DETECTION AND CLASSIFICATION OF ROAD DAMAGE USING DEEP LEARNING APPROACH WITH SMART PHONE IMAGES

BY

SHARMIN SHILA
ID: 191-15-12398

FAHIMA TAHER BAYSHAKHY
ID: 191-15-12062

SUMAIA ZANNAT MITU
ID: 191-15-12809

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

Supervised By

Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

Most. Hasna Hena
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This Project/internship titled **"Detection and Classification of Road Damage Using Deep Learning Approach with Smart Phone Images"**, submitted by Sharmin Shila, ID No: 191-15-12398, Fahima Taher Bayshakhi, ID No: 191-15-12062, and Sumaia Zannat Mitu, ID No: 191-15-12809 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfilment 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 *25 January,2023.*

BOARD OF EXAMINERS

28/01/23

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

Internal Examiner

Dr. Md. Monzur Morshed
Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dewan Mamun Raza
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

25.1.2023

Dr. Ahmed Wasif Reza
Associate Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that, this thesis has been done by us under the supervision of **Abdus Sattar, Assistant Professor, Department of CSE** Daffodil International University. We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By:

Most. Hasna Hena
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Sharmin Shila
ID: 191-15-12398
Department of CSE

Fahima Taher

Fahima Taher Bayshakhy
ID: 191-15-12062
Department of CSE

Sumaia Zannat Mitu

Sumaia Zannat Mitu
ID: 191-15-12809
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 thesis successfully.

We really grateful and wish our profound our indebtedness to **Abdus Sattar**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “Human Computer Interaction in Education” to carry out this thesis. 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 thesis.

We would like to express our heartiest gratitude to **Dr. Touhid Bhuiyan, Professor, and Head,** Department of CSE, for his kind help to finish our thesis and also to other faculty members and the staff of CSE department of Daffodil International University.
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 passion of our parents.

ABSTRACT

In cities, road surface monitoring is mostly done by hand which is a time-consuming and labor-intensive procedure. One of the most critical responsibilities is infrastructure maintenance work for traffic safety. To keep the road network safe, it must be assessed on a regular basis to identify potential threats and risks. We work on detecting and classifying road damage using deep learning approach in this research, which is a low-cost intelligence system. The goal of this work is to investigate the detection and categorization of road damage from road surface photographs using deep learning concept. This study used different transfer learning algorithms to categorize road damage in order to determine which algorithm performed better at detecting and classifying road damage. We divide damages into four groups: potholes, cracks, and revealing and rutting. For this research, we used a smartphone camera to collect data from the streets of Dhaka and processed with it. Our work uses various transfer learning deep neural network algorithms including VGG16, VGG19, ResNet50, MobileNetV2, EfficientNetV2 for classifying road damages, as well as for detection, and it outperforms earlier research. We got the highest 97.15% accuracy for ResNet50 and lowest accuracy 94.88% for MobileNetV2 and EfficientNetV2, 94.31% accuracy for VGG16 and 93.18% for VGG19.

Keywords: Road Surface, Road Damage, Deep Learning, Transfer Learning, Neural Network, Detection, Classification.

TABLE OF CONTENTS

CONTENTS	PAGE

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The metropolitan road system is the primary mode of transportation. Every day, several traffic accidents occur. Bangladesh is ranked 106th in the world for the most road accident related deaths, according to a statistic based on data from 183 countries. In the first eight months of 2021, 3,502 people were killed and 3,479 were injured in 3,701 car crashes. With the extensive building of highways, the government and the general public are concerned about the safety of transportation systems. Road deterioration, notably potholes and cracks, is not only inconvenient but often unsafe. One of the most critical responsibilities is infrastructure maintenance work for traffic safety. To keep the road network safe, it must be assessed on a regular basis to identify potential threats and risks. For real-world application, professional workers are usually in charge of detecting road damages, it is lengthy and expensive operation. As a result, we require an inexpensive computerized process for detecting road surfaces. There are many computerized processes that use sensors to detect road deterioration, but this procedure is expensive. An infrared examination is still the most common method for acquiring data on three dimensional streets. We work on detecting and classifying road damage using deep learning approach in this research, which is a low-cost intelligence system. Past few years have seen considerable advancements in the field of computer vision, displaying great success across many different study fields. In this study, we use transfer learning architectures to detect and classify road damage objects. Previous work on image processing-based road damage detection used a different approach. Many earlier studies in this topic were discovered. They use several CNN and transfer learning methods, and unlike other object detection datasets, there is no common dataset on road surfaces. We used a smartphone camera to gathered data from Dhaka's streets. In order to increase performance, we utilize transfer learning-based architectures VGG16, VGG19, ResNet50, MobileNetV2, and EfficientNetV2 and compare which technique is more effective for identifying and categorizing road damage.

The following is a summary of our work's main contribution:

1. Own dataset comprising 1776 photos that were physically gathered from the streets of Dhaka.
2. A transfer learning model with the highest accuracy of 97.15 percent for identifying road damage
3. An enhanced ResNet50 model for Bangladeshi road damage. Compared to the detection models developed by Maeda et al. [2], our model performed better.

1.2 Motivation

One of the most common causes of traffic accidents and their consequences for a country's economy is road damage. The diagnosis of road diseases is normally the responsibility of experienced professionals. Because this method is both time consuming and costly, maintaining road safety in developing countries is extremely challenging. We need an affordable computerized program to address this problem for detecting road damage that will assist the government in addressing the issue of road safety. There are many automated systems that use sensors to detect road deterioration, but this procedure is expensive. The problem can be solved with a deep learning-based approach.

1.3 Problem Statement

We discovered that there have been numerous earlier studies in this sector. They use a variety of methods, and unlike other object identification datasets, there is no uniform dataset for road surfaces. We used a smartphone camera to collect data from the streets of Dhaka. At this work, we use transfer learning-based architectures to compare which architecture is superior for detecting and classifying road damage into four different categorizations: Pothole, Crack, Rutting and Reveling.

1.4 Research Questions

Following are the key inquiries on which this study is focused:

- Which transfer learning technique performs better for categorizing and detecting deep learning base road damage?

- What type of road damage are there?
- What is the cheapest way to detect road damage?

1.5 Research Methodologies

In this segment of our study report, we go over the Collected Data, Model Structure, Learning Rate, Model Analyzer, and Model Evaluation, as well as Performance Analysis. The effectiveness of the suggested paradigm will be explored at the conclusion of this chapter.

1.6 Research Objectives

There are a number of advantages to employing deep learning approach.

- Create a reliable model to more correctly identify road damage.
- To motivate the software engineers to use this model.
- To determine the transfer learning strategy that performs the best.
- To divide the four types of surface degradation on the road into categories.
- Collect smart phone-based data from Dhaka which can be used for researcher in future.
- Reduce the road maintenance cost.
- Road surface monitoring will be easier.

1.7 Expected Outcome

By implementing our research, we expect that the construction industry will change how they work to ensure that the four categories of damage—potholes, cracks, rutting, as well as raveling—can be readily diagnosed. This will result in quicker, cheaper, and more efficient road repairs. Furthermore, it can help in identifying the areas that require repairs the quickest, hastening the restoration of damaged roadways. Future improvements to the proposed collection will include the use of cellphones and more images from various locations throughout Bangladesh. It is anticipated that working with research institutions and towns in Bangladesh will produce a significantly larger collection for a more accurate estimation method. Additionally, work will be done to create and enhance our model to let CCTVs installed on roadways discover damage in instantaneously and notify the relevant people of the whereabouts and nature of the damages to aid in decision-making. In order to help future study in this field, we create a new dataset that consists mostly of photographs of road damage and evaluate five different strategies to determine which one is most effective at doing so. This task may involve the transportation network. Expand the boundaries to ensure that the region that requires the most immediate replacement project can be determined in order to determine the cost of fixing road damage.

1.8 Project Management

We made the choice to use a smartphone camera to make our own dataset. Every several days, we would travel to various Dhaka city roadways to gather statistics. We used the camera on a smartphone to take 1776 pictures. We have already incurred financial costs there. The building of the model didn't cost us any money. For almost five months, we have been developing our system for identifying and categorizing road damage. In the interim, we have finished writing the entire document and paper.

1.9 Thesis Organization

We covered more study in the following chapters, including the chapter two where we discussed about background of the study. In chapter three, we discussed research methodology. In chapter four, we reviewed the experimental result and discussion. In chapter five we discussed impact on society, environment and sustainability. Finally, we discussed the conclusion and future work in chapter six, which included boundaries, and future employment.

CHAPTER 2

BACKGROUND

2.1 Terminologies

Our main objective is to locate any road defects in the Bangladeshi capital because road damage, particularly potholes and fractures, is not only annoying but frequently hazardous. Maintenance work on the infrastructure is one of the most important responsibilities for traffic safety. Regular inspections of the road network are necessary to find potential threats and vulnerabilities if it is to remain secure. Recently, deep learning techniques have received a lot of attention, especially when building data-driven detection models. With the use of these methods, we created models. We conducted some research in comparison to other works. We were inspired by some eminent scientists who are making significant contributions to this subject.

2.2 Previous Literature

They develop a classifier that is much less susceptible to lighting noise, heavy casting, and other factors. In comparison to conventional methodologies, the current method used by people to undertake road pothole audits [5] learns without requiring any extraction or computation, the feature is applied automatically.

In this investigation, the histogram was the main emphasis. The portrait region feature was then obtained, and a non-linear SVM kernel was added to select the target. The findings demonstrate that it may be simple to locate the pothole in this inquiry. An algorithm based on deep learning classifier, namely a Convolutional Neural Network, was utilized to determine road fracture harm from photographs [14].

For automated fracture recognition, a deep Convolutional Neural Network [15] and an inexpensive sensor were described. A Convolutional Neural Network architecture that can adapt without first doing a feature engineering technique was demonstrated in this experiment. The supplied photos were carefully tagged before the model feed. To solve the problem, deep learning is an inexpensive technique for identifying road potholes.

There are a growing number of papers focused on the work of reporting road deterioration and surface conditions. For road pothole detection tasks, Support Vector Machine (SVM) is the backbone of a common machine learning technique [16].

N. Hoang created a pothole detection intelligence system [17] and evaluated it employs the least square support vector machine (LSSVM) and the artificial neural network as two machine learning algorithms (ANN).

The precise categorization of the LSSVM method was around 89 percent., while Utilizing ANN was around 86 percent. Ryu [18] devised a pothole identification system that reads photos from an optical device deployed in an automobile and then proposes a way for finding potholes from the received data. The suggested method is utilized to remove dark parts from the pothole using a histogram, followed by a morphological filter to close it. The nominee pothole zones are then chosen based on characteristics like volume and compactness. Finally, it is decided whether the applicant certain areas have potholes by contrasting pothole and background. The rating dependability was found to be 73.5 percent using this method.

A. Akagic et al. 2017 [21] offer a pothole depending on an eyesight detection mechanism. The procedure begins by separating Segmenting the item based on the damaged regions within the RGB color model. So, the quest is restricted to the area from which value is derived (ROI). Their method can be used as a first stage for more supervised methods. The accuracy of ROI is crucial to the system's integrity, and 82 percent of them are right.

2.3 Research Gap

- There is no extensive shared dataset for road damage.
- There aren't many studies on road damage in Bangladesh.
- There isn't a transfer learning model available that is more accurate.

Table 1 below shows the research gap in detail, and we used this information to continue the study.

Table 2.3.1: Related Works

Paper	**Author**	**Year**	**Objective**	**Data**	**Methodology**	**Findings & Limitations**
Road Damage Detection and Classification with Detectron2 and Faster R-CNN [1]	Vung Pham, Chau Pham & Tommy Dang	2020	To provide X101-FPN base model to detect road damages.	Global Road Damage Detection Challenge 2020	Faster R-CNN	F1 scores for the consistently producing testing one as well as testing two groups of the challenge come out to 51.0% and 51.4%, respectively. Its primary drawback is that it takes longer to train than other methodologies and that its forecast time is higher.

Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images. [2]	Hiroya Maeda, Toshikazu Seto, Takehiro Kashiyama, Yoshihide Sekimoto & Hiroshi Omata	2018	To create a new, extensive data set of road damage, train a smart phone-based damage detection model, and then assess the model.	Images of Japan road.	CNN	Using MobileNet and Inception V2, they were able to attain recall and precision rates of more than 71% and 77%, respectively. Compared to the approaches that use highly accurate sensors, the model used in this study is not very accurate, however
A Deep Learning Approach for Road Damage Detection from Smartphone Images [3]	Abdullah Alfarrarjeh, Dweep Trivedi, Seon Ho Kim & Cyrus Shahabi.	2018	To create an image-based surveillance system for city streets.	The photographs from the IEEE BigData Cup Challenge depict urban street scenes in seven different geographical regions of Japan.	YOLO	The F1 score for this solution was up to 0.62. Although this method is cheap, it might not be very accurate.

Transfer Learning-based Road Damage Detection for Multiple Countries. [4]	Deeksha Arya, Hiroya Maeda, Sanjay Kumar Ghosh, Durga Toshniwal, Alexander Mraz, Takehiro Kashiyama & Yoshihide Sekimoto	2020	To provide a universally applicable, consistent paradigm for identifying and classifying road damage	Road images of Japan, India, Czech.	16 deep neural network models.	Their system was capable of finding damages in more than one country with less expensive automated system for road damage detection. According to them same prototype can be extented for other countries.
Road Damage Detection using Deep Ensemble Learning [6]	Keval Doshi, & Yasin Yilmaz	2020	To offer an ambient model for effective damage classification and detection for roads.	Collected image using smartphone from India, Japan.	YOLO-v4	On the test 1 dataset and the test 2 dataset, respectively, received an F1 score of 0.628 and 0.6358.Need more possible extensions for improving the dataset

Road Damage Detection Acquisition System based on Deep Neural Networks for Physical Asset Managemen t [8]	A.A. Angulo, J.A. Vega-Fernánd ez, L.M. Aguilar-Lobo , S. Natraj & G. Ochoa-Ruiz	2019	Offer a service that is intended to be the acquisition part of a property manageme nt solution that can assist governmen t organizatio ns with planning ahead or infrastructu re maintenanc e companies with using preventativ e maintenanc e techniques.	To create the dataset, crowdsour ced photos that had been manually labeled for further processing were added to photograp hs from numerous public databases.	DNN and Generic object dictection.	They achieved 0.98 accuracy.
Road Surface Classificatio n with Images Captured From Low-cost Camera - Road Traversing Knowledge (RTK) Dataset [9]	Thiago Rateke, Karla Apareci da Justen & Aldo von Wangen heim	2019	To identify the surface types and surface quality.	RTK dataset which collected images by an inexpensiv e camera with low resolution.	CNN	Result accuracy from Rtk dataset is 95.73% In their work Dataset was imbalanced , in Future they want to focus on the crack category more.

An Approach for Extracting Road Pavement Disease from HD Camera Videos By Deep Convolutional Networks [19]	Wei Xia	2018	From camera videos, extract road surface diseases in a complicated setting.	Using the labeling procedure, created their own dataset.	CNN	They designed a poorly controlled way to generate road damage dataset to reduce the manual burden. More model need to apply.
Convolutional neural networks based pothholes detection using thermal imaging [12]	Yukti Bhatia, Rachna Rai, Varun Gupta, Naveen Aggarwal & Aparna Akula	2019	Implementation of information improvement approaches in the problem area to introduce a deep study approach to CNN, a novel approach is thermal imaging.	Used thermogra p-hic camrea to gather the data.	Thermal Imaging	They had a 97.08% accuracy rate.
Road Damage Detection Based on Unsupervised Disparity Map Segmentation [13]	Rui Fan & Ming Liu	2019	introduces an alternative unattended disparity map segmentation-based method for detecting road deterioration.	ApolloScape, KITTI, EISATS Dataset	Roll angle estimation	They did so by exerting as little skill as they could in respect of streo sliding angle and road differential.

Road Damage Detection and Classifiaction with Faster R-CNN	Zhixiang Wang, Wenzhe Wang, Bin Wu & Sixiong Yang	2019	To identify road damage, include a smart phone-based model.	Road Image of Japan	Faster R-CNN	They achieved Mean F1-score 0.6255 According to them Network structure didn't optimized. Need to try other techniques for improving performance.
A Deep Learning-Based Approach for Road Pothhole Detection in Timor Leste [11]	Vosco Pereira, Satoshi Tamura, Satoru Hayamiz & Hidekazu Fukai	2018	A model-based solution for detecting photos of road potholes at a reasonable cost.	Road images from South Africa, Bangalore and Rangoon.	CNN	Achived Mean F1-score 0.9960.

2.4 Scope of the Problem

- Low cost automated system for detect road damage for Bangladesh.
- Detect different type of road damage.
- Classify road Damage to understand the damage type.
- Get more accurate result than the previous work.
- Determine which approach for deep learning-based road damage detection performs best.

2.5 Challenges

There seem to be numerous models that may be applied to the identification of road damage, which makes it essential that we pick the most dependable one. Finding a model that could produce the most accurate outcome possible was not a simple feat. Additionally, we assess how well our models perform generally across a variety of different platforms.

Deep learning calculations demand enormous volumes of data, yet there is no standard dataset for studies on road damage. Therefore, we manually gathered data from Dhaka's streets. We had a very difficult time collecting statistics on road damage.

2.6 Summary

Our research is driven by the question of which strategy performs better on our own dataset. In this study, we will use the CNN-based models to evaluate which outperforms and categorize images into four groups.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

With 4 multiple categories of images and transfer learning techniques, this work's best accuracy was 97.15%.

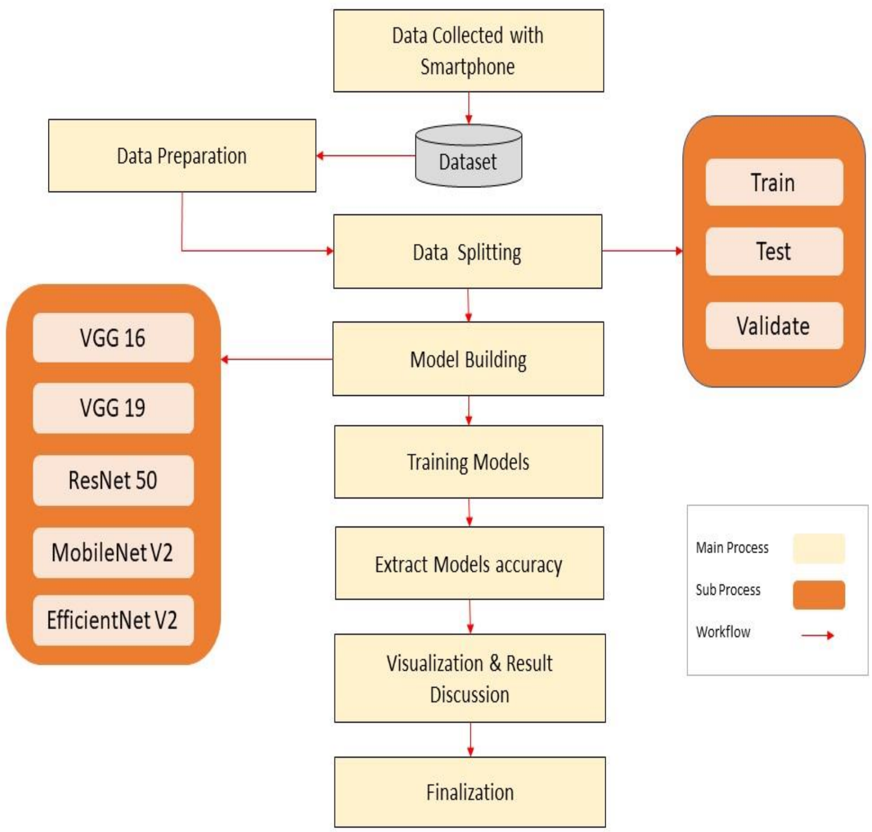

Figure 3.1 Methodology Process

3.2 Research Subject and Instrumentation

Our study concentrates primarily on Deep Learning Approaches and is titled "Detection and Classification of Road Damage Using Deep Learning Approach with Smart Phone Images." The experiment's models were created using the Google Collaboratory's Intel Xeon CPU. The scikit-learn, pandas, and NumPy libraries were used to develop a deep learning-based model.

In general, the programming language called Python was employed. We made advantage of Google Collaboratory to build and simulate our models.

3.3 Research Method

There are mainly two types of method that researcher commonly use :-

1. Qualitative Method
2. Quantitative Method

The focus of qualitative research is on sentiments and how they are construed, whereas the focus of quantitative research is on quantities and statistical facts. The comparative evaluation of parameters and testing of hypotheses is made easier by quantitative methodologies. We may explore concepts and encounters more deeply thanks to qualitative methodologies. Therefore, the nature of our research is quantitative.

3.4 Data Collection

There is no large-scale shared collection for road damage detection like there is for other item detection datasets. We decided to create our own dataset from 1776 photos which we collected using a smartphone camera from the streets of Dhaka for this study. Disastrous part was how laborious it was to personally capture the data from the field. This dataset was produced in order to test a theory about image classification. Over 1776 square images make up this dataset, which is divided into four categories. Approximately 1335 photos are used to train the model, The remaining 176 photos are utilized for testing, while the remaining 265 photos are used to perform validation.. There are four classes of training and validation data sets in the study. Then we resize and label images with four category Crack, Pothole, Raveling and Rutting. Our sample data is given below.

Figure 3.4.1 Sample Collected Dataset

3.5 Data Pre-Processing

The entire data set was gathered using a smartphone on the streets of Dhaka, Bangladesh, where the photographs were not all the same size. It was quite challenging to train and test the dataset. A fixed dimension and quality of images are present in this dataset version. All of the photographs have been square-sized converted in accordance with the project specifications. To prepare our original data for our models, we resize the image to 224 X 224 pixels and use the vgg16 pre- processing filtering method. Additionally, we used the RGB color mode to train our model.

Figure 3.5.1 Sample Pre-Process Data

3.6 Transfer Learning Based Classification

Transfer learning allows us to represent features from a model that has already been trained, saving us the time and effort of having to create a new model from scratch. Frequently, a periodically re- model is educated on a sizable dataset like ImageNet, and our own neural network can leverage the weights generated from the trained model for any additional applications that are similar. These recently created models can be utilized to provide predictions on relatively untested jobs directly or to train algorithms for related applications. This strategy lowers the generalization error while also cutting down on training time. Our proposed models were mainly developed by transfer learning concept with four popular architectures: VGG16, VGG19, ResNet-50, MobileNetV2 and EfficientNetV2. Our models were trained by our own dataset that contain 1335 train images, 265 validation images and 176 test images of four road damage classes. The typical transfer learning workflow is given below:

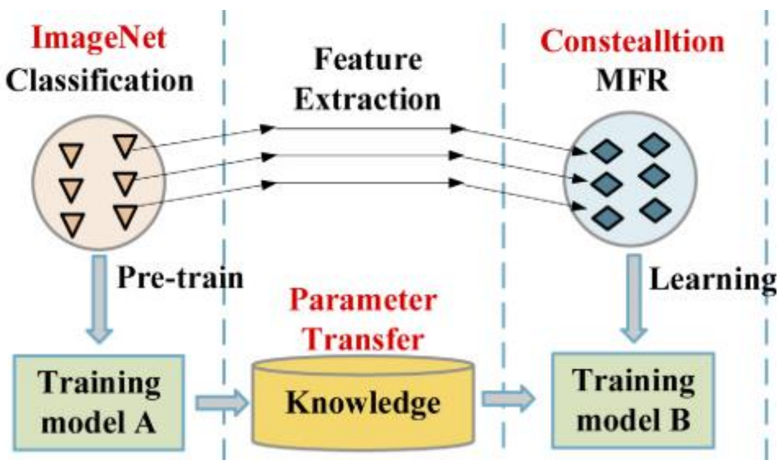

Figure 3.6.1 Transfer Learning Workflow

3.7 VGG16

K. A. and Simonyan proposed the convolutional neural network model VGG16. Professor Zisserman of Oxford University wrote an essay titled "Very Deep Convolutional Networks for Larger Data Recognition."

Input is routed to the channel from a measured picture (224, 224, 3). T wo of the initial layers include the same spacing and 64 bands with a sampling frequency of 3*3. after a maximum pool level of stride (2, 2), two phases have fully connected polarizer having 128 levels size and batch size (3, 3). A max-pooling step (2, 2) phase serves as the next level that is an exact duplicate of the level above it. Thereafter, 256 filtration are dispersed across 2 convolution layers with 3 and 3 filter widths. Afterwards when, a maximum pool coat is applied 2 different groups of three convolution levels. The same spacing and 512 filtration are prevalent in each filtration (3, 3). After that, that photo is sent to the bundle of two convolution levels. We develop 3 and 3 lenses in such convolution as well as maximum pooling levels instead of 7*7, ZF-11*11, likewise AlexNet filtering. Several of the levels also exploit 1*1 frames to adjust the system that uses count. To remove the spectral characteristics of the image, 1-pixel spacing is added after each convolution level.

After introducing a convolution once added, the stack's maximum pooling level, we obtained a (7, 7, 512) saliency map. That result is flattened to create a feature representation also with value 1, 25088. The following are the three fully linked levels: the first stack yields a vector with a size of (1, 4096), the second level as well generates a vector with a size of (1, 4096), and the third layer output values a vector with a size of (1, 1000), which is used to create the generative model to classify classes. Every hidden layer makes use of ReLU's activation function. ReLU promotes speedier learning and lessens the likelihood of fading gradient problems, making it more computationally efficient.

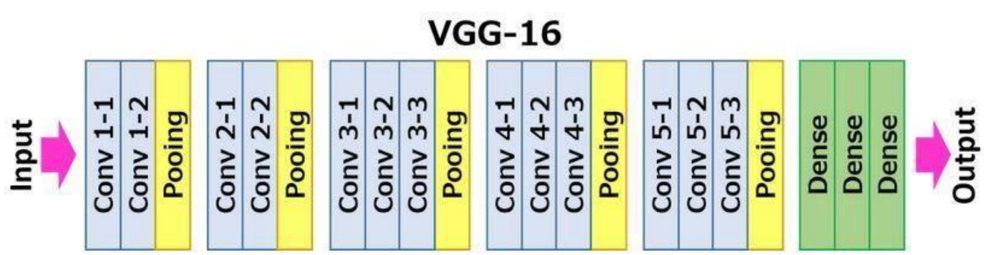

Figure 3.7.1 VGG-16 Architecture

3.8 VGG19

The Multilayer Perceptron neural Centre inside the Department of Engineering Science at the University of Oxford is where the programmed Convolutional Neural Network VGG-

19 was developed. The number of trainable layers is represented by the number 19. 16 Convolutional layers and three completely connected layers.

This connectivity was given a predetermined (224 224) RGB picture to show the matrix's shape (224,224,3). The sole step in the preparation process was to discover the average Color that each pixel in the training set had been given. We were able to cover the entire image using strides of 1 pixel and cores of length (3 * 3). To maintain the image's spatial resolution, spatial spacing was applied. Stride 2 was used to do max pooling on a 2 by 2 pixel frame. Following this, a batch normalization unit (ReLu) was utilized to introduce nonlinearity into the model, improving classification precision and computation efficiency. This model performed far better than previous models utilized hyperbolic tangent or sigmoid processes. three completely linked layers were used, the very first two of which consisted of a size of 4096, the third of which was a softmax function and contained 1000 units for classification utilizing the 1000-way ILSVRC.

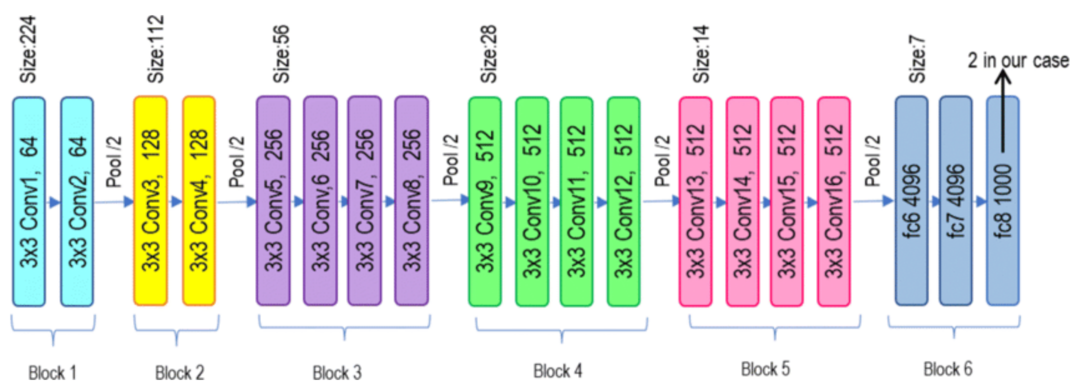

Figure 3.8.1 VGG-19 Architecture

3.9 ResNet 50

Every one of the 5 phases of the ResNet-50 architecture has an ideological wall and a convolution. Each ideology wall and convolution wall contains three convolution levels. There are around 23 million generative model parameters in the ResNet-50.

The Residual Frames theory was introduced in this design to address the fading gradient problem. In this network, we employ a method known as skip connections. The skip link skips over some intermediary levels in order to connect phase action potentials to subsequent layers. As a result, there is a leftover block. To form resnet, these extra blocks are piled. Instead of having levels learn the underlying mapping, the approach used by this network is to allow the network adopt the residual mapping.

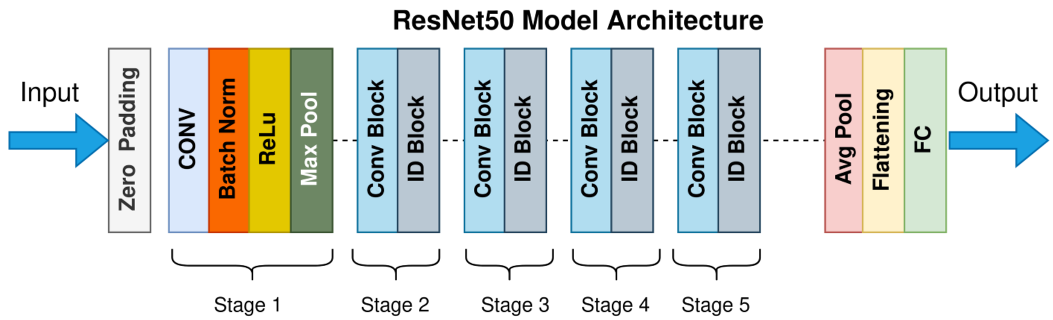

Figure 3.9.1 ResNet50 Architecture

3.10 MobileNet V2

MobileNetV2 is a neural network convolutional architecture that intends to perform well on mobile platforms. The bottle - neck layers are linked together by residual connections, and it is constructed on an inversion residual structure.

In MobileNetV2, there are two distinct types of blocks. There is one stride left in block one. Another one is a block that can contract by two strides. There are 3 stages for each type of block. This time, the first layer consists of 11 convolution layers with ReLU6. The second layer is the depth-wise convolution. Once more using an 11 convolution for the third layer, non-linearity is not present this time. If ReLU is applied once more, deep networks are considered to only contain the ability of the a linear predictor on the non-zero volume zone of the return area.

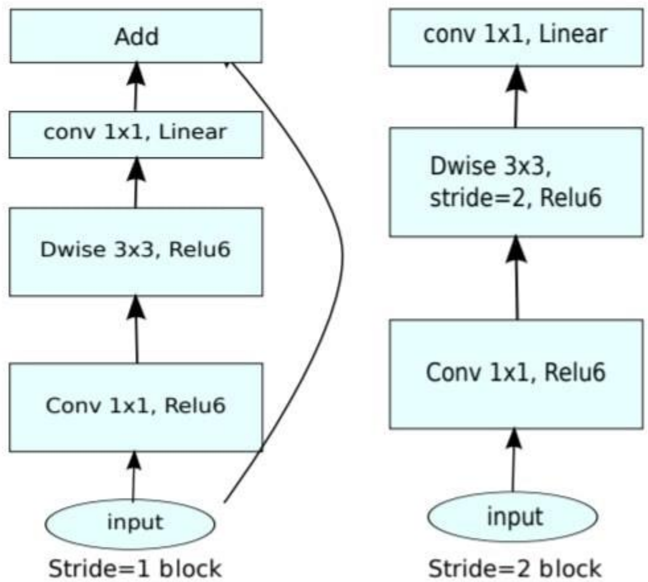

Figure 3.10.1 MobileNetV2 Architecture

3.11 EfficientNet V2

In terms of training time and parameter efficiency, a brand-new category of convolutional networks dubbed EfficientNetV2 beats older models. These models were developed by Mingxing Tan and Quoc V. Le using training-aware neural network models search and scaling, which jointly improved training time and parameter effectiveness. A search region that had been widened to include fresh processes like Fused-MBConv was utilized to hunt up the models. This study shows that EfficientNetV2 architectures train upto to 6.8 times faster despite being smaller than state-of-the-art models.

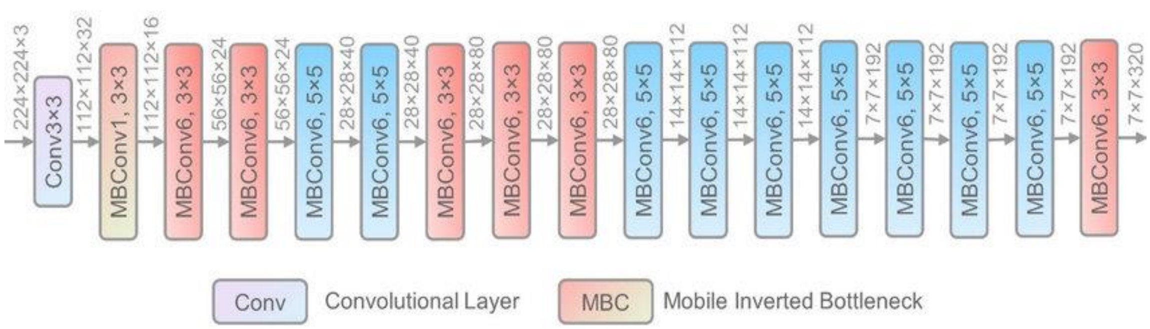

Figure 3.11.1 EfficientNetV2 Architecture

3.12 Model's learning rate and optimizer

For quicker outcomes at the start of learning, Adam Optimization 0.001 adjusts network weights according to the training data. The authors of the research altered the link weights by switching from one loop to the following. Revised weights in accordance with k (1), in which k is the iterative process index.

$$w_{lij}(k+l) = w_{lij}(k) - \mu g_{lij}(k) \qquad l = 1,2,\ldots L \tag{1}$$

Model size appears to be immaterial because the simulation time performs similarly well for modeling techniques of sizes 1x and 10x. The selected learning rate is the most important hyperparameter for the model. The procedure (2) for computing each upgrade with learning rate decline is as follows:

$$LR = initial\_lrate \times (1 / (1 + decay \times iteration)) \tag{2}$$

We quickly decreased the error with the lowest learning ability decrease of 0.000001. We previously accessed Keras' sponsored learning rate through a callback mechanism. The callback method changes the modeling weight to slow learning when the model no longer improves.

3.13 Summary

In order lamenting the drop, which serves as an error function for multiple classification issues, we use categorical cross-entropy. Then, considering that this is a categorical variable problem, the optimizer, which aids in performing gradient descent, and the measurements set accuracy.

The performance of each Convolutional Neural Network design technique will be compared in the following chapter, along with a four-group classification of road damage.

CHAPTER 4

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Evaluation Technique

Our investigation's major focus is on deep learning techniques and the topic is the detection and classification of road damage using a deep learning approach with smartphone images. We used Google's Tensor Flow and the Keras deep learning framework for image processing to carry out our research in Python. The experiment's models were created using the Google Collaboratory's CPU. The scikit-learn, pandas, and NumPy libraries were used to develop a deep learning-based model. We made advantage of Google Collaboratory to build and simulate our models. The total number of photos is around 1776 where we differentiate between training, test, and validation data. We combined RGB color mode with category mode. The model that was suggested in the study performed well.

4.2 Experimental Result & Analysis

Measuring metrics requires the use of a confusion matrix. A table structure called a confusion matrix, commonly referred to as an error matrix, is used in deep learning to demonstrate the efficacy of a method of controlled learning. We determine True Positive (TP), False Positive (FP), True Negative (TN), and False Negative(FN) using a confusion matrix.

Accuracy

The most obvious performance indicator is accuracy. The ratio of precisely anticipated observational data to all observations is, to put it simply, what it is.

$$Accuracy=(TP+TN)\ /(TP+FP+FN+TN) \qquad (3)$$

Precision

The precision ratio measures how well all anticipated positive studies match the measurements that were successfully predicted as positive. How many individuals have survived unscathed, to use metrics.

$$Precision=TP/(TP+FP) \qquad (4)$$

Recall

Remember the ratio of all actual class observations to precisely forecasted good results? How many of all the things that have truly true information did we mark?

$$Recall = TP/(TP+FN) \quad (5)$$

F1 Score

It is simple to recall and exact to use the weighted F1 Ranking. When determining the score, both false positives and false negatives are taken into account. F1 is usually more useful than accuracy despite being tougher to explain on the surface, especially if your class distribution is asymmetrical.

$$F1\ Score = 2*(Recall*Precision) \ / \ (Recall*Precision) \quad (6)$$

Sensitivity

The proportion of positive samples that result in a favorable finding when a particular test is added to a model without changing the samples is known as the real positive rate, also known as sensitivity.

$$\text{Sensitivity} = TP/\ (TP + FN) \quad (7)$$

Specificity

The percentage of samples that test negative when the test is used is known as the true negative rate, or specificity, in the setting of an unaffectedly negative model.

$$\text{Specificity} = TN/\ (TN + FP) \quad (8)$$

Model Analysis

By using deep learning techniques to develop models, we looked for the greatest performance. Table 4.2.1 shows the effectiveness of the model analysis. Therefore, it is evident that when it used with our dataset, transfer learning techniques performed better. They are all between 93% and 97% accurate. We must now choose the best model out of the numerous available options.

Table 4.2.1 Performance Analysis Metrics

Serial	Algorithm	Accuracy (%)	Validation Accuracy	Validation Loss	F1 Score	Precision	Recall
1	VGG16	94.31	0.9585	0.2329	0.9407	0.9435	0.9385
2	VGG19	93.18	0.9623	0.2837	0.9303	0.9372	0.9268
3	ResNet50	97.15	0.9396	0.2561	0.9701	0.9715	0.9691
4	MobileNet V2	94.88	0.9434	0.2470	0.9525	0.9542	0.9517
5	EfficientNetV2	94.88	0.9623	0.2032	0.9478	0.9524	0.9442

Here are the F1 score, Precision, Recall, and testing accuracy, validation accuracy, validation loss ratings for the five CNN-based architectures (VGG16, VGG19, ResNet50, MobileNetV2, and EfficientNetV2) that were used.

Equations 3, 4, 5, and 6 were used to calculate the accuracy, recall, precision, and F1 score, which are all listed in table 4.2.1. There, we can observe that ResNet50 has the highest accuracy (97.15%), making it the best performer. The top two models' accuracy differences are 2.27 percent, suggesting ResNet50 can detect objects more accurately than MobileNetV2 and EfficientV2. The remaining algorithms are also capable of identifying objects close to the ResNet50 model, with accuracy values of 94.31%, 93.18%, 94.88, and 94.88 for VGG16, VGG19, MobileNetV2, and EfficientV2, respectively. Which method can more reliably identify the four target classes must be taken into account.

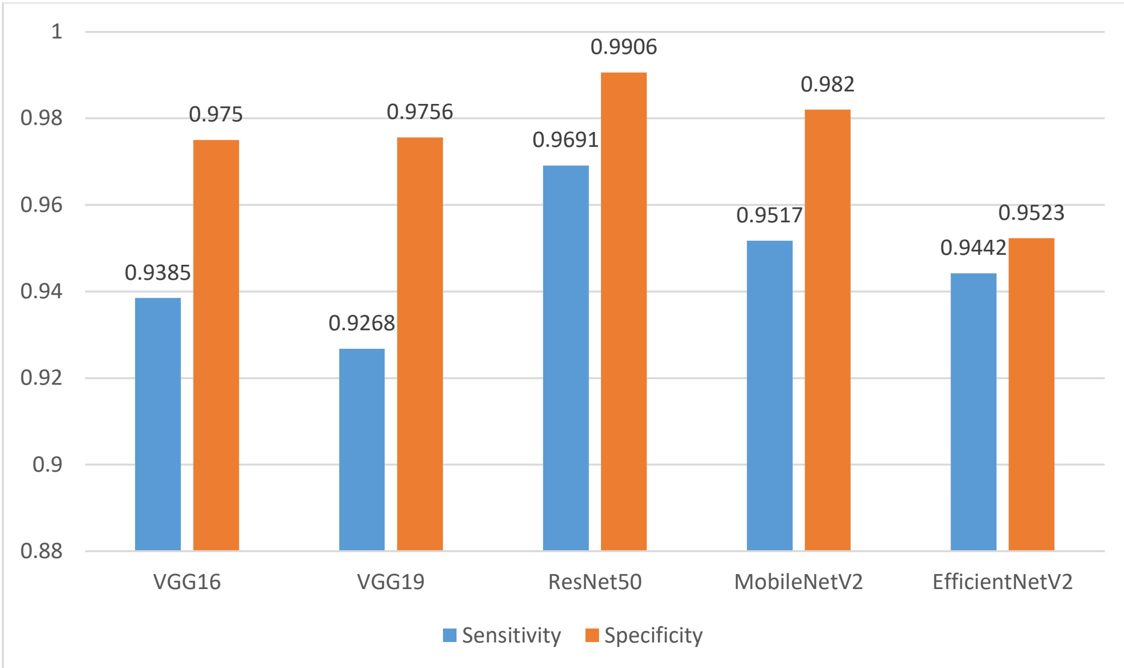

Figure 4.2.1 Sensitivity and Specificity of Models

After our model had been trained, we evaluated it using our test set of data. We already know that ResNet50 performs better than the other five models from earlier discussions. The sensitivity and specificity were calculated using Equations 7 and 8, and the results are shown in Figure 4.2.1, where we can see that ResNet50 has the highest score.

4.3 Visualization

Upon completion of the terms of classification accuracy, the task obtains a consecutive accuracy rate. Depending on the accuracy and confusion matrix generated by this work, it can be said that the suggested model is suitable for the task of identifying road damage. This is how this job has been performed.

VGG16 Plot: The graph demonstrated that accuracy grows with the number of epochs. We run 150 training epochs on our model. After 14 epochs we got 94.31% accuracy which is shown in figure 4.3.1.

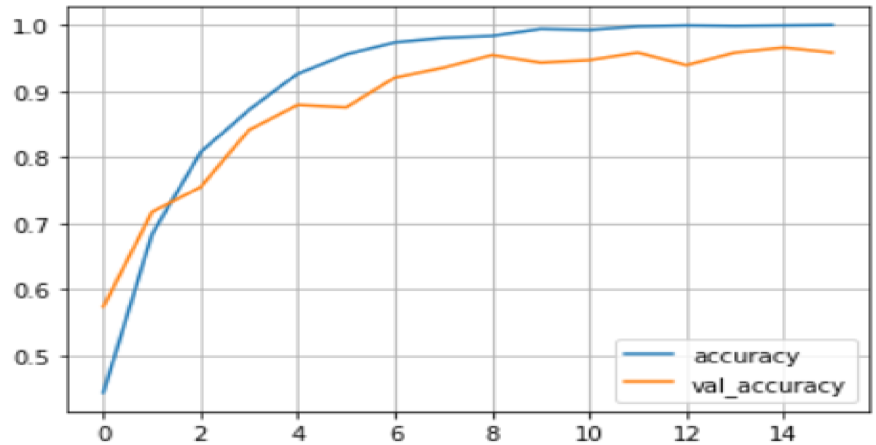

Figure 4.3.1 VGG16 Model Accuracy

VGG19 Plot: The graph demonstrated that accuracy grows with the number of epochs. We run 150 training epochs on our model. After 12 epochs we got 93.18% accuracy which is shown in figure 4.3.2.

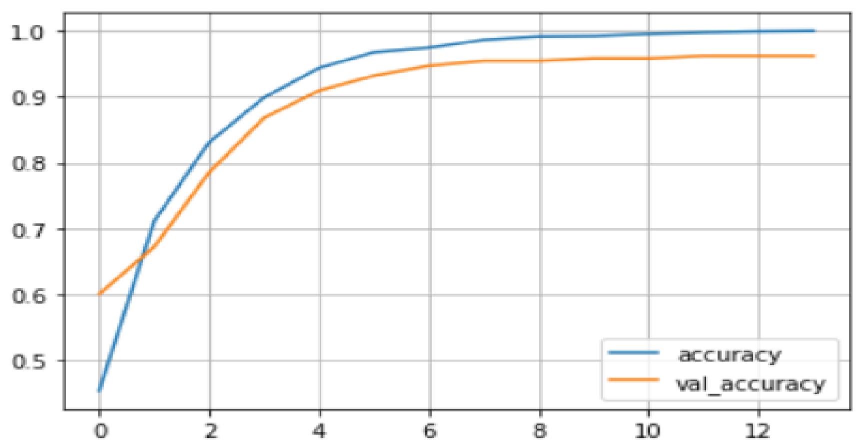

Figure 4.3.2 VGG19 Model Accuracy

ResNet50 Plot: The graph demonstrated that accuracy grows with the number of epochs. We run 150 training epochs on our model. After 6 epochs we got 97.15% accuracy which is shown in figure 4.3.3.

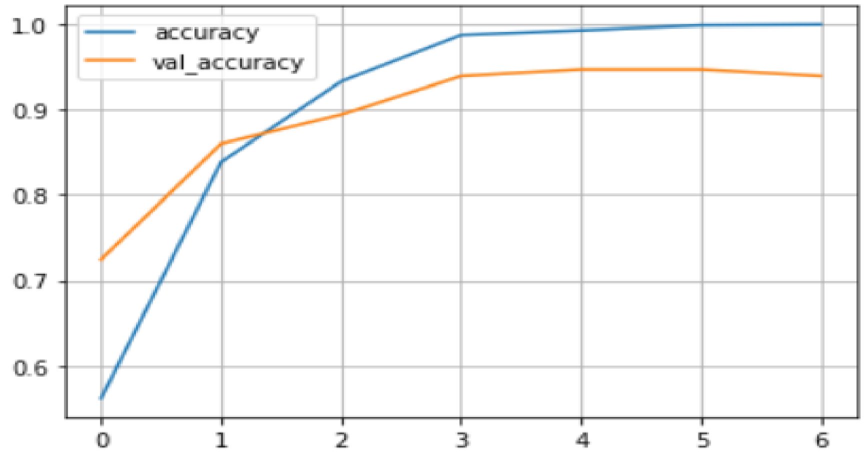

Figure 4.3.3 ResNet50 Model Accuracy

MobileNetV2 Plot: The graph demonstrated that accuracy grows with the number of epochs. We run 30 training epochs on our model. After 20 epochs we got 94.88% accuracy which is shown in figure 4.3.4.

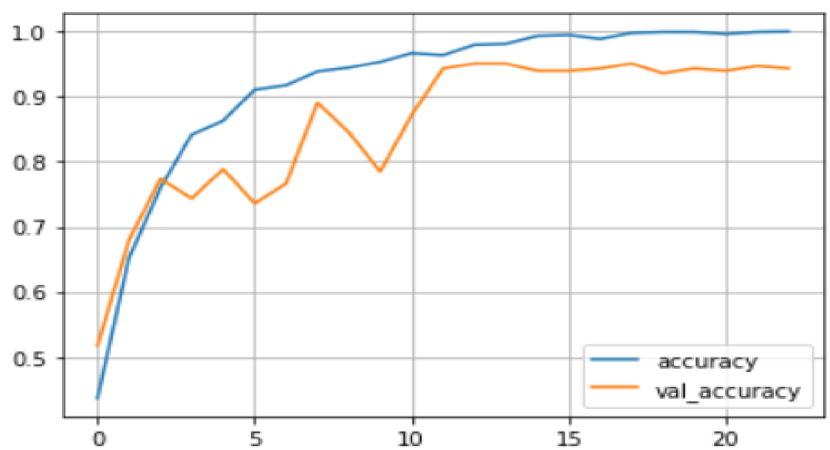

Figure 4.3.4 MobileNetV2 Model Accuracy

EfficientNetV2 Plot: The graph demonstrated that accuracy grows with the number of epochs. We run 150 training epochs on our model. After 17 epochs we got 94.88% accuracy which is shown in figure 4.3.5.

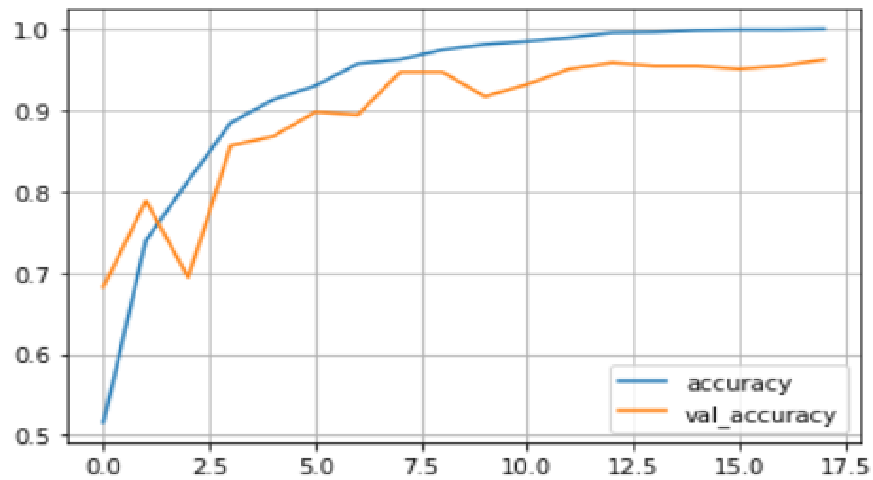

Figure 4.3.5 EfficientNetV2 Model Accuracy

4.4 Result Comparison

In comparison to other paper, this work provides us with the highest accuracy. Some researchers did work that was almost identical to this work, but no comparable analysis has been done for any Bangladeshi dataset. Table 4.4.1 reveals some connections between our study and similar earlier research.

Table 4.4.1 Comparative analysis of some earlier works

Works	Algorithms	Accuracy
Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images. [2]	MobileNet Inception V2	71% 77%
Transfer Learning-based Road Damage Detection for Multiple Countries. [4]	16 deep neural networks	77% - 94%
Detection and Classification of Road Damage Using Deep	ResNet 50 MobileNet V2 EfficientNet V2	97.15 % 94.88% 94.88%

Learning Approach with	VGG 16	94.31%
Smart phone Images	VGG 19	93.18%

From the following table, we can see that several studies used various varieties of algorithms and obtained varying degrees of accuracy, but the highest accuracy overall is 94%. For quicker running speed, 16 CNN-based algorithms with deep networks were deployed. They employ images with high resolution as well, but our accuracy is still considerably superior to theirs. Another researcher utilized MobileNet to appropriately categorize the image and obtained an accuracy of 71%; they also employed the Inception V2 model for obtain a higher accuracy, but they did not achieve a higher accuracy.

4.5 Result Discussion

For categorization and detection of road damage, we can recommend the ResNet50 model above all others because it outperforms them. For the ResNet50 model, we achieved the greatest accuracy of 97.15%. The ResNet50 model likewise has the top scores for sensitivity and specificity.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

State of disrepair roads have a negative impact on many aspects of daily life, from increasing traffic congestion to gradually decreasing property prices. Lack of maintenance roads are the primary cause of traffic accidents, which result in severe injuries to people's bodies and minds. Numerous emotional problems are brought on by it, including sadness, PTSD, anxiety, and behavioural changes. This even causes monetary issues. Using our Deep Learning-based technique, we can quickly provide a solution to this enormous issue. This strategy maintains a safe route while also using up time and money. We can help with minimal maintenance using our technology for detecting road degradation. A fast repair of damaged roads can prevent many fatal accidents and can improve current societal quality of life by estimating the cost of correcting the road damage and specifying the placement of the most necessary repairs.

5.2 Impact on Environment

Due to modifications in the ground flow conditions, construction projects and excavating can uncover soil and cause erosion. The habitat of wildlife may also be disturbed. The removal of surface vegetation may lead to a fall in plant species and affect the structure and operation of the ecosystem. In addition to poor traffic conditions, it contributes to a variety of imbalanced environmental factors, such as inclement weather, precipitation, and wind gusts. But implementing our suggestions could assist to enhance the environment. Like It can reduce the length of the construction process and help determine where the most important repairs should be made. makes for a more comfortable, safe, and smooth trip.

5.3 Ethical Aspects

We should be truthful when conducting research. For better study outcomes, some academics have created fictitious datasets or altered data. It is utterly unethical. Since in this instance, conducting such types of research for government road repair work may be against the law. We collected data for our investigation firsthand from Dhaka's streets. 1776 photos of road damage made up the dataset we used. To help our models identify more precisely, we built them using a tremendous amount of actual data.

5.4 Sustainability Plan

Every time we want to complete a task, we must think on how to keep it last. The goal of this work was to develop models using deep learning that could detect road deterioration. This kind of study could have a big impact on society. In light of this, we thought about our research's sustainability approach. The grounds of all damages now include many more categories. That is the reason we have such plan in place. We're about to incorporate supervised learning into our models. This will allow us to keep adding new models and train our model with different kinds of data. If the kind of road damage varies over this process, there won't be any problems.

In the near future, we'll also expand our dataset so which we can build a more resilient model to recognize and categorize road damage more precisely.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

This investigation set out to determine which transfer learning model performed better at identifying and categorizing road damage. We gathered data from the city of Dhaka's street. We gather a total of 1776 images of the road's surface and classify the image data into crack, pothole, rivaling and raveling. We made an effort to match the data using several transfer learning-based architectures. We train our model using 1335 images, test it using 176, and validate it using 265 data before using evaluation measures to assess its performance. We employ the F1 Score, Precision, and Recall measures for evaluation. ResNet50 performed better than other models in terms of evaluation, and results were also improved than any previous work. Deep learning-based solutions for road damage detection can assist in low-cost road maintenance. We construct a new dataset on primarily road damage images and test five methods to see which one works best for detecting road damage in the hopes that it will aid future research in this area. Further, the transportation system can be included in this job. To estimate the cost of repairing road damage, enlarge the parameters so that the location of the most urgent repair needs can be identified.

6.2 Limitation

It is important to keep in mind that there are a few restrictions. Huge amounts of data are required for deep learning predictions, but there is no shared dataset for work on road damage. Road damage data collection was exceedingly challenging.

6.3 Future Work

This road damage dataset can be useful for future studies. To estimate the cost of repairing road damage, enlarge the parameters so that the location of the most urgent repair needs can be identified.

In the hereafter, the projected dataset will be improved by utilizing cellphones and more images from various regions of Bangladesh. Collaboration with research institutes and municipal governments in Bangladesh is predicted to result in a substantially bigger dataset for a more precise estimator. Additionally, efforts will be made to develop and improve our model so that CCTVs deployed on roads can spot destruction in instantaneously and alert the appropriate authorities to the location and type of damage to help them make decisions. We build a new dataset, largely made up of images of road damage, and test five alternative approaches to see which one performs the best in order to aid future research in this area.

In the future, we intend to create a detection-capable Android application. road degradation and provide better classification results.

REFERENCE

[1] V. Pham, C. Pham, and T. Dang, "Road Damage Detection and Classification with Detectron2 and Faster R-CNN," 2020 IEEE International Conference on Big Data (Big Data), Dec. 2020, doi: 10.1109/bigdata50022.2020.9378027.

[2] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images," Computer-Aided Civil and Infrastructure Engineering, vol. 33, no. 12, pp. 1127–1141, Jun. 2018, doi: 10.1111/mice.12387.

[3] A. Alfarrarjeh, D. Trivedi, S. H. Kim, and C. Shahabi, "A Deep Learning Approach for Road Damage Detection from Smartphone Images," IEEE Xplore, Dec. 01, 2018. https://ieeexplore.ieee.org/abstract/document/8621899

[4] D. Arya et al., "Transfer Learning-based Road Damage Detection for Multiple Countries," arXiv:2008.13101 [cs], Aug. 2020, Accessed: Dec. 29, 2022. [Online]. Available: https://arxiv.org/abs/2008.13101

[5] V. Hegde, D. Trivedi, A. Alfarrarjeh, A. Deepak, S. Ho Kim, and C. Shahabi, "Yet Another Deep Learning Approach for Road Damage Detection using Ensemble Learning," IEEE Xplore, Dec. 01, 2020. https://ieeexplore.ieee.org/abstract/document/9377833 (accessed Dec. 29, 2022).

[6] K. Doshi and Y. Yilmaz, "Road Damage Detection using Deep Ensemble Learning," 2020 IEEE International Conference on Big Data (Big Data), Dec. 2020, doi: 10.1109/bigdata50022.2020.9377774.

[7] H. Tsuchiya, S. Fukui, Y. Iwahori, Y. Hayashi, W. Achariyaviriya, and B. Kijsirikul, "A Method of Data Augmentation for Classifying Road Damage Considering Influence on Classification Accuracy," Procedia Computer Science, vol. 159, pp. 1449–1458, Jan. 2019, doi: 10.1016/j.procs.2019.09.315.

[8] A. Angulo, J. A. Vega-Fernández, L. M. Aguilar-Lobo, S. Natraj, and G. Ochoa-Ruiz, "Road Damage Detection Acquisition System Based on Deep Neural Networks for Physical Asset Management," Advances in Soft Computing, pp. 3–14, 2019, doi: 10.1007/978-3-030-33749-0_1.

[9] T. Rateke, K. A. Justen, and A. Von Wangenheim, "Road Surface Classification with Images Captured From Low-cost Camera - Road Traversing Knowledge (RTK) Dataset," Revista de Informática Teórica e Aplicada, vol. 26, no. 3, pp. 50–64, Nov. 2019, doi: 10.22456/2175-2745.91522.

[10] Y. J. Wang, M. Ding, S. Kan, S. Zhang, and C. Lu, "Deep Proposal and Detection Networks for Road Damage Detection and Classification," IEEE Xplore, Dec. 01, 2018. https://ieeexplore.ieee.org/abstract/document/8622599 (accessed Dec. 29, 2022).

[11] V. Pereira, S. Tamura, S. Hayamizu, and H. Fukai, “A Deep Learning-Based Approach for Road Pothole Detection in Timor Leste,” 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), Jul. 2018, doi: 10.1109/soli.2018.8476795.

[12] “Convolutional neural networks based potholes detection using thermal imaging,” Journal of King Saud University - Computer and Information Sciences, Feb. 2019, doi: 10.1016/j.jksuci.2019.02.004.

[13] R. Fan and M. Liu, “Road Damage Detection Based on Unsupervised Disparity Map Segmentation,” IEEE Transactions on Intelligent Transportation Systems, pp. 1–6, 2019, doi: 10.1109/tits.2019.2947206.

[14] Y.-J. Cha, W. Choi, and O. Büyüköztürk, “Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks,” Computer-Aided Civil and Infrastructure Engineering, vol. 32, no. 5, pp. 361–378, Mar. 2017, doi: 10.1111/mice.12263.

[15] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, “Road crack detection using deep convolutional neural network,” IEEE Xplore, Sep. 01, 2016. https://ieeexplore.ieee.org/abstract/document/7533052

[16] J. Lin and Y. Liu, “Potholes Detection Based on SVM in the Pavement Distress Image,” IEEE Xplore, Aug. 01, 2010. https://ieeexplore.ieee.org/abstract/document/5571563 (accessed Dec. 29, 2022).

[17] N.-D. Hoang, “An Artificial Intelligence Method for Asphalt Pavement Pothole Detection Using Least Squares Support Vector Machine and Neural Network with Steerable Filter-Based Feature Extraction,” Advances in Civil Engineering, vol. 2018, p. e7419058, Jun. 2018, doi: 10.1155/2018/7419058.

[18] A. Akagic, E. Buza, and S. Omanovic, “Pothole detection: An efficient vision based method using RGB color space image segmentation,” 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2017, doi: 10.23919/MIPRO.2017.7973589.

[19] W. Xia, “An Approach for Extracting Road Pavement Disease from HD Camera Videos by Deep Convolutional Networks,” IEEE Xplore, Jul. 01, 2018. https://ieeexplore.ieee.org/abstract/document/8455601 (accessed Dec. 29, 2022).

[20] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, “A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure,” Advanced Engineering Informatics, vol. 29, no. 2, pp. 196–210, Apr. 2015, doi: 10.1016/j.aei.2015.01.008.

[21] S. Mathavan, K. Kamal, and M. Rahman, “A Review of Three-Dimensional Imaging Technologies for Pavement Distress Detection and Measurements,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2353–2362, Oct. 2015, doi: 10.1109/tits.2015.2428655.

Sharmin Shila

04/01/23

ORIGINALITY REPORT

18%	15%	8%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	9%
2	Submitted to Daffodil International University Student Paper	3%
3	Submitted to Liverpool John Moores University Student Paper	1%
4	YAN M. LOPES, ROBERTO G. MOORI. "THE ROLE OF IoT iN THE RELATIONSHIP BETWEEN STRATEGIC LOGISTICS MANAGEMENT AND OPERATIONAL PERFORMANCE", RAM. Revista de Administração Mackenzie, 2021 Publication	1%
5	Submitted to Segi University College Student Paper	<1%
6	paperswithcode.com Internet Source	<1%
7	link.springer.com Internet Source	<1%