an interest in various recognition tasks as example as: object detection [6], activity recognition[4], scene classification[10]. Object classification indicates the field of computer vision that shows in images into cabalistic categories.A CNN is a sector of a deep neural network, most commonly applied to analyzing visual pictures. CNNs have acquired astonishing advancement in the field of vehicle detection. [7]. In our proposed method, we developed a new model to detect the local vehicle. At first, we train our Vehicle-NN model with our own raw data of 7 classes whose names are: van, thelagari, bus, CNG, rickshaw, leguna, and pickup. Collecting our raw data we face some challenging tasks to find few local vehicles like: van, thelagari around the cities. These vehicles are on the verge of extinction is the speed of movement of slow. Towards this our classifier can predict local vehicles using CNN and it will be work as the way we train. The outcome of our classifier relies on the datasetIn this work, we also train

MobileNet, InceptionV3, VGG16 models for comparison with our developed Vehicle-NN model for more reliability.

1.2 Motivation

In this modern era, preserving our local vehicle is really a challenging task. Nowadays we are floating in the tide of technology. As a result, our interest in retaining local resources is waning. We have lost our cultural resources by keeping pace with technology. Local vehicles are also among our local resources. So, our motivation behind this work is preserving our culture with the help of our traditional model. We want to hold our traditional roots and reserving our local culture.

1.3 Objective

1. The objective of this project to develop a new technique that can support future generations to gain knowledge about local vehicles of our country.

2. We want to represent our local vehicle to the outside world.

3. We are going to use CNN architecture and our own data set to classify the vehicles properly.

4. Wide diversity of vehicle and image recognition.

5. Apply convolutional neural network configuration (CNN) architecture to detect local vehicles more efficiently.

1.4 Research Question

1. Which type of challenges did we face for the vehicle classification approach?
2. How can we collect raw data in this pandemic situation?
3. What is the impact of vehicle detection in our day-to-day life?
4. How can we apply machine learning to implement our model?
5. What is the better pathway for vehicle detection?
6. Which sector is beneficial to this research?
7. How can it perform in future automation in the vehicle sector?

1.5 Expected Outcome

Our project is research-based. This research plays an effective role in the vehicle sector. Vehicles are part of our lives. Because we can't imagine a single communication without the help of vehicle. In the future for faster and easier lifestyle, we apply computer science in the vehicle sector. Below we described our research expected outcome-

1. Identify the vehicle I primary stage.
2. Generating an efficient CNN-based model to recognize the vehicles more precisely.
3. Comparing the model developed with other existing models to improve accuracy.
4. Our research will be used in smartphone applications if we trained our model with various types of dataset. It will be able to detect the vehicle instantly and get information about acceleration, velocity, capacity, etc.

1.6 Report Layout

Chapter 1 : Introduction

We will describe the introduction of our work, Motivation of our work, Objective of our work, research question, Expected outcome from our work and Research layout in this chapter.

Chapter 2 : Background

We will describe the background of our work. We will also describe the prior work, The research summary, Scope of the problem, Challenges of our work in this chapter.

Chapter 3 : Research Methodology

In this chapter we will describe how we can build up our model. At first we will describe our data collection procedure, data pre-processing method, some theoretical term that will be related to our work. Our proposed CNN model, number of parameters of our model and the procedure of train the model with training dataset.

Chapter 4 : Experimental Results and discussion

This chapter will be containing summary of the results of our model and the comparison among performance evaluation of our model with other pre-trained models.

Chapter 5: Conclusion

This chapter will be containing the conclusion of our work and future implementation of this work.

CHAPTER 2

BACKGROUND

2.1 Introduction

In this chapter, we represent some researcher's work about the sector of object detection, vehicle type identification.

In this technological era, people are less using slow-motion vehicles for their transportation. As a result, they have less knowledge about local vehicles. Our proposed model can easily identify the local vehicles and from images, it can classify the vehicles more accurately. For vehicle classification, many researchers are using many approaches. Their ultimate goal is image processing.

Image processing is considered a subfield of signals. This sector focus on images only. By using algorithm image transfer into digital form. To acquire some particular model it acts on some operation and analyzing raw images, gather information from images is called image processing. Computer vision can understand the visual world. Our research is to classify local vehicles from raw images and it is based on computer vision.

In this chapter, we described prior work of our research, a summary of research, scope of the problem, and challenges.

2.2 Prior Work

Using CNN we proposed a vehicle NN model to detect local vehicles of Bangladesh. In the community of computer vision, CNN has been studied and developed day by day.

In paper [1], The authors proposed a neural network that identifies the type of vehicle. In this paper, they proposed a framework based on various geometrical parameters. They implement it on 9 different classes of vehicles and achieve 95% of accuracy.

In paper[2], they represent the CNN in the classification and detection of vehicles using a traffic camera which resolution is low. Their system achieved an average of 94.72%.

In paper[3], the authors introduced a lean CNN for vehicle classification. They developed a lean CNN with smaller parameters and keeping the best accuracy on vehicle classification. They divide their dataset into 5 different classes. They constructed 6 different models and in the future, they want to increase validation accuracy.

In paper[5], they focuses on the survey on vehicle detection based on CNN. The CNN-based vehicle detection in the content of satellite imagery and monocular vision.

In paper [6], the authors proposed a classifier on the base of CNN. It classifies our traditional Bengali games. They retrained the last layer of the Inception V3. They identify 5 different classes and the average accuracy is 80%.

In paper[9], the authors introduced a model vehicle type classification using CNN. Their main focus is contemporary road safety and intelligent transportation system. This dataset has 2400 images and is divided into 4 classes. Their model constituted two steps. Firstly they augmented their dataset and finally build a CNN model with different architecture.

In paper [10], they proposed CS-CNN model. They used a pre-trained model which is VGG network. Their proposed model compared with VGG-s and VGG-Verydeep-16. They getting 97% accuracy.

In paper[11], the authors proposed a CNN model for image classification(CNNVA). They collect images from traffic camera which is used for classification of automatic vehicles. Their proposed method is tested on a surveillance nature dataset.

In paper [12], the authors proposed a system for Real-Time Traffic Surveillance which is Vehicle Detection and Counting System. For image processing, they developed their algorithm and that was implemented and tested on smart camera which platform is embedded.

In paper[13], Lucan firstly proposed it and worked on digit recognition which is handwritten. And it depends on a small scale.

2.3 Research Summary

We collect data from various places for the purpose of our research. We use CNN architecture to develop our model. We create our own Vehicle-NN model. We trained our model as it can be capable of recognizing an image. For developing this model we gather lots of knowledge and learning many techniques from other related work. Our classification system was developed by python through jupyter notebook. After the training session has been done, our model is ready to test images to recognize the vehicles. We create a Vehicle-NN model, which can capable to classify the local vehicles more precisely.

2.4 Scope of the problem

When we attempt to do a new task, at first it is very difficult to identify the existing problem. We can get a clear idea of these problems during the research period. In the beginning, we don’t understand which platform is suitable for our research work. After that, we realize any kind of platform will be used for implementation. Our proposed model gives a basic idea of data processing and how the algorithm works. Actually how we process the data, which model is perfect, and how programming can be implemented are the main problems we found.

2.5 Challenges

We face many challenges to complete our research. To develop our model we need lots of images of local vehicles. We collect our local vehicle images from several roadsides. In covid 19 situation lockdown is strictly followed for a few months. So, it is a big challenge for us to collect images from the roadside in this pandemic situation.

Then another challenge is capturing a good quality image with high resolution. We can’t process the image perfectly if those images are low resolution. The other concerning issue is RGB value which is essential for image processing. We pass an enormous time collecting proper images to make our own dataset.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Vehicles vary from country to country. The recognition of vehicles are all over the country is a tough job. Many researchers applied many techniques to recognize vehicles from traffic videos, classified vehicles from CCTV footage. In the before, there has no research about the classification of local vehicles of Bangladesh. In this paper, we proposed our Vehicle-NN model that developed with some method to get a response from the computer system. Our model classified the local vehicles of Bangladesh.

3.2 Research Methodology

A research project is dependent on research methodology. To achieve our target research methodology give us instructions on how can we step forward. Perfect planning is needed to develop a new model. So, we follow some planning and methodology to get our expected outcome. The process development of our model Data collection, Data pre-processing, Apply CNN to our model, Train model (MobileNet, Inception V3, VGG16), Comparison.

Work flow of our proposed model ---

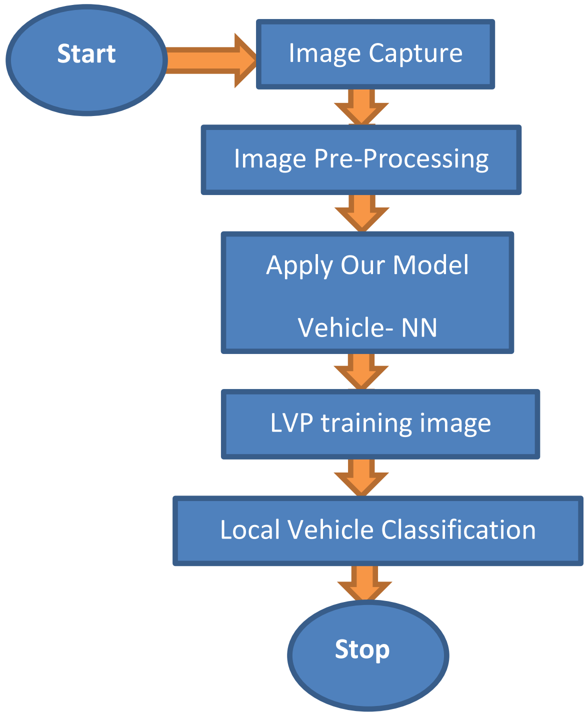

Figure 3.2.1 : Proposed system workflow

3.3 Procedure of Data Collection

An adequate dataset is the core part of the machine learning technique. For our proposed model, we collect raw data and make a new dataset. There are 1304 images with 7 classes like Bus, CNG, Leguna, Pickup, Rickshaw, Thelagari, van. Each class has more than 120 images. Our collected image format is Jpg. Our dataset was collected manually from the roadside. We considered 80% data for training purposes and 20% data for testing purposes.

Bus

CNG

Leguna

Pickup

Rickshaw

Thelagari

Van

Figure 3.3.1 : Sample of our dataset

3.4 Data Pre – Processing

The bottom of machine learning is data. We collect our raw data with different heights and weights. Resolution is different for all images including our dataset. So, it is necessary to prepare them for our work purpose. In the next part, we resize all of our collection images into 224*224 pixels.

Our all considered images are RGB. For getting better accuracy we used RGB color. Because it helps to detect features easily. We can diminish values of RGB divided by 255 for normalizing.

3.5 Proposed Methodology

In the below we describe our proposed methodology step by step.

3.5.1 CNN

In a machine learning algorithm, CNN is known as a feed-forward type. CNN depends on the perceptron model. CNN architecture is inspired by the process of primary visual cortex treatment of the biological brain. The structure of CNN is constructed with the help of the input layer, several hidden layers, and output layer. CNN takes a picture as input then it is processed by a hidden layer and finally classifies the picture into trained categories. A CNN is composed of different kinds of layers.

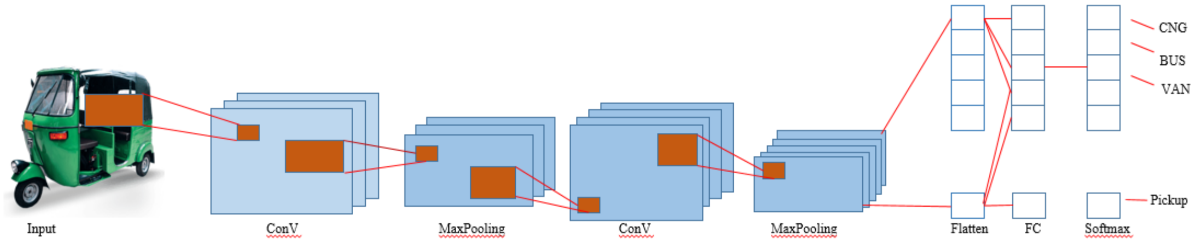

Figure 3.5.1.1 : Demo of Vehicle-NN

3.5.2 Convolutional Layer

The core layer of this algorithm is the convolution layer. It is liable for feature extraction of the input picture. Scanning of machine language is happening in convolutional. A machine can identify the pixel form of any image. The pixel is scanning by filter and gathers all information. The filter can be any kind of and it can be applied to detect images edge filter, curve filter, color filter intensity. There are different sizes of filters. There has a random value in the filter. When we convolve the filter, besides with the pixel value of images and finally we get a new set of values. It can help to identify the property of images. In this way, the rest of the properties are fit into the convolutional layer. The output of the convolutional layer fitted in the fully connected layer. So, all over the process is helpful to image classification. The mathematical operation is performed between the input picture and a filter of an individual size M*M. The mathematical term is:

$$S= \max(0,X*K)$$

3.5.3 Kernel

Convolutional using kernel. The kernel is a matrix and acts as a filter. It is used to extract features from input images. Edge detection, blurring, sharpening, embossing, and more are implemented by the kernel. Kernel pas the input data by the value of stride.

The general expression is:

$$g\ (x, y) = w * f\ (x, y) = \sum_{dx\ =-a}^{a} \sum_{dy=\ -b}^{b} w\ (dx, dy) f(x + dx, y + dy),$$

Where $g\ (x, y)$ is the filtered image, $f\ (x, y)$ is the original image, w is the filter kernel is considered by $-a \leq dx \leq a$ and $-\ b \leq by \leq b$

3.5.4 Stride

In CNN, Stride is considered an element and also it another building block. The input matrix is shift over the number of pixels is called stride. We move the filters based on stride numbers. Stride numbers can't be fraction or decimal. It is always an integer value. By default, the stride number is 1. The movement of a filter depends on stride number. When stride = 1 then filter movement in 1 pixel. Increasing the number of strides, then we will get a small resulting output.

3.5.5 Padding

In the process, when we adding an extra layer to our input matrix it is called padding. The extra layer consists of zero. We considered the extra layer of zero because we don't want to lose any information on our input image. If we want the same kind of output as like as input image pixel. Then we have to apply padding. We use padding to get the same image size after applying a filter operation.

3.5.6 Activation Function

In a Neural network, function is called as an activation function. The important task of the neural network is to find out the value of gradiant and slope at each point. ReLU is the most applied as activation function in CNN. In activation function, we used ReLU when CNN is non-linear. Solving the regression problem we used ReLU as an activation function. It is much more popular from sigmoid function. Back propagation allows ReLU. ReLU define: Max(y,0)

Suppose,y=-ve max(-ve,0)

Output= 0

y=+ve max(+ve,0)

Output= that particular value.

Output 1 means neuron is activated and output 0 means neuron is not activated. Neuron activated means it is transferring the signal and helping to classify the final output.

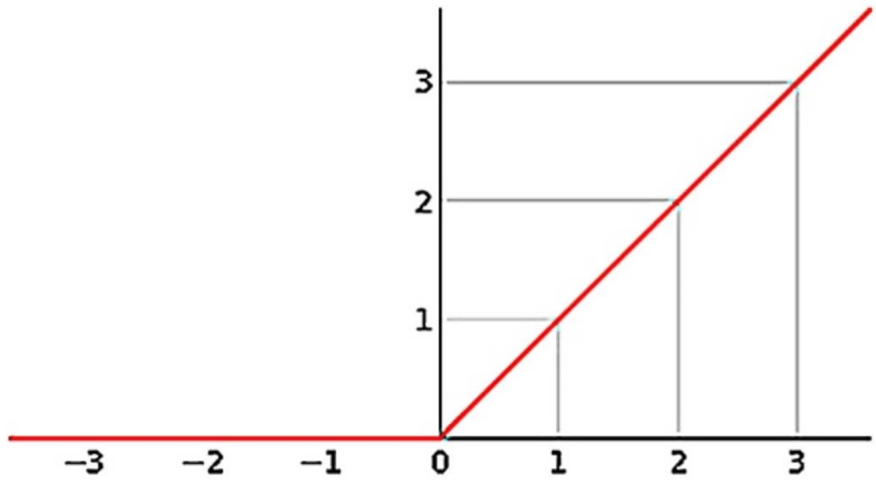

Figure 3.5.6.1: ReLU Function.

3.5.7 Maxpooling Layer

The maximum value is picked up from the convolution layer output. Proper detection is happening in this layer. Because it only picked up the high pixel value. Actually maxpooling extracts the maximum value. The pooling layer is implemented after the convolutional neural network. There are many types of pooling layers. Max, min, sum, average. We use max polling in our CNN model. The number of the parameter is increased in the convolution layer. When the pooling layer is applied, the number of the parameter is reduced. It reduces the special resolution of feature map but it keeps the information properly. In this layer, the size of feature map is also reducing. When we apply this layer, the computational performance is increasing. The max pooling acts as a strong enhancer of the CNN to handle the transaction [4].

3.5.8 Flatten

For simplification purpose, the image is portioned into a pixel. Flatten means converting all the resultant 2D arrays into a 1D feature vector. When a number of pooling layers is occurred, we are getting a final feature map. After that, we applied to flatten in this feature map and flatten according to row. Flatten value will be passing into a fully connected layer as input. In or model, to fit into the fully connected layer, we flatten our final matrix.

3.5.9 Fully Connected Layer

In NN (neural network) input of one layer is connected to the next layer using an activation unit is called fully connected layer. It is also called the Dense layer. This layer is 1 dimensional. It is very important layer when it comes to classifying an image into a label for classification problem perspective or getting the numerical predictions for regression-based problem. This layer is considered feed-forward network. The calculation of this layer is element-wise multiplication. A fully connected layer in CNN is the one that takes the end result of the convolution or pooling layer why a flattened layer and reaches a classification decision in fully connected layer every input is connected to every output by of weight it serves the purpose of doing the actual classification without this layer a traditional CNN would be unable to spit out the predicted classes.

3.5.10 Dropout Layer

It used to obstruct a model from overfitting. Overfitting in the neural network occurs when it gets turned on training data such that it doesn't perform well on test data. Drop layer drops out a random set of Neurons or activations setting them to Zero. It's an approach that helps out reducing interdependent learning amongst the neurons actually. It is computationally cheap and dropout regularization is an efficient way to diminish overfitting and rectify generalization error in deep neural networks of all kinds.

3.5.11 Softmax

Generally, softmax used the last layer of CNN for various class detection. If we have a network that classifies some different classes, there softmax is used in the last layer to produce different probability for each class. It converts the real value into probability form. When we get a high probability then we considered it as output.

3.6 Proposed CNN Model

We considered 12 layered CNN for training purpose. For reducing overfitting we used dropout layer= 0.8 in our model. In our model, ReLU is used in the first layer, and Softmax is used in the last layer. In **Figure 3.6.1** we show our proposed CNN model.

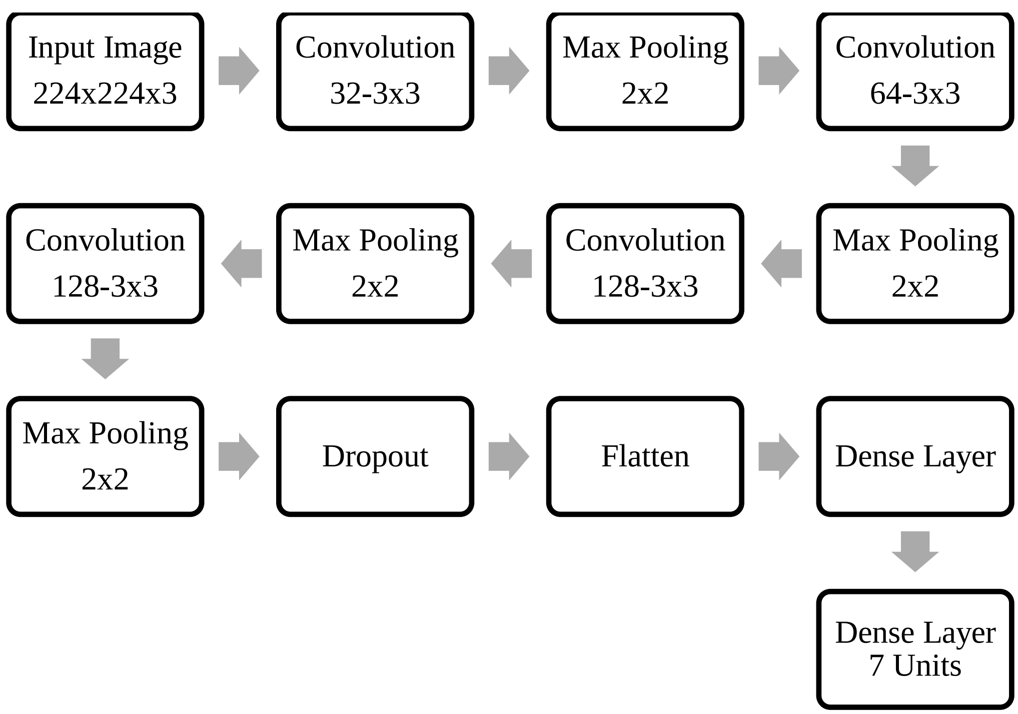

Figure 3.6.1: Our Proposed CNN model

3.7 Number Of Parameters

In our Vehicle-NN model, we equipped up our model with different layers, number of filters, activation shape, activation size, parameters, etc. Then we get the total number of parameters 1,420,999. We used RGB color images and our first image input size is (224*224). Our Vehicle-NN model is decorated with 4 numbers of Conv2D layers and MaxPool layers, one Dropout layer, one flatten layer, 2 Dense layers. We used ReLU as an activation function and Softmax is used in the output layer. Our model stride= 1 and padding =0. Kernel size is(3,3). The number of filters (32,64,128.128).

Step 1:

Pixel to pixel image reading by input layer. The activation shape for the input layer is (224*224*3). Here, 3 is RGB color image, input image height is 224, and input image width is 224. There has no parameter in the input layer.

Step 2:

The first layer of our model is Conv2D(1). Here filter size is 32 and kernel (3,3). (224+2p-3/1)+1= 222. Here, 224=input size, padding=0, 1=stride. After that we get (222,222,3) as output shape. The first Conv2D produce the number of parameter= (((3*3*3)+1)*32)=896. Here, 3= RGB color image,(3*3)= kernel size. The activation shape (222,222,32) and activation size 1,577,088.

Step 3:

In the MaxPooling layer, the number of the parameter is reduced. The activation shape of the first maxpooling(1) layer is (222/2)=111. Here, we divided it by 2 because our maxpooling is 2D. The activation shape is(111,111,32) and the activation size is 394,272. Here filter size is 32. There is no parameter is generated by the pooling layer.

Step 4:

The second layer of our model is Conv2D(2). Here filter size is 64 and kernel size (3,3). (111+2p-3/1)+1=109. Here 111=input size, padding=0, 1= stride. After that we get (109,109,64) as output shape. The second Conv2D produce the number of parameter (((3*3*32)+1)*64) = 18496. Here 3= RGB color image, (3*3)= kernel size. The activation shape (109,109,64) and the activation size 760,384.

Step 5:

The second layer of this model is Maxpool (2). The activation shape(109/2)=54 and we get (54,54,64). The activation size is 186,624. Here, the filter size is 64.

Step 10:

Flatten layer converting all the resulting 2D arrays into a 1D feature vector. When we completed the convolutional layer and maxpooling layer then we flatten our layer. The activation shape is 18432. There is no parameter and activation size in this layer.

Step 11:

The next layer is the fully connected layer. We called it dense layer also. There has no activation size and the activation shape is 64. The number of ((18432+1)*64)= 1,179,712 parameter is produced in this layer.

Step 12:

The softmax layer is the last layer of our model and this layer is used for final classification. ((64+1)*7) =455 is the number of parameters.

Summation of all layer parameters and we get a total number of the parameter is 1,420,999.

Table 3.7.1: Number of parameters

Number	Operation	Number of Filter	Activation Shape	Activation Size	Parameters
1	Input Size	-	224,224,3	150,528	-
2	Conv2D 1	32	222,222,32	1,577,088	896
3	MaxPooling 1	-	111,111,32	394,272	-
4	Conv2D 2	64	109,109,64	760,384	18496
5	MaxPooling 2	-	54,54,64	186,624	-
6	Conv2D 3	128	52,52,128	346,112	73856
7	MaxPooling 3	-	26,26,128	86,528	-
8	Conv2D 4	128	24,24,128	73,728	147584
9	MaxPooling 4	-	12,12,128	18432	-
10	Dropout	-	12,12,128	18432	-
11	Flatten	-	18432	-	-
12	Dense	-	64	-	1179712
13	Softmax	-	7	-	455
Total parameter					1,420,999

3.8 Train Model With Training Dataset

For getting the expected outcome, we are now going to describe our proposed model met hodology. In our dataset, we subsume 7 different classes. Our dataset contains 1304 imag es. We used 1043 images for training our model and 261 images using for testing .To get more valid accuracy, we using random augmentation to prepare our dataset. We update o ur CNN model many times and change the optimizer, learning rate, loss function for incre asing and decreasing loss as possible. We used the ‘adam’ optimizer to train state =5 in both the train set and test set. In a pre-processing step, we didn’t apply any our model to reduce loss function as possible. Our training dataset is 80% and the test dataset is 20%. The number of the epoch is 80 is used in our model. For final test evolution, our model is ready to predict unseen data and finally we get validation loss 0.2001 percent , validation accuracy 96.17%, training accuracy 99.42 percent.

```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_20 (Conv2D)           (None, 222, 222, 32)      896

max_pooling2d_19 (MaxPooling (None, 111, 111, 32)      0

conv2d_21 (Conv2D)           (None, 109, 109, 64)      18496

max_pooling2d_20 (MaxPooling (None, 54, 54, 64)        0

conv2d_22 (Conv2D)           (None, 52, 52, 128)       73856

max_pooling2d_21 (MaxPooling (None, 26, 26, 128)       0

conv2d_23 (Conv2D)           (None, 24, 24, 128)       147584

max_pooling2d_22 (MaxPooling (None, 12, 12, 128)       0

dropout_4 (Dropout)          (None, 12, 12, 128)       0

flatten_5 (Flatten)          (None, 18432)             0

dense_10 (Dense)             (None, 64)                1179712

dense_11 (Dense)             (None, 7)                 455
=================================================================
Total params: 1,420,999
Trainable params: 1,420,999
Non-trainable params: 0
```

Figure 3.8.1 Model Summary

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

We propose our own Vehicle-NN model to classify local vehicles. We tried our best to get a better accuracy. We compare our Vehicle-NN model with other pre-trained InceptionV3, MobileNet, VGG16. In this chapter 4, we showed the confusion matrix, accuracy, classification result, classification report, result analysis, and comparison of all models.

4.2 Performance Evaluation

Above we describe our CNN architecture. The average accuracy of our 7 class is 97%. In **Figure 4.2.1** the upwards part is validation accuracy and **Figure 4.2.2** the downwards part is training loss. At that point, we said that the training model acquired perfect knowledge from the training dataset.

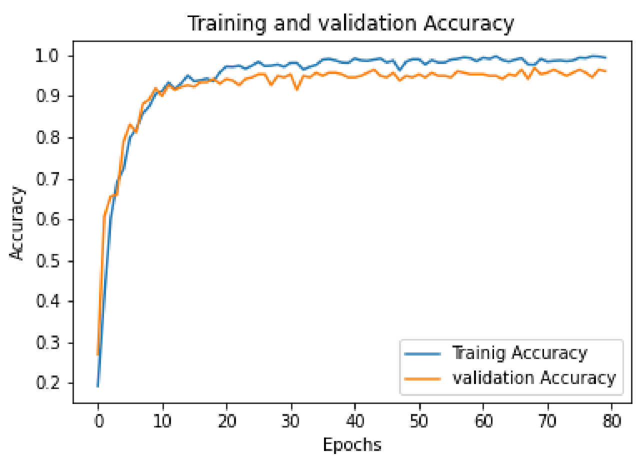

Figure 4.2.1: Training and validation accuracy(CNN)

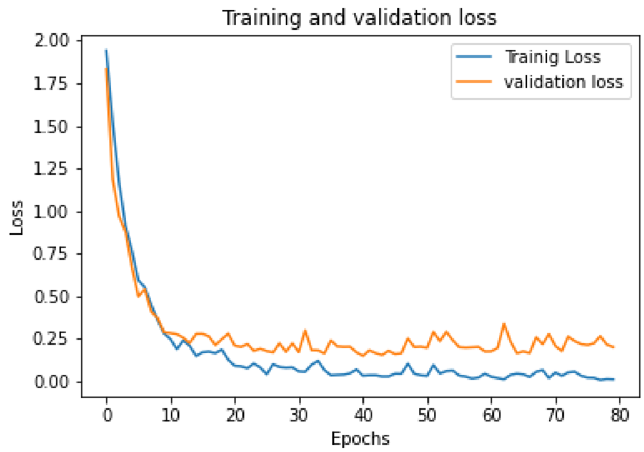

Figure 4.2 :Training and validation loss(CNN)

Table 4.2.1: Classification report

Model	Accuracy	Precision	Recall	F1-Score	No of epoch
CNN	96	95.8	95.8	95.8	80
MobileNet	99	98.7	98.7	98.7	80
InceptionV3	97	96.8	96.8	96.8	80
VGG16	97	97.7	97.2	97.2	80

4.3 Result Discussion

Our CNN model achieves a high accuracy of precision, recall and every weighted average up to 96.The images of the total dataset are 1304. After classification, The correct prediction is 251and the false prediction is 10. Finally, our model performs 96% accuracy. Observing the above confusion matrix for unseen data, our model gives better accuracy

Table 4.3.1 : CNN classification report

Class	Precision	Recall	F1-score
Bus	0.98	0.95	0.96
CNG	1.00	0.90	0.95
Leguna	1.00	1.00	1.00
Pickup	0.98	1.00	0.99
Rickshaw	0.98	0.96	0.97
Thelagari	0.90	1.00	0.95
Van	0.87	0.90	0.88

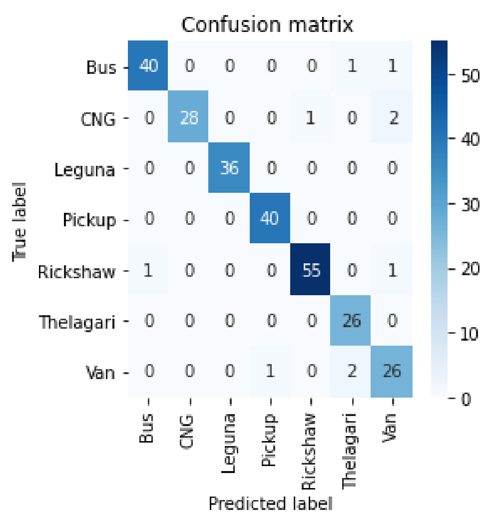

Figure 4.3.1: CNN Confusion matrix

4.4 Comparison

In this paper, we compare our proposed CNN model with other pre-trained models like MobileNet, VGG16, Inception v3. We used the same number of epoch in all models. We divided our dataset by 80:20 for training and testing purpose. We summarized all model accuracy in table [2]. MobileNet achieved the best accuracy from another two pre-trained models.

4.4.1 MobileNet

MobileNet achieved 99% accuracy. This is the best accuracy from other models.

Classification Report:

Classification Report:

	precision	recall	f1-score	support
0	0.94	1.00	0.97	30
1	1.00	1.00	1.00	35
2	1.00	0.91	0.96	35
3	0.97	1.00	0.98	30
4	1.00	1.00	1.00	56
5	1.00	1.00	1.00	35
6	1.00	1.00	1.00	40
accuracy			0.99	261
macro avg	0.99	0.99	0.99	261
weighted avg	0.99	0.99	0.99	261

Figure 4.4.1.1 MobileNet Classification Report

Confusion matrix:

Here Bus, CNG, Pickup, Rickshaw, Thelagari, Van are identified accurately.

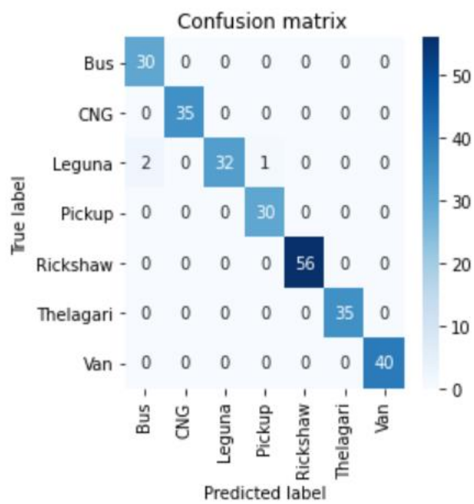

Figure 4.4.1.2 MobileNet Confusion Matrix

Training and Validation Accuracy:

In this Figure 4.4.1.3 blue line is training accuracy and orange line is validation accuracy. We saw that training and validation accuracy are almost same. Considering the gap we saw that gap range is less and less noisy.

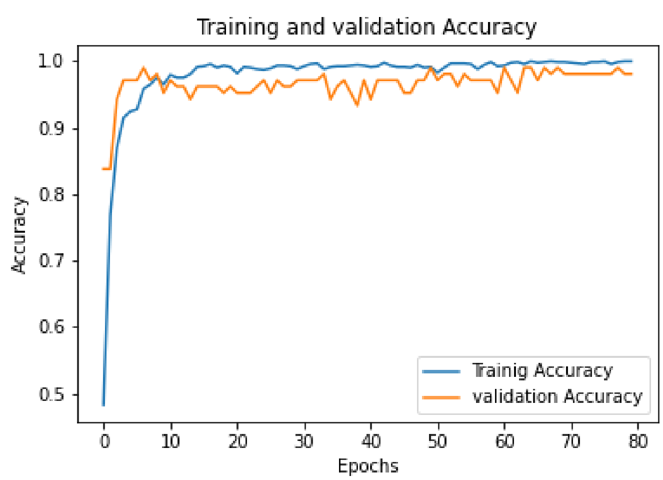

Figure 4.4.1.3 Training and Validation Accuracy (MobileNet)

Here validation accuracy is little bit noisy. But our training accuracy is good.

Training and Validation loss:

Here validation loss is noisy.

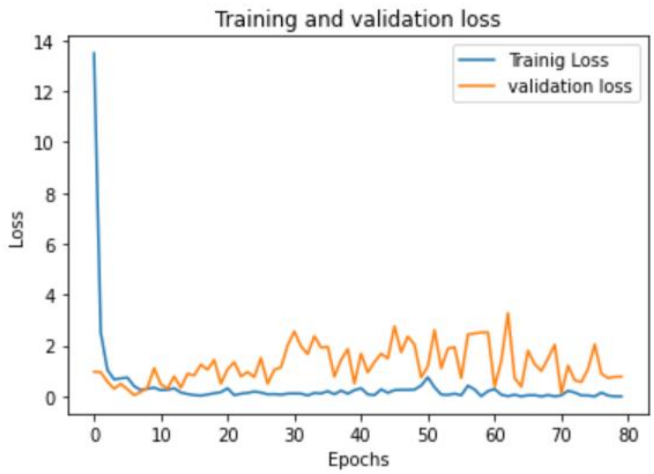

Figure 4.4.1.4 Training and Validation loss (MobileNet)

4.4.2 VGG16

VGG16 achieved 97% accuracy.

Classification Report:

Classification Report:

	precision	recall	f1-score	support
0	0.97	0.97	0.97	36
1	1.00	0.97	0.98	29
2	0.94	0.97	0.95	32
3	1.00	1.00	1.00	42
4	1.00	1.00	1.00	50
5	1.00	0.90	0.95	31
6	0.93	1.00	0.96	41
accuracy			0.98	261
macro avg	0.98	0.97	0.97	261
weighted avg	0.98	0.98	0.98	261

Figure 4.4.2.1 VGG16 Classification Report

Confusion matrix:

Here Pickup, Rickshaw and Van are identified accurately.

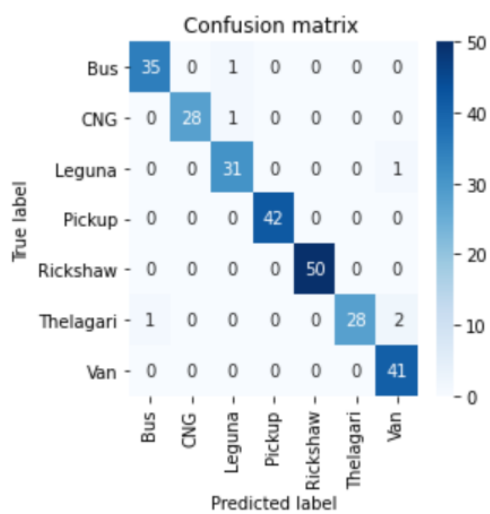

Figure 4.4.2.2 VGG16 Confusion Matrix

Training and Validation Accuracy:

We saw that training and validation accuracy are almost same. Considering the gap we saw that gap range is less and less noisy.

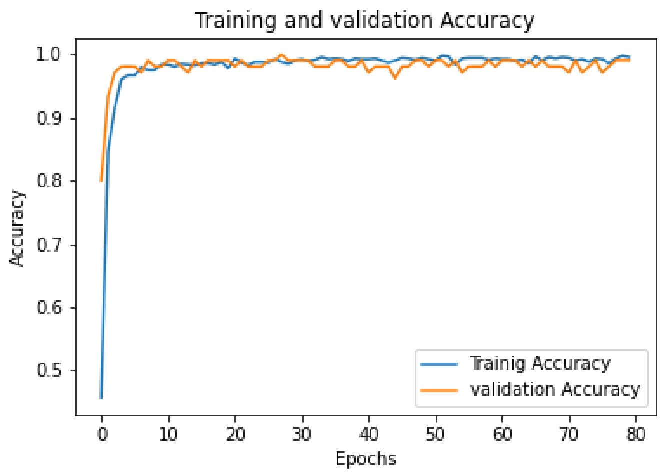

Figure 4.4.2.3 Training and Validation Accuracy (VGG16)

Training and Validation loss:

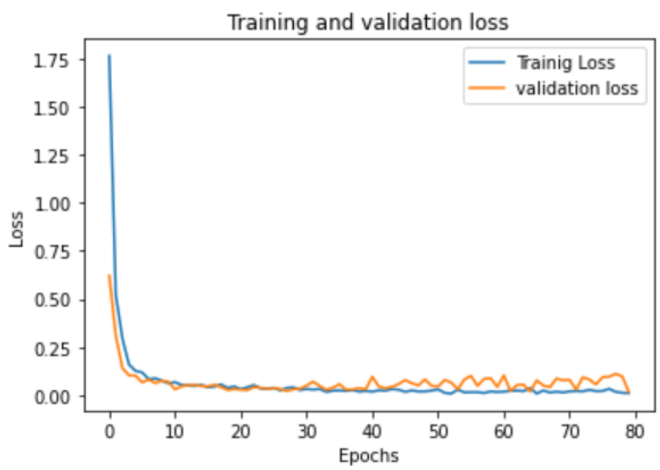

Figure 4.4.2.4 Training and Validation loss (VGG16)

4.4.3 InceptionV3

InceptionV3 achieved 97% accuracy.

Classification Report:

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.97	0.99	35
1	0.96	1.00	0.98	25
2	0.96	0.96	0.96	25
3	0.97	0.97	0.97	36
4	0.97	0.98	0.97	59
5	0.97	0.92	0.95	39
6	0.95	0.98	0.96	42
accuracy			0.97	261
macro avg	0.97	0.97	0.97	261
weighted avg	0.97	0.97	0.97	261

Figure 4.4.3.1 InceptionV3 Classification Report

Confusion matrix:

Here CNG is accurately identified.

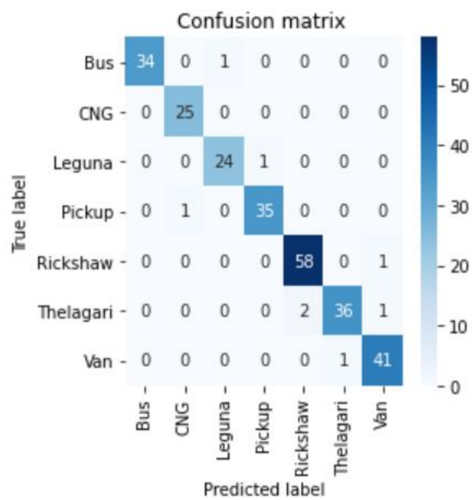

Figure 4.4.3.2 InceptionV3 Confusion Matrix

Training and Validation Accuracy:

We saw that training and validation accuracy are not close. Considering the gap we saw that gap range is little bit more.

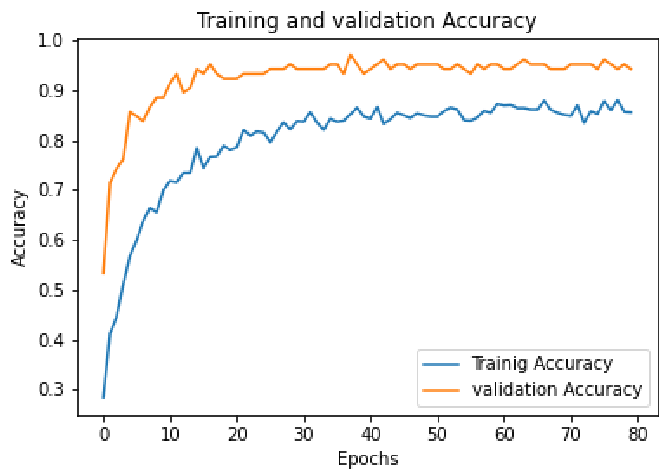

Figure 4.4.3.3 Training and Validation Accuracy (InceptionV3)

Training and Validation loss:

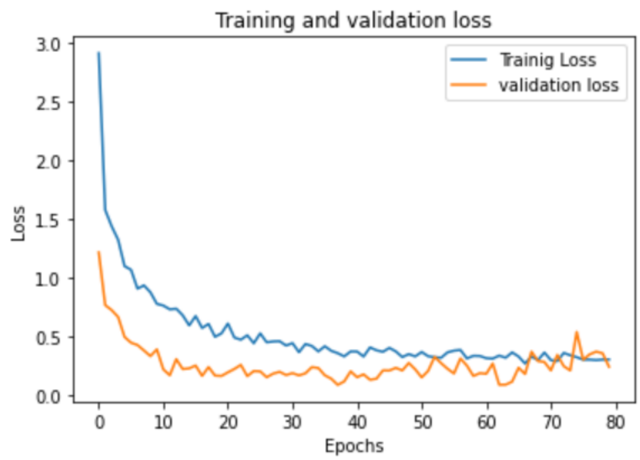

Figure 4.4.3.4 Training and Validation loss (InceptionV3)

4.4.4 CNN

Our Vehicle-NN model achieved 96% accuracy.

Training and Validation Accuracy:

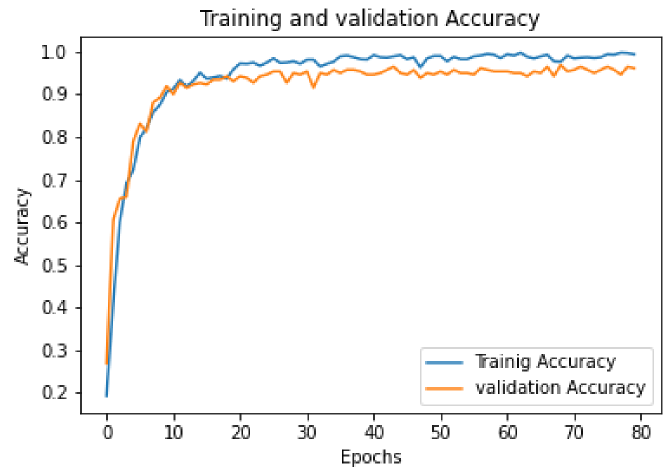

Figure 4.4.4.1 :Training and validation Accuracy(CNN)

Training and Validation Loss:

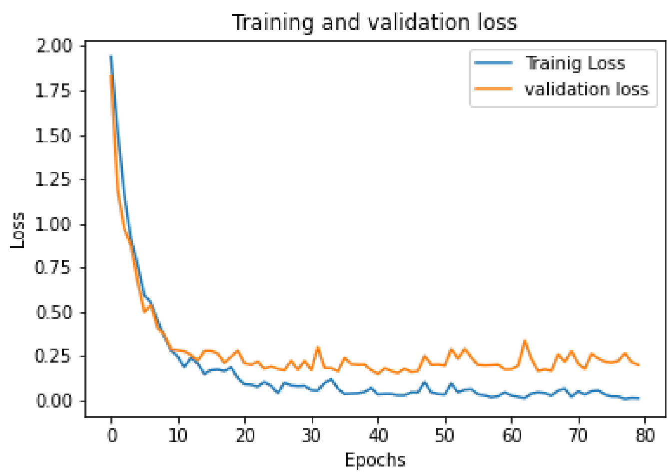

Figure 4.4.4.2 :Training and validation loss(CNN)

CHAPTER 5

CONCLUSION

5.1 Future Work

Our paper, we show the accuracy of our own Vehicle-NN along with MobileNet, InceptionV3, VGG16 models and present a comparison among these models. We collect only 1304 data in this pandemic situation. In future, we want to collect more data and update our Vehicle-NN model with various ways for transfer learning.

5.2 Conclusion

In this article, we developed Vehicle-NN model for Bangladesh local vehicle classification using our own data. In this study the result of gain is encouraging. Our proposed classifier is effective in object detection. And in future, it will be a contribution in our Bangladesh vehicle sector. [8]

REFERENCE

[1] B. Daya and P. Chauvet, "Neural network approach for the identification system of the type of vehicle," 2010 Int. Conf. Comp. Inte. and Com. Net. IEEE, pp. 162-166, November 2010.

[2] C.M. Bautista, C.A. Dy, M.I. Mañalac, R.A. Orbe, M. Cordel, " Convolutional neural network for vehicle detection in low resolution traffic videos," 2016 IEEE Reg. 10 Symp. (TENSYMP), pp. 277-281, May 2016.

[3] J.J. Sanchez-Castro, J.C. Rodríguez-Quiñonez, L.R. Ramírez-Hernández, G. Galaviz, D. Hernández-Balbuena, G. Trujillo-Hernández, W. Flores-Fuentes, P. Mercorelli, W. Hernández-Perdomo, O. Sergiyenko and F.F. González-Navarro, "A Lean Convolutional Neural Network for Vehicle Classification," 2020 IEEE 29th Int. Symp. Ind. Elec.(ISIE), pp. 1365-1369, June 2020.

[4] M. Hasan, and J.F. Nima, "Exer-NN: Cnn Based Human Exercise Pose Classification," 2019.

[5] M. Manana, C. Tu, and P.A. Owolawi, "A survey on vehicle detection based on convolution neural networks," 2017 3rd IEEE Int. Conf. Comp. and Com. (ICCC), pp. 1751-1755, December 2017.

[6] J.H. Bappy and A.K. Roy-Chowdhury, "CNN based region proposals for efficient object detection," 2016 IEEE Int. Conf. Img. Pro. (ICIP), pp. 3658-3662, September 2016.

[7] X. Hu, X. Xu, Y. Xiao, H. Chen, S. He, J. Qin, and P.A. Heng, "SINet: A scale-insensitive convolutional neural network for fast vehicle detection," IEEE Trans. Int. Trans. Sys. 20, vol. 3 ,pp. 1010-1019, 2018.

[8] M.S. Islam, F.A. Foysal, N. Neehal, E. Karim, and S.A. Hossain, "InceptB: a CNN based classification approach for recognizing traditional bengali games," Pro. Comp. Sci. 143, pp. 595-602, 2018.

[9] B. Hicham, A. Ahmed, and M. Mohammed, "Vehicle type classification using convolutional neural network," 2018 IEEE 5th Int. Cong. Inf. Sci. and Tech. (CiSt), pp. 313-316, October 2018.

[10] K.C. Wang, Y.D. Pranata and J.C. Wang, "Automatic vehicle classification using center strengthened convolutional neural network," 2017 Asia-Pacific Sig. and Inf. Pro. Asso. Annu. Sum. and Conf. (APSIPA ASC), pp. 1075-1078, December 2017.

[11] D. Zhao, Y. Chen and L. Lv, " Deep reinforcement learning with visual attention for vehicle classification," IEEE Trans. Cog. and Dev. Sys. 9, vol. 4, pp. 356-367, September 2016.

[12] B.A. Alpatov, P.V. Babayan and M.D. Ershov, "Vehicle detection and counting system for real-time traffic surveillance," 2018 7th Med. Conf. Emb. Comp. (MECO), pp. 1-4, June 2018.

[13] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel, "Backpropagation applied to handwritten zip code recognition," Neu. Comp.1, vol. 4, pp. 541-551, December 1989.

[14] What is Modern Technology and how is it changing << https://aging.com/what-is-modern-technology-and-how-is-it-changing/ >> Last Access on 4.4.2021 at 3.10 PM

Vehicle-NN

ORIGINALITY REPORT

21%	15%	10%	16%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	6%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	4%
3	Submitted to UNITEC Institute of Technology Student Paper	1%
4	Submitted to An-Najah National University Student Paper	1%
5	Benjamin Dowden, Oscar De Silva, Weimin Huang, Dan Oldford. "Sea Ice Classification via Deep Neural Network Semantic Segmentation", IEEE Sensors Journal, 2020 Publication	1%
6	Submitted to Uttar Pradesh Technical University Student Paper	1%
7	Renxi Chen, Xinhui Li, Shengyang Li. "A Lightweight CNN Model for Refining Moving Vehicle Detection from Satellite Videos", IEEE Access, 2020	1%