AUTOMATIC DETECTION AND RECOGNITION OF FISH TO HELP VISUALLY IMPAIRED PEOPLE

BY

ASHIF RAIHAN
ID: 171-15-9389

AND

MD. ZAHED HOSSAIN MONJU
ID: 171-15-9395

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

Supervised By

Md. Taker Habib

Assistant Professor

Department of CSE

Daffodil International University

Co-Supervised By

Abu Khalid Mubashshir Mahmud

Lecturer

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JUNE 2021

APPROVAL

This project titled "**Automatic Detection and Recognition of Fish to Help Visually Impaired People**", submitted by **ASHIF RAIHAN** and **MD. ZAHID HOSSAIN MONJU** 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 (BSc) and approved as to its style and contents. The presentation has been held in June 2021.

BOARD OF EXAMINERS

Chairman

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

Abdus Sattar — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Jueal Mia — **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Dewan Md. Farid — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
United International University

DECLARATION

We hereby declare that this thesis has been done by us under the supervision of **Md. Trek Habib, 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 the award of any degree or diploma.

SUPERVISED BY:

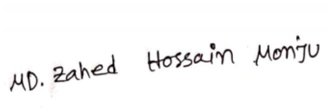
Md. Tarek Habib
Assistant Professor
Department of CSE
Daffodil International University

CO-SUPERVISED BY:

Abu Kahlid Mubashshir Mahmud
Lecturer
Department of CSE
Daffodil International University

SUBMITTED BY:

Name: Ashif Raihan
ID: 171-15-9389
Department of CSE
Daffodil International University



Name: Md. Zahed hossain Monju
ID: 171-15-9395
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to Almighty God for His divine blessing makes it possible to complete the final thesis successfully.

We are really grateful and wish our profound indebtedness to Md. Tarek Habib, Assistant Professor, Department of CSE Daffodil International University, Dhaka, Bangladesh. Deep Knowledge & keen interest of our supervisor in the field of "Deep Learning and Image processing" to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting us at all stages have made it possible to complete this thesis.

We would like to express our heartiest gratitude to o Prof. 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 discussion while completing the course work.

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

ABSTRACT

This paper explores fish identification for blind people utilizing multi-image recognition including deep learning techniques. Our system was built using Speech Recognition to support blind people in knowing about various fish in the market or somewhere else. According to our preliminary study, the majority of blind people in Bangladesh face difficulties purchasing various products, especially fish, from the market. However, there is no automated device available that can identify the fish and offer an interpretation. For this generosity of spirit, we created a system for visual imperial people that will automatically identify and explain the fish using speech recognition. We employed Deep Learning (DL) resources like Matploatlib, TensorFlow, Keras, and others throughout the process. We utilized TensorFlow for image preprocessing and classification. We used the most popular image processing algorithm, Convolutional Neural Network (CNN), to evaluate the reliability of our work. We tested three CNN models to see which one offered the best performance. We compared all of the models before identifying multiple fishes with the model that provided the highest precision and efficiency.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

People are ecstatic to be able to purchase necessities using their own hands and preferences. People who are blind or visually impaired are no different. According to the World Health Organization (WHO) in 2010[1, there are 285 million people worldwide who are visually impaired of all ages, with 39 million of them blind, which is a big amount. There are several visually disabled people in Bangladesh. According to the initial survey, most blind people are pleased to purchase their glossaries, but there is no automated system that enables them to recognize the object and have it interpreted for them using speech recognition. We felt bad for them knowing these things, so we decided to assist them in exploring the market in a systematic manner, especially in identifying the fish. With the advancement of a camera-enabled device's computational capability, object recognition on such devices is shifting away from oriented client-server methods, in which phones or devices with cameras only function as input/output front-ends, and toward local on-device classification methods. To fix this problem, we built a model for visual imperial people that automatically recognizes the object/fish and explains it using a speech recognition technique. ML is a type of artificial intelligence (AI) that allows a system to understand and use knowledge to improve implementation without having to be explicitly programmed. We identify the image using the Deep Learning (DL) method on the image dataset to build our model. We've used several deep learning tool libraries, including Matplotlib, Tensorflow, Keras, and others. The image was preprocessed using the TensorFlow Keras API. To evaluate the validity of our findings, we used three CNN models focused on deep learning. We examined all three models to see which one had the best accuracy. For our model to recognize and classify the picture, we first cover twelve groups. We used the best accuracy generated model for image classification to detect the fish perfectly. Our model would also substantiate to be beneficial for regular peoples. It can accommodate any kind of user with the information of a particular fish. To detect the fish perfectly, we practiced the most accurate image classification model we could discover. Our model will also determine that it is helpful to ordinary citizens. It can provide any kind of consumer with knowledge about a specific fish.

1.2 Motivation

In our wonderful planet, the habitat for all types of people. Some are born with the gift of having all of their characteristics or attributes, but it is fate that not everyone is equal. Some of them cannot comprehend or some of them cannot see. They have a variety of deformations. There aren't nearly enough services to help blind people. When a blind person wants to go to the market, he or she is unable to do so. We decided to use Deep Learning, Image Processing to develop an application that handles images, transforms photograph into texts, and converts texts into voice-form. This app is a tremendous help to the blind.

1.3 Problem Definition

Sightless people can't see anything, they don't know much what fish they buy from the market and don't know the fish as well. This can be tackled by image classification by recognizing the image while they visiting the market. By using image recognition, we can detect and reorganize fish or artifacts in front of the camera, blind people can hear details or things, and blind people can understand what fish they buy. This system can be beneficial for the identifying the local and coastal fish for normal people also.

1.4 Research Questions

- Can the program detect the fish precisely?
- How detecting image can be helpful to blind people.
- Can the detection accurately define the fish?

1.5 Research Methodology

This section will go over the Experiment data, data pre-processing, algorithm implementation, training the model, and model evaluation. The model's success will be addressed at the end of the chapter.

1.6 Research Objectives

Our main goal is to create a deep learning-based system that will assist visually disabled people in visiting fish markets and identifying fish using the system.

1.7 Research Layout

The report will have appeared as regards:

The overview of this study is presented in Chapter 1. This first section is the main stage of the introductory analysis. Besides, it is also well illustrated in this chapter what inspired us to make such study. The problem definition is the most important part of this chapter. Then, what are the research questions and what are the anticipated results in the last chapter.

Chapter 2 discusses what has previously been done in this area. Various forms of work involving machine learning, deep learning, and some image processing.

The research methodology is discussed in Chapter 3; this chapter elaborates on the statistical methods used in this work. In addition, this chapter demonstrates the procedural methods of the CNN models. There are raw data, pre-processed data, data analysis, implementation, and all the steps required to complete the work.

The assessment of the outcome is linked to Chapter 4. This implies that the results or results of the research are reported. The details of the proposed work are presented in this chapter.

The conclusion subjects of the study are founded on Chapter 5. The entire report on results is presented in this chapter in accordance with the recommendation. The chapter ends with a display of weaknesses in our work which could be in the future for other people who wish to work in this area.

1.8 Expected Outcome

The expected results of this research are to consider or cumulate the resulting strategy, which determines a given fish picture as regards the welded model of a prepared data set. In order to detect fish in which region the predicted result will be divided into different categories.

CHAPTER 2

BACKGROUND

2.1 Introduction

Machine learning and Deep Learning are widely used to solve image processing problems. The processing of images plays a crucial role to help people make life easier. A lot has been done to make this in our everyday lives more accessible.

2.2 Related Works

There are some of the previous works that related to our works is given below:

The system for identifying the artist, kind, stuff, and creation year of a work of art was proposed by T. Mensink et. al. [2]. The exhibit of the Rijksmuseum exhibited in Amsterdam, Netherlands, used a dataset of 112,039 images, freely available. The system precision was computed using the precision of the frequency classified. The mean average precision (MAP) algorithm was used for measuring the kinds of categorization of artworks. Exactitude was also measured in MAP for the labeling of the minerals. They encrypted the images using state of-the art Fisher vectors (FV). For classification and regression models, 70 percent of the data were used. 10% of the data set was used as a validation set to adjust the hyper-parameters of the models. Test set prepared by 20% data. The tests were carried out using liblineara SVM.

O. Bimber et. al. [3] explained their mobile museum guidance system Phone Guide. The procedure was dealt with by using an adaptive image classification to achieve virtually feasible object recognition rates under realistic terms. To recognize objects, a 3-laying neural network was used to extract global color characteristics from the captured images. Following the classification with a nearest neighbor technique of separate key frames, data base images are analyzed to determine if they are related to nearly far between frames similar to those of the key frames adopted. The user's location was judged to enhance the mobile image classification system by 93.47 percent accuracy.

J. S. Hare et. al. [4] a new approach was adopted to run the retrieval of content-based images and the recognition of objects from input images which were degraded by noise and subjected to image system transformations. They had used model of vector space to index each image
effectively. After this, the correct image is defined by the two-step ranking process. The algorithm is highly resistant to modifications between images of the query and database. In the implementation of this system is used Lowe's Invariant Transform Feature Scale (SIFT) descriptor13. The images of the query are rebuilt in a vector space based on 'visual' words in the image. Its dataset included 850 images from the photograph collections of the National Gallery. They have randomly selected 200 photos to test the system.

T.-Y. Lin et. at. [5] By exerting dominance of the inherent multi-scale, pyramidal hierarchy of deep convolutional networks the authors-built feature pyramids with minimal additional cost. A high-level semantic feature map at all measures was formed utilizing a top-down architecture with lateral attachments. The architecture Feature Pyramid Network (FPN) bestowed much development as a generic feature extractor. They obtained state-of-the-art single model results on the COCO apprehension benchmark externally bells and whistles. The system relied on a structure that blends low-resolution, semantically weak features through a top-down pathway, semantically strong features with high resolution, including lateral connections.

P. Dollár et. al. [6] have transformed three distinct visual recognition methods for applying fast feature pyramids. They examined the consequences of object detection and pedestrian detection. They employed the Caltech, TUD-Brussels, ETH datasets, INRIA, also PASCAL VOC. Their method is broadly relevant for vision algorithms requiring fine-grained multiscale study. Their assessment is possible in large-scale images, but it fails in small spectrum bandpass pictures. They demonstrated competence in pedestrian detection of Aggregated Channel Features (ACF). The ACF uses standardized gradient magnitude, histogram, and color channels of the oriented gradients.

C. Szegedy et. al. [7] attempted to organize and specifically localize objects by employing the power of the DNNs. They used DNN-based regression to create a system that outputs a binary mask of the object-bounding box. They imply that a DNN-based regression will learn features for classification as well as geometric details. After recovering masks of various objects in a range using DNN regression, the system utilizes a DNN localizer to a small set of large sub-windows. It employs one network to mask the object box and four additional networks to foretell the box's four halves. The layers were taught using ADAGRAD using a stochastic gradient to estimate the learning rate of the layers naturally. The system had reviewed on the Pascal Visual Object Challenge (VOC) 2007 dataset, which included 5000 images from 20 different classes. The system was trained using the VOC2012 training and validation dataset, which included 11000 images.

X. F. Hermida et. al. [8] suggested a system for identifying the Braille Code using the ideas of Optical Character Recognizer (O.C.R.). They practiced a variety of image processing tools, including adaptive thresholding and skew angle detection. The zones were converted into black and white spots using an approach based on two adaptive thresholds. The luminance histogram was used to calculate the thresholds. The system was also capable of detecting fake and lost points.

G. Sainarayanan et al. [9] aimed to create a system that could assist blind people by identifying barriers in their path of travel. The system consists of a single board processing system (SBPS), a vision sensor mounted headgear and a pair of stereo earphones. The vision sensor gathers information about the surroundings in front of the user. The image is then processed in real-time using fuzzy clustering algorithms in accordance with a real-time image processing scheme. The processed image is then transformed into a specially designed stereo acoustic pattern and sent to the stereo earphones. A fuzzy-based LVQ (FLVQ) is used to identify objects in the background.

Based on the topic above, we should distinguish our work based on the function of a speech recognition device that can tell about different fish. As a result, blind people can recognize and know what fish they purchase.

2.3 Research Summary

The above study was conducted on a wide range of research works from various research teams, and it is being produced to demonstrate that new research work on image is progressing day by day. This argument has already been validated by a number of positive outcomes. While adequate resources are not available, there is hope that this region will become more resourceful with each passing day.

2.4 Challenges

The main difficulties in this work are concerned with the datasets. We need some effective entrances to perform the dataset refinement, but there aren't enough known approaches to do it. Another difficulty of this work is a lack of funding for this subject and the collection of accurate data from markets and fisheries.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The methodology includes a total of six steps which accomplishes our research that is presented in Figure 3.1. The steps are the following:

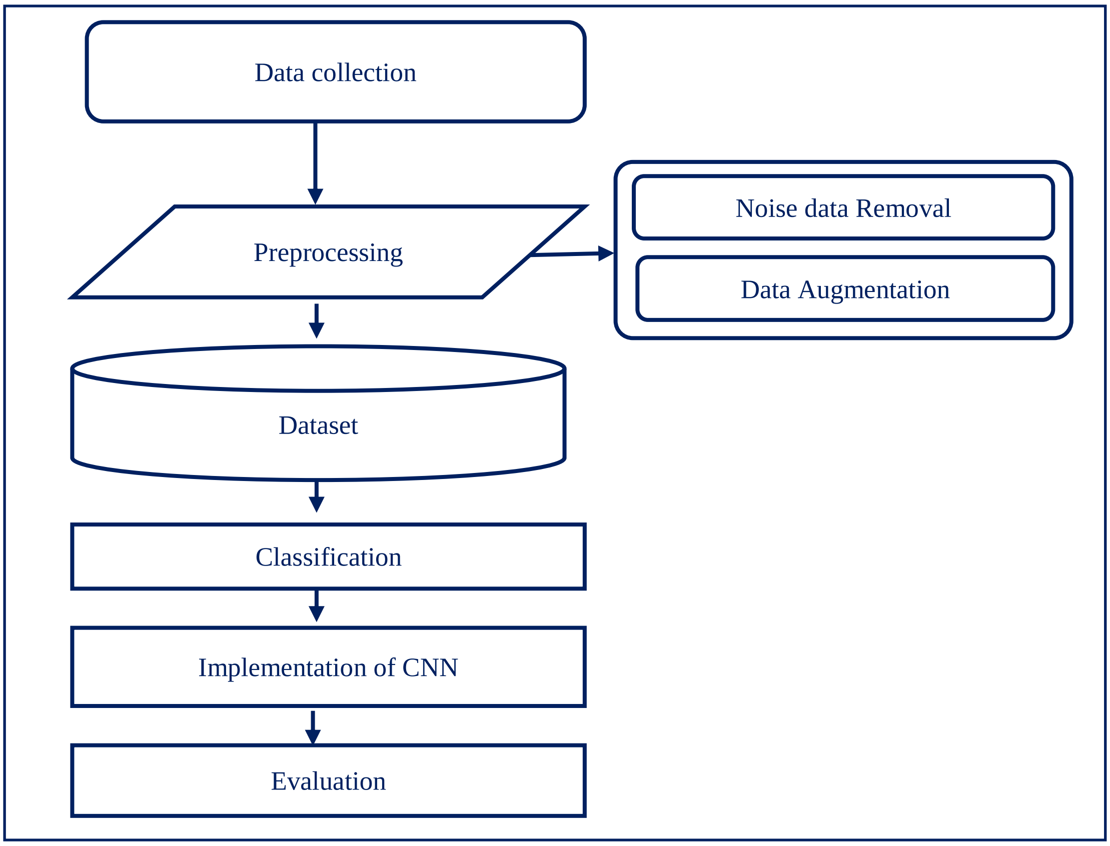


Figure 3.1: Methodology diagram

3.2 Data Collection

We went to a lot of local fisheries and markets to get real raw fish photos for our work. We gathered twelve different types of images that depict twelve different fishes (local and coastal fish). We accumulated 1200 images in total. We photographed the fishes with a Canon EOS 700D digital single-lens reflex (DSLR) sensor. We get pictures with high resolution and low noise.

3.3 Pre-Processing

The method of transforming raw data into a pixel array format that can be employed to train various algorithms is recognized as image data preprocessing. We practiced Tensorflow Keras to construct our model. It's a dataflow and differentiable programming library that's open-source and available. The noise data was removed for building a reliable dataset. For this project, we augmented our raw 1200 data with about 3000. Our gathered data was split into two parts. Our model was trained by the first Part of the dataset that has 3077 images, and we tested it in the second that has 345 images. The photographs are all colorful and distinct.

3.4 Dataset

All the images are converted into 224*224 height and weight. Our dataset contains a total of around three thousand and five hundred images. An example image in our dataset is given in Figure 3.2. All of our image data were regenerated in the similar way that is bestowed below:

Figure 3.2: Sample of dataset.

3.5 Classification

We gathered data of twelve different kinds. As a result, our entire dataset was divided into 12 groups. We used the name “ Koral, Kholse,Hilsa, Boumach, Baila, Tailla, Rupchada, Puti, Poa, Pabda, Mini, Loitta “ to represent the 12 separate types of fish. The total number of images in every class is about 250.

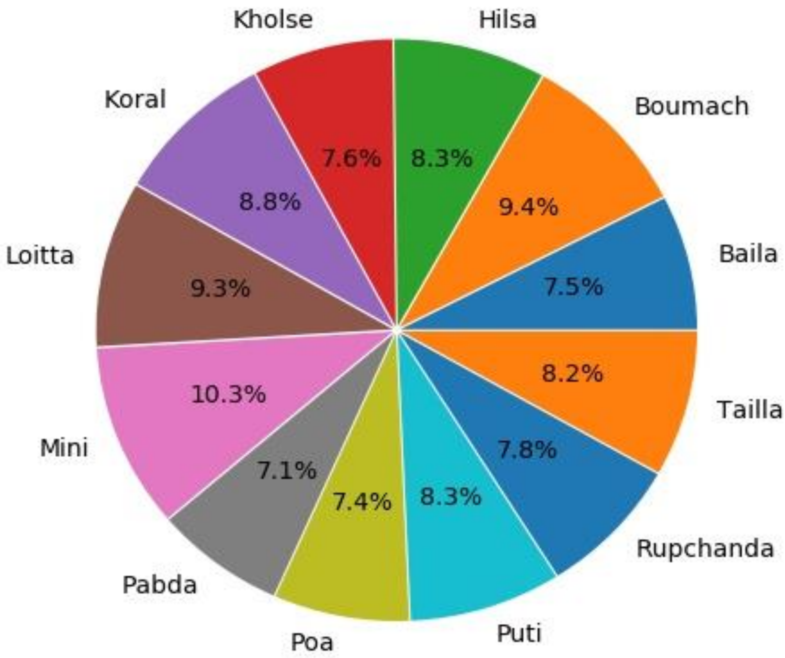


Figure 3.3: Classification of dataset.

Figure 3.3 depicts the data distribution for each group.

3.6 Implementation of CNN

Our model is based on image classification. Holding this in mind, we decided that CNN is the best alternative for developing our model. Model1, Model2, and Model3 were the names we assigned to our deep learning CNN models. We evaluated the efficiency of all three models to decide the appropriate one for training and testing so that our model could accurately classify the fish. All of the models produced satisfactory results.

3.7 Evaluation

We evaluated our model in the following manner:

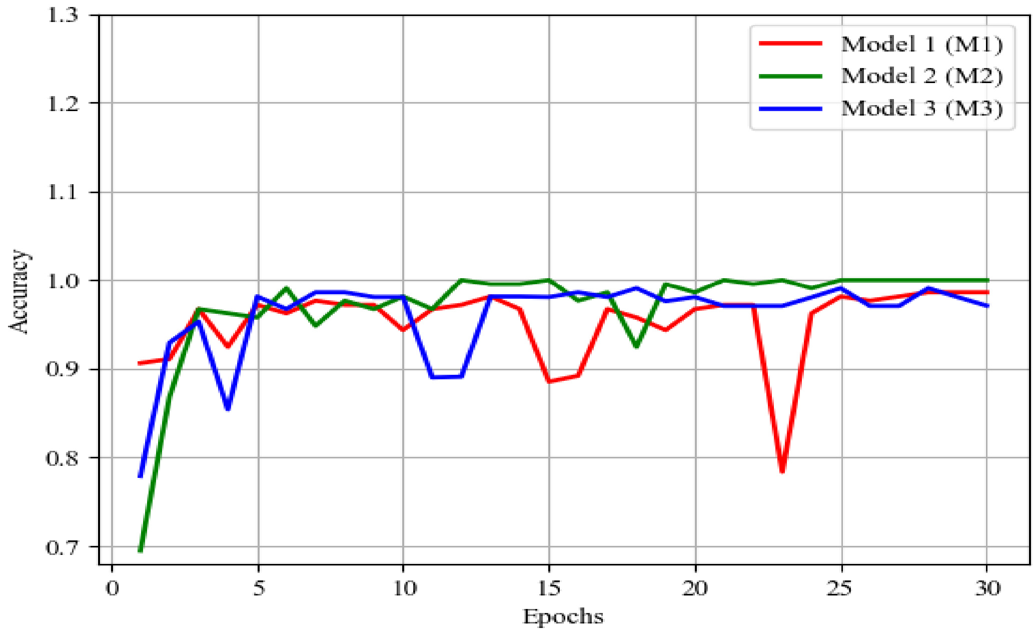


Figure 3.4 Accuracy comparison of three models

From Figure 3.4 we assume that the achievement of our three models is much anticipated and less distinct. In this figure the red, green and blue line depicts Model 1, Model 2 and Model 3 correspondingly. Their performance is extremely high that Model 2 got 100% accuracy and the other models are also performed very admirably.

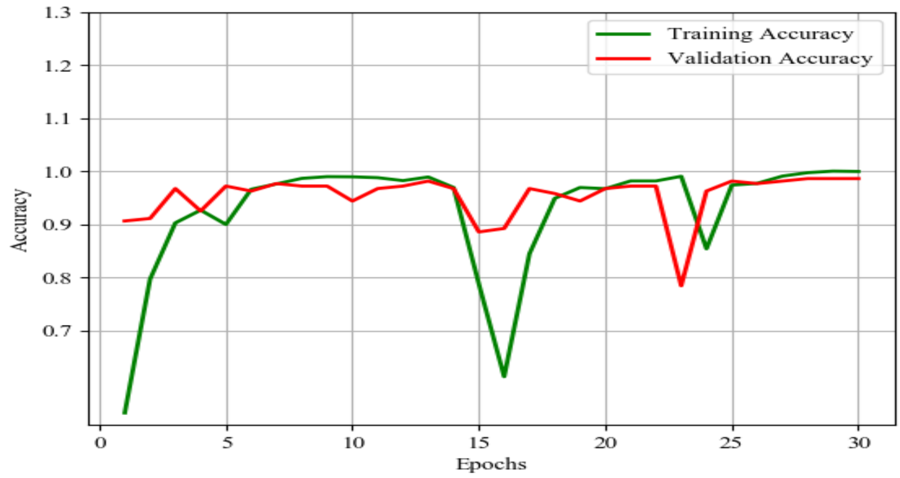


Figure 3.5 Training vs validation accuracy of Model 1

Figure 3.5 depicts CNN Model M1's training & validation accuracy. It creates a slight distance between validation and training at certain stages, but most of the time the difference between training and validation accuracy was minimal.

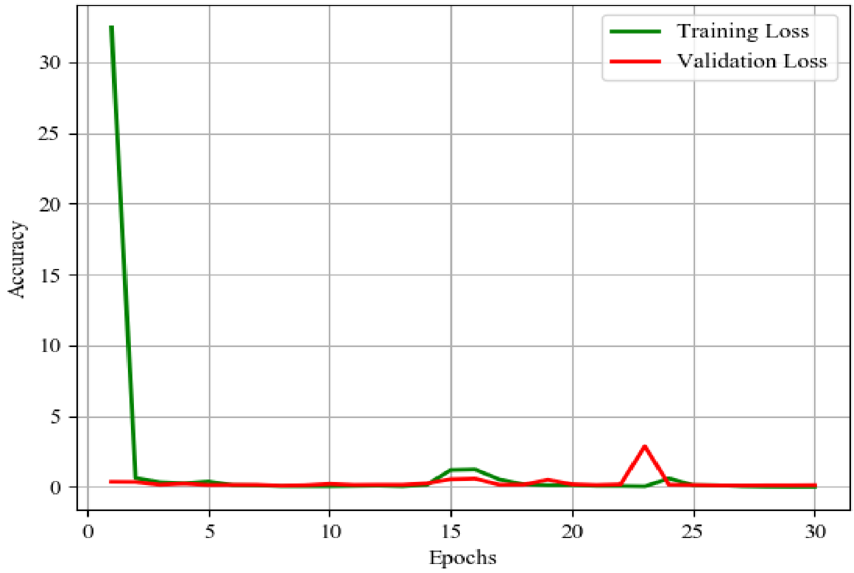


Figure 3.6 Training vs validation loss of Model 1.

Model M1's training and validation losses are depicted in Figure 3. 6. From this graph, we can see that at certain stages, validation loss is greater than training loss, but the difference is very small.

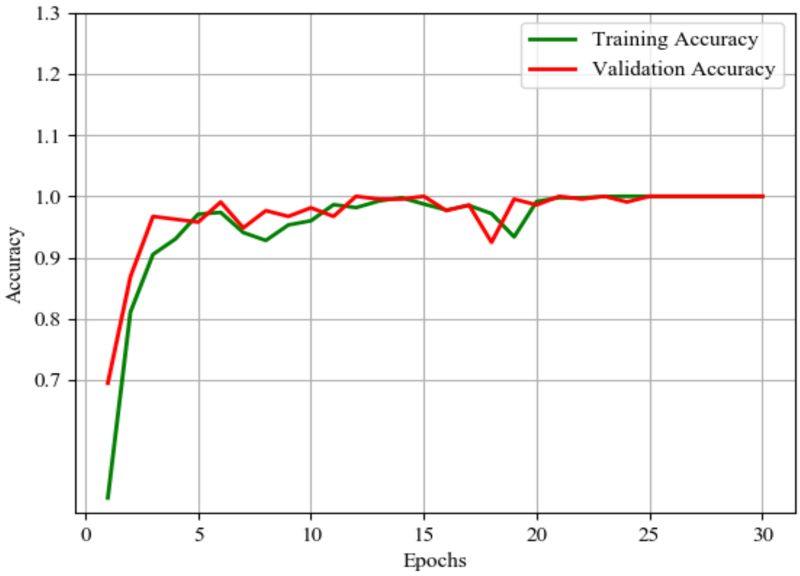


Figure 3.7 Training vs validation accuracy of Model 2.

The training and validation accuracy of CNN Model M2 is depicted in Figure 3.7. It generates a small gap between validation and training at a particular epoch stage, but most of the time the gap between training and validation accuracy was marginal or overlapped.

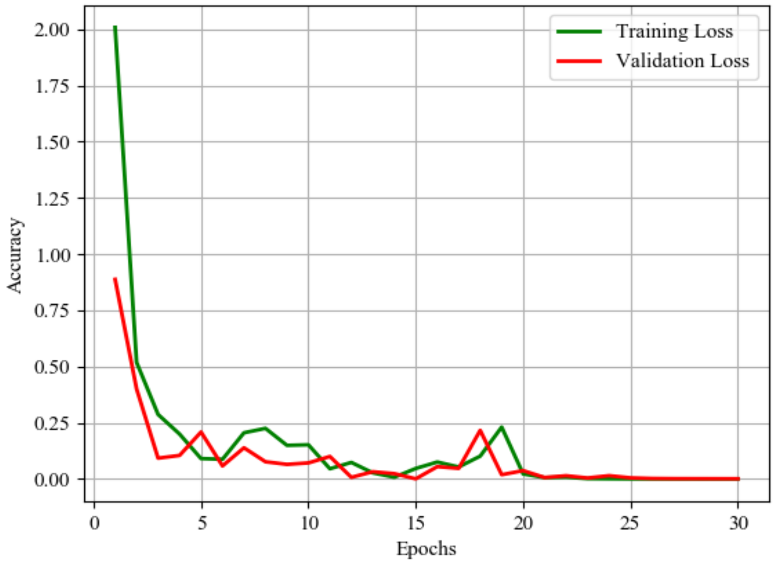


Figure 3.8 Training vs validation loss of Model 2.

Model M2's training and validation losses are depicted in Figure 3.8. From this graph, M2 did pretty well and we can observe that at certain stages, validation loss and training loss create a little difference, but the variance is very small.

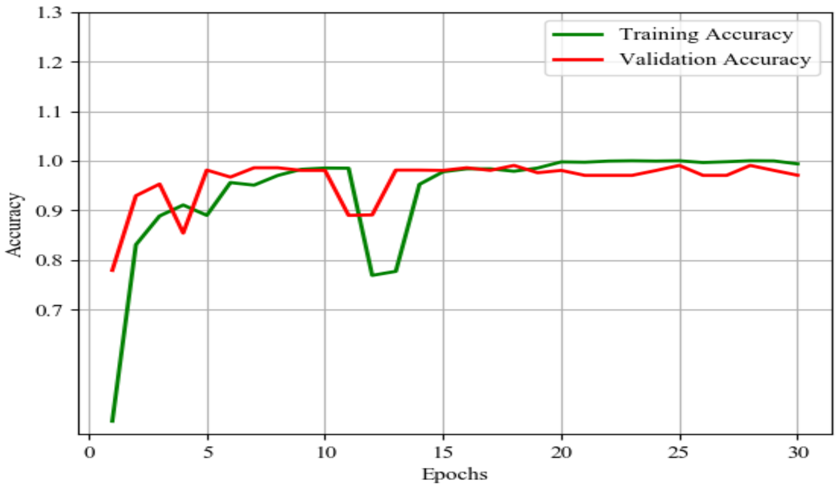


Figure 3.9 Training vs validation accuracy of Model 3

Figure 3.9 depicts CNN Model M2's training & validation accuracy. It creates some gap between validation and training at some epoch points, but most of the time the distinction between training and validation accuracy was insignificant.

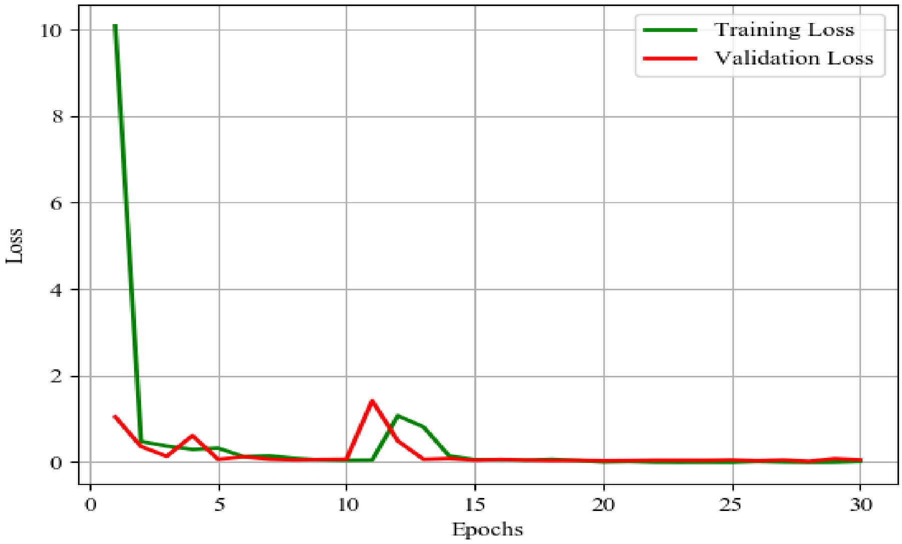


Figure 3.10 Training vs validation loss of Model 3.

Model M2's training and validation losses are depicted in Figure 3.10. From this graph, we can observe that at certain stages, validation loss and training loss create a little difference, but the diversity is very tiny.

As we can see that Model M2 is good for this dataset and performed best among all three models. So, we decided to use model M2 for further process.

CHAPTER 4

RESULT ANALYSIS

4.1 Introduction

This Chapter 4 primarily emphases on the evocative analysis of the data used in the research as well as the experimental results of our research.

4.2 Experimental Result

We tested three CNN models on the gathered dataset to see how well they performed. We inserted the data into the CNN Models Table in order to achieve a crystal representation of precision for future evaluation of this experiment. Our models were trained on 3077 images and tested on 345 images. A very impressive finding was found by corresponding three models. All of the algorithms performed in a similar manner. The three CNN model explanations that we generated are described in Tables 4.1, 4.2 and 4.3.

Table 4.1 CNN MODEL M1

Model Layout				Training Accuracy(%)	Validation Accuracy(%)
No	Layers	Filters	Kernel size		
M1	Convolutional	64	(3,3)	98.59	98.14
	MaxPooling	-	(2,2)		
	Convolutional	128	(3,3)		
	MaxPooling	-	(2,2)		
	Convolutional	256	(3,3)		
	MaxPooling	-	(2,2)		
	Dense	1024 units	Acti..= relu		
	Dropout	0.2	-		
	Dense	512 units	Acti.. =relu		
	Dropout	0.2	-		
	Dense	12	Acti..= softmax		

Model 1 has three convolutional layers with 64, 128, and 256 filters, as seen in Table 4.1. M1 achieved a training accuracy of 98.59 percent and a validation accuracy of 98.14 percent.

Table 4.2 CNN MODEL M2

M2	Convolutional	64	(3,3)	100	100
	MaxPooling	-	(2,2)		
	Convolutional	128	(3,3)		
	MaxPooling	-	(2,2)		
	Convolutional	256	(3,3)		
	MaxPooling	-	(2,2)		
	Convolutional	512	(3,3)		
	MaxPooling	-	(2,2)		
	Dense	1024 units	Acti..= relu		
	Dropout	0.2	-		
	Dense	512 units	Acti.. =relu		
	Dropout	0.2	-		
	Dense	12	Acti..= softmax		

We employed four convolutional layers with 64, 128, 256, and 512 filters in the corresponding method for Model 2 in Table 4.2, and we obtained 100 percent training and validation accuracy, which is exceptional.

Table 4.3 CNN MODEL M3

M3	Convolutional	64	(3,3)	99.96	98.06
	Convolutional	64	(3,3)		
	MaxPool	-	(2,2)		
	Convolutional	128	(3,3)		
	Convolutional	128	(3,3)		
	MaxPool	-	(2,2)		
	Convolutional	256	(3,3)		
	MaxPlooling	-	(2,3)		
	Convolutional	512	(3,3)		
	MaxPlooling	-	(2,3)		
	Dense	1024 units	Acti..= relu		
	Dropout	0.2	-		
	Dense	512 units	Acti.. =relu		
	Dropout	0.2	-		
	Dense	12	Acti..= softmax		

Model 3 is depicted in Table 4.3. We utilized 6 convolutional layers including 64, 64, 128, 128, 256, and 512 layers for model M3, and we got 99.97% and 98.06 percent training and validation accuracy for this model.

We can observe from the aforementioned analysis that CNN Model M2 achieved the highest accuracy of 100% among all the models. For the following procedures, we picked Model M2.

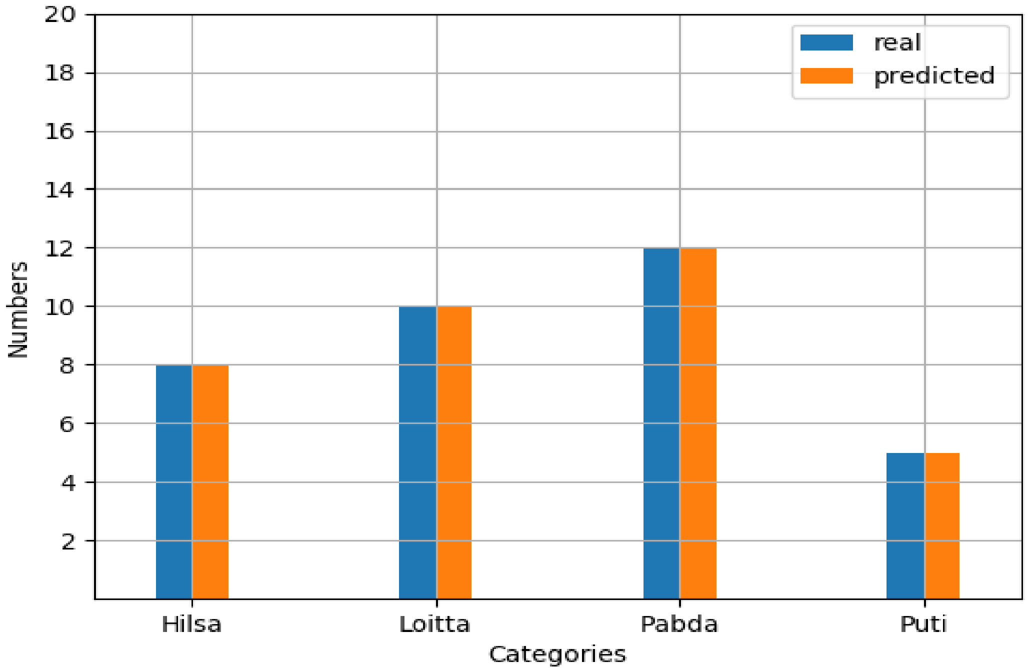


Figure 4.1 Real Vs Predicted Comparison

During this process, we checked our model with images of fish that were selected at random. We picked 35 images at random to compare against the CNN algorithm's prediction. The outcome of our assessment is depicted in Fig. 11. Four classes' images were included in the test data. The predicted and actual classes were approximately similar. The blue bar depicts the real image class, while the orange bar depicts the predicted image class. It's really clear to see from the graph that there was 100% efficiency in our model.

From the above discussion, we can accomplish that the Model used in this work performed fairly well. This interprets our work implicitly.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

In our society, there are several people around us. Some are born with the benefit of holding all of their characteristics or qualities, but it is the fortune that not everyone is alike. Some of them cannot comprehend or some of them cannot see things. They have a mixture of deformations. There isn't approximately adequate assistance to help blind people. Our system can help them to understand things by image processing and speech reorganization. They can buy their necessary things by their hand with the help of this system.

5.2 Impact on Environment

When we say the environment at first, we think of all the things around us. The people are the main part of the environment. Blind people generally think of them as a burden because they normally can do their thing by their hand. So, our system at least helps them to recognize something.

5.3 Ethical Aspects

- With the help of this system, blind people and normal people initially know the details of the recognized fish.
- Blind people can buy their necessary things by their hands and can visit the market.
- The system has a speech reorganization system so that everyone can understand things.
- Can differentiate the local and coastal fish.

5.4 Sustainability Plan

This system will help the Blind and normal people as well now and future also. In near future, the system will increase the features and include many objects for recognizing to help the people.

CHAPTER 6

SUMMARY, CONCLUSION AND FUTURE WORK

6.1 Summary of the Study

It has no doubt that there are numerous research activities on image processing. When the outcome of such a wide range of works causes a cataclysmic shift in our computing lives, this type of research has recently extended. We get some notable real-world applications from such forms of research works. However, it is extremely concerning that there is no such research work for blind people. However, we anticipate that several researchers from various countries have begun researching this area. In our study, we develop methods for detecting fish, categorizing them, and explaining them to both normal and blind people.

6.2 Conclusion

Our device can detect fish with 100% accuracy using data collected from various markets and fisheries in Bangladesh. Model 2 out of three CNN models generated the best results. The output of all the models was quite close. The blind would be able to classify fish using our proposed model. Normal people may benefit from our model as well. It is capable of providing information about a specific fish to any person who has this system.

6.3 Recommendations

A few notable recommendations for this are as follows

- To create the data set more efficiently and organized, can produce an enhanced output of this research work.
- Growing the data in the dataset can also make a healthier output.

6.4 Future Work

On the basis of our proposed model, we have already begun to create a mobile application. To make our dataset more useful to a wider range of users, we intend to expand our prediction classes and dataset size. We would attempt to cover all of the market's glossaries in order to provide more valuable material for both blind and non-blind users.

REFERENCE

[1] WHO. (2010). Blindness and vision impairment prevention. Available: https://www.who.int/blindness/publications/globaldata/en/T. Ghorpade and L. Ragha, "Featured based sentiment classification for hotel reviews using NLP and Bayesian classification," in 2012 International Conference on Communication, Information & Computing Technology (ICCICT), 2012, pp. 1-5: IEEE.

[2] T. Mensink and J. Van Gemert, "The rijksmuseum challenge: Museum-centered visual recognition," in Proceedings of International Conference on Multimedia Retrieval, 2014, pp. 451-454.

[3] O. Bimber and E. Bruns, "PhoneGuide: Adaptive image classification for mobile museum guidance," in 2011 International Symposium on Ubiquitous Virtual Reality, 2011, pp. 1-4: IEEE.

[4] J. S. Hare and P. H. Lewis, "Content-based image retrieval using a mobile device as a novel interface," in Storage and Retrieval Methods and Applications for Multimedia 2005, 2005, vol. 5682, pp. 64-75: International Society for Optics and Photonics.

[5] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2117-2125.

[6] P. Dollár, R. Appel, S. Belongie, P. J. I. t. o. p. a. Perona, and m. intelligence, "Fast feature pyramids for object detection," vol. 36, no. 8, pp. 1532-1545, 2014.

[7] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in Advances in neural information processing systems, 2013, pp. 2553-2561.

[8] X. F. Hermida, A. C. Rodriguez, and F. M. J. P. o. I.-B. Rodriguez, "A Braille OCR for Blind People," 1996.

[9] G. Sainarayanan, R. Nagarajan, and S. J. A. S. C. Yaacob, "Fuzzy image processing scheme for autonomous navigation of human blind," vol. 7, no. 1, pp. 257-264, 2007.

[10] Md. Tarek Habib, Md. Jueal Mia, Mohammad Shorif Uddin, Farruk Ahmed "An in-depth exploration of automated jackfruit disease recognition". In Journal of King Saud University – Computer and Information Sciences, 25 April 2020

[11] M. M. Hasan, M. T. Zahara, M. M. Sykot, A. U. Nur, M. Saifuzzaman and R. Hafiz, "Ascertaining the Fluctuation of Rice Price in Bangladesh Using Machine Learning Approach," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-5, doi: 10.1109/ICCCNT49239.2020.9225468.

APPENDIX

We faced numerous challenges while conducting the analysis, the first of which was determining the methodological methodology for our study. It was not conventional work, and there had not previously been much work done in this region. As a result, we couldn't get much support from anybody. Another impediment was data processing, which proved to be a significant challenge for us. Since there was no available source for the dataset and image pre-processing method, we created a corpus for data collection. In addition, we began manually collecting data. We were able to accomplish this after a long period of hard work.

PLAGIARISM REPORT

Plagiarism Report

ORIGINALITY REPORT

8%	3%	7%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Md. Rejaul Alam, Afsana Akter, Minhajul Abedin Shafin, Md. Mehedi Hasan, Antara Mahmud. "Social Media Content Categorization Using Supervised Based Machine Learning Methods and Natural Language Processing in Bangla Language", 2020 11th International Conference on Electrical and Computer Engineering (ICECE), 2020 Publication	1%
2	www.arxiv-vanity.com Internet Source	1%
3	Samira Pouyanfar, Muneeb Saleem, Nikhil George, Shu-Ching Chen. "ROADS: Randomization for Obstacle Avoidance and Driving in Simulation", 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019 Publication	1%
4	Hao Sun, Zehui Meng, Pey Yuen Tao, Marcelo H. Ang. "Scene Recognition and Object	1%