LOCAL FISH RECOGNITION: A MACHINE VISION APPROACH

BY

JAHANGIR ALAM
ID: 163-15-8276
AND
SUBROTA DAS SHUVA
ID: 163-15-8277
AND
PRIANKA SAHA
ID: 163-15-8270

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

Supervised By

ANUP MAJUMDER
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

SHAON BHATTA SHUVO
Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

SEPTEMBER-2019

APPROVAL

This Project/internship titled **"Local Fish Recognition: A Machine Vision Approach"** submitted by **Jahangir Alam, ID: 163-15-8276, Subrota Das Shuva, ID: 163-15-8277, Prianka Saha, ID: 163-15-8270** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **12th September**.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Narayan Ranjan Chakraborty — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Shaon Bhatta Shuvo — **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Saddam Hossain — **External Examiner**
Assistant Professor
Department of Computer Science and Engineering
United International University

DECLARATION

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

Supervised by:

ANUP MAJUMDER
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

SHAON BHATTA SHUVO
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Jahangir Alam
ID: 163-15-8276
Department of CSE
Daffodil International University

Subrota Das Shuva
ID: 163-15-8277
Department of CSE
Daffodil International University

Prianka Saha
ID: 163-15-8270
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

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

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

We would like to express our heartiest gratitude to **Professor Dr. Syed Akhter Hossain**, Head**,** Department of CSE, and **Md. Tarek Habib,** Assistant Professor**,** Department of CSE, Daffodil International University, Dhaka and also to other faculty member 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 patients of our parents.

ABSTRACT

Our local fish are becoming extinct day by day. Known for the extinction of local fish in Bangladesh and the people of our country, we will detect local fish.in this paper a methodology has been detect to local fish. This paper will be identified by using color, shape, texture processing. In this research work we have achieved a huge accuracy of 88% using Random Forest Classifier. In this paper an automated approach is developed to detect local fish and recognition by using machine vision based image processing techniques which is implemented in “PYTHON” programming language on “ANACONDA” platform including a machine learning algorithm with the Random Forest Classifier.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES **PAGE NO**

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Introduction

Bangladesh is a country with thousands of rivers and lakes and is known for a fish-loving nation with the name "Machh-e Bhat-e Bangali" meaning "Fish and rice Bengali." Ilish is the country's domestic fish, contributing 13 percent of total fish manufacturing in the nation .Our country is blessed with a great variety of Sweetwater fishes. There are so many features available provided by the modern technologies that can make a great change in our fishing sector. If we say about developed countries they are practicing many advanced technologies in their fishing where in a fish-loving nation like Bangladesh has not given much attention to the modern technologies that can contribute an enormous amount of productivity in fishing. There are noble and prominent research possibility and opportunity are available in our fishing science that are not being risen properly. In this research an automated system is developed which can certainly help to detect local fish by capturing the image with a camera and then they should input the image to the processing system. The system analyzes the image is segmented base on color shape texture processing and then classified by Random Forest Classifier to detect the predicted local fish.

1.2 Motivation

Fish play a central role in livelihood and culture in Bangladesh. Some 76 species of fishes are often included both as freshwater and marine in Bangladesh. In Bangladesh, local fish is about to go extinct because of habitat change due to the change of River flow. Fish is a daily meal of Bengali. Without fish we cannot think, but we do not know the exact name of fish.

Our youth society does not know the name of those fish. Because Day by day our local fish are extinct. About 54 species of fish are extinct in our country.

There is no research on this in our country. On the way to the extinction of local fish in Bangladesh, we will do this research to introduce the youth to the local fish.

In Bangladesh, local fish is about to go extinct because of habitat change due to the change of River flow. Fish is a daily meal of Bengali. Without fish we cannot think, but we do not know the exact name of fish.

Our youth society does not know the name of those fish. Because Day by day our local fish are extinct. About 54 species of fish are extinct in our country. There is no research on this in our country. On the way to the extinction of local fish in Bangladesh, we will do this research to introduce the youth to the local fish.

To build the system many images are collected to train the system in PYTHON using color, shape, texture processing and Random Forest Classifier.

1.3 Objectives

The objectives of this research to find out and analyze a technique which can help various to detect the local fish from the farms or market and thus make a suitable profits. The proposed system is user friendly, fast to execute and can be implemented easily in any system.

- ✓ To Indicative the local fishes to our country man
- ✓ To Making extinct fish known to the people of our country
- ✓ To contribute to the research field in our country
- ✓ To retreat to extinct fish
- ✓ To use Random Forest Classifier scenario.

1.4 Expected Outcome

The fish of Bangladesh is a favorite food of the people. But fish like local fish such as rui, katla, chital etc. are going extinct. And with these we are not well-known to the current

Youth society. But we should know about these fish. And in this endeavor we did this research. We thought of a procedure which will be available to youth society to identify the local fish by capturing images .The problem will be identified by using color, shape, texture processing and Random Forest Classifier.

- ✓ We will get a research based project (There is no work on local fish detection in our country until now. We believe that lot of information will come out in our research)
- ✓ We will also getting a Apps (User simply take a picture with their phones camera and in seconds the program identifies their fish)
- ✓ Can know about local fish (This system not only detect fish, but also provides information about fish)
- ✓ Know about endangered and extinct fish
- ✓ Show the local fish to future generations and country

1.5 Report Layout

This report is divided into six chapters. The chapters are summarized below:

- Chapter one covers the introduction, motivation, objective and expected outcome of the project work and the report layout in this chapter.
- Chapter two covers the background state of our work in this chapter. We also provide a review of the literature, a summary of the research, scope of the problem and system challenges.
- Chapter three contains the proposed system methods and steps, classification, data collection procedure, algorithms and statistical analysis.

- Chapter four we have discussed experimental result, performance analysis and summary.
- Chapter five includes the conclusion and Implication of further study.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

In this section, we address several study work conducted by researchers in the field of image categorization, fish recognition. At present, people spend less time in the family's work. For those people who are now a day less recognized with local market. And for this, their ideas about local fish are low. This system is able to detect local fish and also can recognize the specific fish from images. Image processing is a technique that acquires an image and analyze it, enhance or collect useful information from images and finally it output the result in an explainable or apprehensible format. The image may be analyzed to discover samples that aren't visible by the human eye. People can take decisions after getting the output, sometimes the decisions also can be made by the machine itself.

In this research ANACONDA platform is used to train and analyze the data from numerous images of fish. ANACONDA is a platform where the digital image processing algorithms are implemented.

2.2 Literature Review

Many image processing strategies have been developed with the help of many years of research for the identification of objects and the use of classifiers. Histogram based techniques are effective for image splitting because it requires only one pass through pixels. The picture acquisition process can collect an image from some hardware-based sources, which can be used for subsequent processing.

MKSA et al. [1], Proposed a machine learning based on novel fish classification methodology to detect fish. In their research work they have achieved a large accuracy of 96.4% using Artificial Neural Network and Decision Tree. In their work, they have

combination between robust feature selection, image segmentation and geometrical parameter techniques.

UA.B et al. [2], Establish a machine learning based a hybrid method to optimize the performance with Back-propagation classifier for fish classification. They have used memetic algorithm, Back-propagation classifier (BPC), Color Histogram Technique & Gray Level Co-Occurrence Matrix (GLCM) in this research. in their work they have compared the classifier with other classifier.so the overall accuracy In this research work they have achieved a huge accuracy of 84% & 91% using BPC &HGAGD-BPC method. They have used 410 fish images for training and the rest 200 for testing.

MKA et al. [3], Proposed a Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements. They have used 350 fish images for training and the rest 150 for testing. In this research work they have achieved a large accuracy of 86% using neural networks (NN) algorithm. This research an isolated pattern of interest in the image based on the combination between robust features extraction.

KBO et al. [4], Establish a Fish Classification Based on An automatic fish image recognition system. Generally they have discussed image segmentation based on color signatures. They have used different method like Grey Level Co-occurrence Matrix, RGB color space, color histogram, image analysis, and digital image recognition. But they have achieved a huge accuracy of 84% using PBC.

S.O.O [5], Establish a SVM-based platform for fish recognition and classification is presented in this research. SVM-based fish classification, Principal Component Analysis , K-means Clustering, K-Nearest Neighbor Algorithm are they have used.in this research they have achieved a big accuracy of 78.59% where 150 fish divided into 76 training and 74 testing sets.

UA.B et al. [6], Proposed a machine learning based methodology to recognize isolated fish. in their work they have compared the classifier with other classifier.so the overall accuracy In this research work they have achieved a huge accuracy of 82% & 85% using

BPC &GAGD-BPC method ,here they have used 220 images for training and 100 images for testing phase.

NJCS [7], Develop a machine learning based methodology for the color and shape analysis of fish. They have used Fish species recognition, color and shape analysis, and fish sorting method. In this research the algorithms operate with a reliability of 100% and 98%, where 18 species of demurral and 5 species of pelagic fish are used.

KN et al. [8], Proposed a Shape-Based Fish Recognition to detect shape space. In this research Shape-Based Fish Recognition, Geodesics-based metric, marine shapes method are they have used. But in their work they have not compared the classifier with other classifier. Because in this research work they have not achieved accuracy.

M.A et al. [9], Proposed to base on Robust Extraction Hybrid approach for optimizing the performance of BPC by the ability of Memetic algorithm. In this research Back-propagation classifier, potential local geometric feature, and edge detection method are they have used.in their work they have compared the classifier with other classifier. The overall accuracy in this research work they have achieved a huge accuracy of 86% & 96% using BPC &HGAGD-BPC method.

PX.H et al. [10], Proposed a hierarchical classification approach to recognize live fish from underwater video.in this research Hierarchical classification ,SVM method are they have used .in their work they have compared the classifier with other classifier.so the overall accuracy In this research work they have achieved a huge accuracy of 86.32% , 88.08% , 90.01% using flat SVM classifier, baseline tree , automatically generated hierarchical

2.3 Research Summery

In this research we collect data from different places. Then the data set are pre-processed for next processing .For the processing of data we have made the leveling arrangement, through which there is a specific foldering system for a specific fish. Random forest classifier we found out the features of the data from different images and clusters.

Data set is ready and trained for each classification performed in Python through the Anaconda platform together with a particular label. The system is prepared to use when all the operation is completed. After pre-processing, the test picture is recorded and provided as input. Extract the characteristics and compare the chosen cluster picture to the prior dataset.

We have trained and tested several data through the image of this system, And the images are identify we used to Random Forest Classifier.

2.4 Scope of the Problem

The proposed system enables individuals identify local fish. The entire process of capturing images is implemented, even though the image is inserted as input and the implementation of the function can be true to a non-professional user who has no knowledge about this type of system.

But it is a proposed framework whose main purpose includes the main idea of how to process data properly and the right steps about how algorithms should be applied.This system can be applied to any type of platform as it is necessary. Using this technique, it is possible to readily develop mobile apps or online web applications to reach local individuals.

2.5 Challenges

To create such a system, first you need some fish pictures to be trained and then there are some more pictures to test. It was a lot of trouble for us to collect fish picture in the various fish market. Large challenges with this approach were image quality. Low resolution images can’t be processed easily. Due to the number of pixels that have a faulty area, the image becomes obscure. RGB value is most important for image processing, which should be explained first. Also the background of the image sometimes makes it difficult to identify the affected parts of the fish. If background colors are the same as fish content, segmentation can’t be done properly.

CHAPTER 3

Research Methodology

3.1 Introduction

We discuss, in this part in this part, a framework for the automatic fish recognition. The architecture of local fish recognition a machine vision approach is show in figure 3.1.Someone i.e. a people or who does not know name of the local fish name or user captures an image of a local fish using mobile phone then user send the images to our proposed system by using proposed app. After getting the images through the internet then the system process the input and give its output based on input image and at list the system send a notification into user mobile.

The input of the system is regarded as a fish picture of particular size and format, then the picture is processed. The program then recognizes the given fish. User simply take a picture with their phone, camera as input and identify the fish. Fig 3.1 shows how the proposed system works.

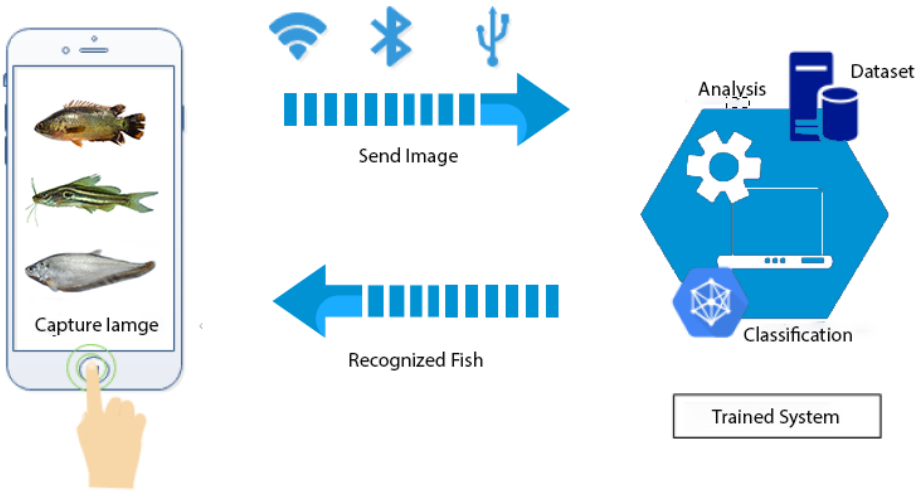


Figure 3.1: Working procedure of the system

Figure 3.2 shows the steps of approach for fish recognition, proper defect segmentation is required. Otherwise the feature non-infected part will bring under control over the features of the infected part.

In this paper K-means clustering defect segmentation is completed to detect the ROI (Region of Interest). Figure 3.2 operates in two phases testing and training. Both phases require image segmentation and feature extraction

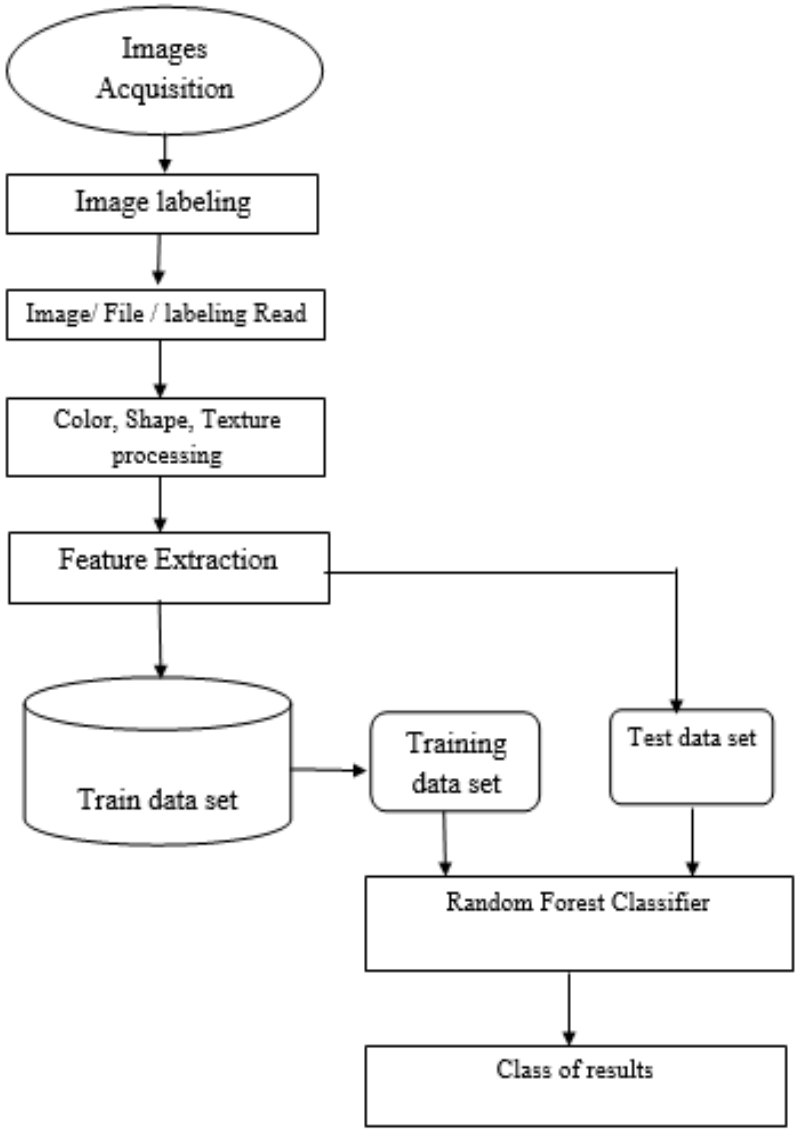


Figure 3.2: Fish recognition system

3.2 Methods and steps

3.2.1 Algorithms

Step 1 : Start

Step 2 : import all library

Step 3 : Variable and array Define

array (urls, levels)

fixed size, bins

Step 4 : images path #images folder location

Step 5 : Define methods all files name()

Step 6 : Feature selection approach

hu_moments() # Image moments capture information about the shape of a blob in a binary image

haralick() # Haralick Texture is used to quantify an image based on texture.

histogram() # for RGB or HSV.

Step 7 : feature extraction() # call define methods and processing the color image save the value in array

Step 8 : Define another variable for classification train and test data size, random state

Step 9 : Call classifier # call Random Forest classifier and given all images value

Step 10: Output # report of the classifier and accuracy

Step 11: End

3.2.2 Import Library

OpenCV2: OpenCV2 that means Open source computer vision. It is a library function. It mainly used to do all the operation related to Images.

Scikit-learn: Scikit-learn is a Learning library for python. In this program it use for Random Forest Classification

NumPy: NumPy is a Scientific Computing package in Python. NumPy package is used for various activities. NumPy Array is a multidimensional array that is used to store the same data type values.

Mahotas: Mahotas is a Python library for computer vision and image processing. The NumPy Array is a multidimensional array that is used to store the same data type values.

3.2.3 The feature selection approach

Feature extraction relates to a method using distance and geometric instruments to calculate and collect fish characteristics from color, shape and texture measurements. The objective of extracting a feature determines a larger set of characteristics.

3.2.3.1 Shape Measurement

Using shape measurements, to calculate the distance measurements and the edge (Angle measurement) of the fish object to determine the important same and different parts (like as fin, tail shape) for each different fishes .Moreover; using the distance and angle tools, the following option are often determined: The size of the mouth, dorsal fin length, caudal angle, length of caudal, angle of caudal, and the angle between the mouth and the eye. In figure 3.4 shows the Fish External Anatomy.

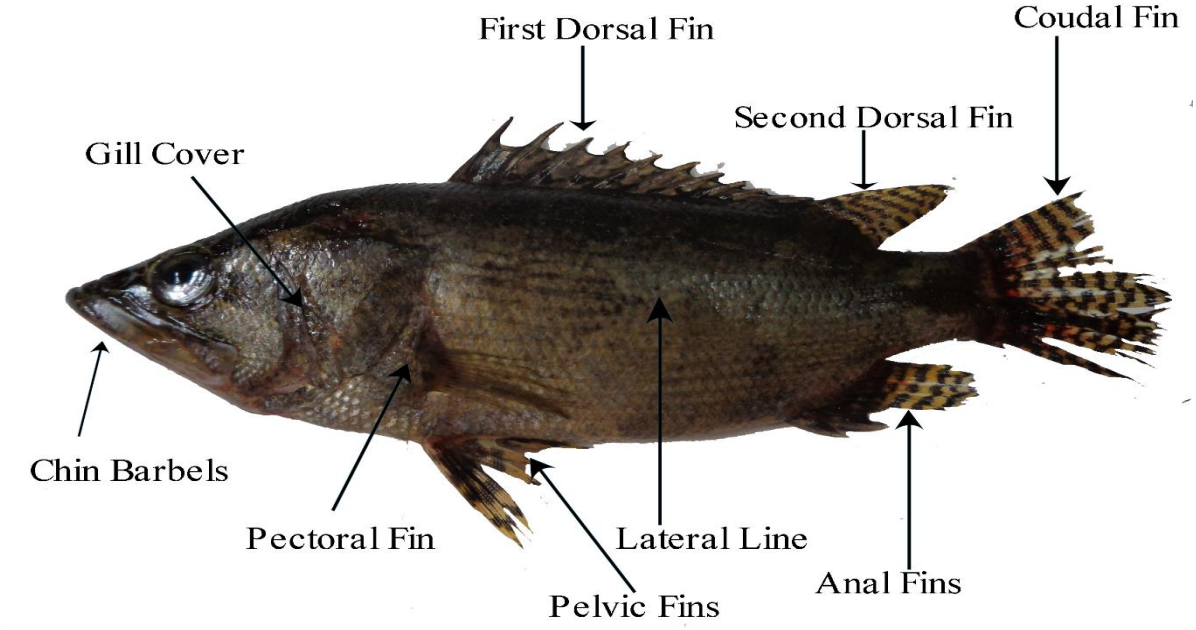


Figure 3.3 Fish External Anatomy

3.2.3.2 Hu Moment:

Image moments are an average pixel intensity weighted. For a binary image the place pixel intensity (x, y) is determined by I (x, y).

$$M = \sum_{x} \sum_{y} I(x, y) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad i$$

Image Moment Calculation where i and j are integers (1.2.3….)

$$M_{ij} = \sum_{x} \sum_{y} x^i \, y^j \, I(x, y) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad ii$$

Hu Moments calculation set 7 set of moments. The first 6 moment are provide invariant to translation, scale, and rotation, and reflection. While the sign of the seventh moment shifts for reflection of the picture.

$$I_1 = \eta_{20} + \eta_{02} \ldots\ldots\ldots\ldots\ldots\ldots iii$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \ldots\ldots\ldots\ldots\ldots\ldots iv$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad v$$

$$I_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad vi$$

$$\begin{aligned} I_5 = {} & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\ & - (\eta_{21} + \eta_{03})^2] \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad viii \end{aligned}$$

$$\begin{aligned} I_6 = {} & (\eta_{20} - \eta_{02})[(\eta_{30} - \eta_{12})^2 - (\eta_{21} - \eta_{03})^2 \\ & + 4\eta_{11}^2(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad ix \end{aligned}$$

$$\begin{aligned} I_7 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} - \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & - (\eta_{30} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\ & - (\eta_{21} + \eta_{03})^2] \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad x \end{aligned}$$

3.2.3.3 Anchor points location detection

In the previous couple of years detection of anchor point has been the goal in several research works within the pattern identification sector. Anchor points number ought to be determined of fish shape measurements in figure 3.4.

The objective of detecting anchor points in this work is to determine twenty one designated points that will assists to determine the place of each feature to recognize picture of fish.

After the landmark or anchor points have been detected over the picture, significant features will be extracted using shape and size measurements.

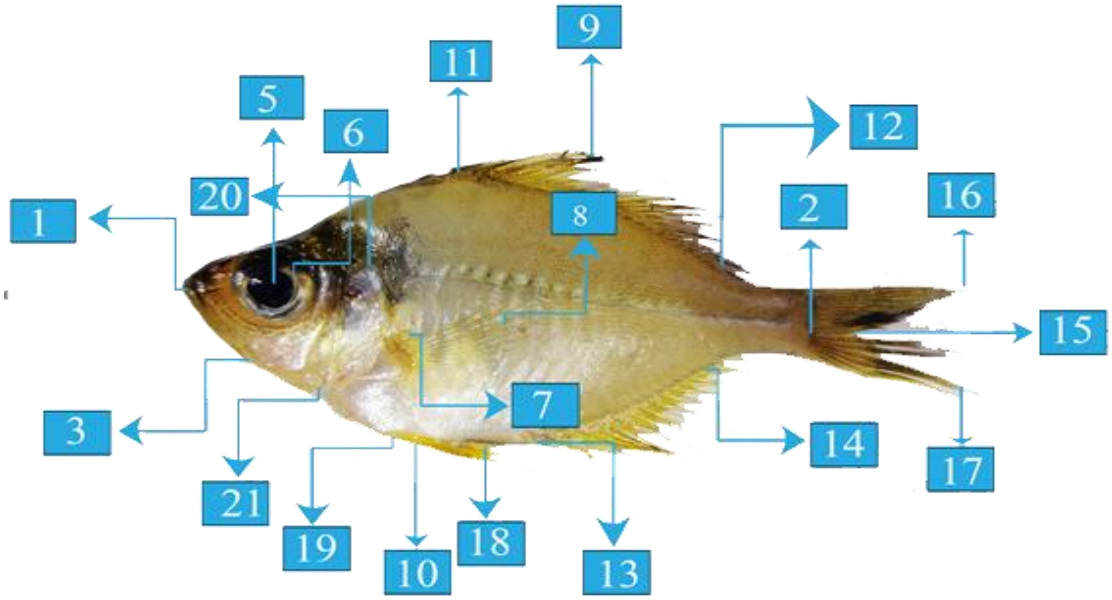


Figure 3.4: Anchor point locations

The distance of the distance measurements between 21 anchor points (P1,P2,P3,P4,P5,P6,P7,P8,P9,P10,P11,P12,P13,P14,P15,P16,P17,P18,P19,P20,P21)

3.2.3.4 Distance measurements

Distance measurements are thought-about very helpful tools within the field of pattern recognition to extract robust features so as to improve the classification accuracy. In the area of algebraic geometry, distance 'D' between the points A= (m1, n1) and B= (m2, n2) will be determined by equation 1. [15] [3]

$$D = \sqrt{\Delta m^2 + \Delta n^2} = \sqrt{(m_2 - m_1)^2 + (n_2 - n_1)^2} \quad \ldots\ldots\ldots\ldots\ldots \quad xi$$

The 23 anchor points as in fig 2 shows the length between anchor points as in table 1

Table 3.1: Measurement of distance

D.N	Feature Name	Anchor Point
D1	Length of fish without caudal fin	Dist. (P1, P2)
D2	Width of fish without fins at the top and bottom	Dist. (P11, P15)
D3	Length of the fish mouth	Dist. (P1, P3)
D4	Length of the pectoral fines	Dist. (P7, P8)
D5	Length of first dorsal fin	Dist. (P11, P12)
D6	Length of the anal fin	Dist. (P13, P14)
D7	Length of the caudal	Dist. (P2, P15)
D8	Length of the pelvic fin	Dist. (P19, P10)
D9	Width of head	Dist. (P20, P21)
D10	Length of the second dorsal fin	Dist. (P22, P23)
D11	Distance between the right _end of first dorsal fin and the start of second dorsal fin	Dist. (12, P23)
D12	Distance between end of the pelvic fin and the start of the anal fin	Dist. (18, P13)

3.2.3.5 Calculating the Angels

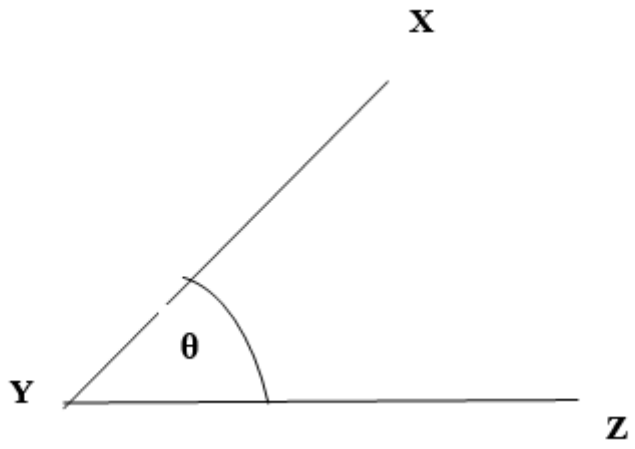


Figure 3.5: Calculating the Angels

An Angle occurs when two vectors meet a common end point. Common endpoint is characterizes because angle vertex. Figure 2 shows the angles between two vectors, can be represented as $\angle XYZ$ or$\angle ZYX$. This angle you can conjointly write as $\angle Y$ that names the vertex. Using the cosine rule we will be calculated the angle θ by the following equation: [16] [3]

$$b^2 = c^2 - 2ac\,cosB \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ xii$$

$$\cos B = \frac{a^2+c^2-b^2}{2ac} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ xiii$$

Table 3.2 demonstrates calculated the five angle features from angle measurement supported the anchor points in fig 3.5

Table 3.2: Measurement of angle

A.N	Feature Name	Anchor Description
A1	The smaller triangle angle	P15, P10, P17
A2	The bottom triangle angle	P15, P9, P16
A3	Angle of caudal fin	P16, P15, P10
A4	Angle of fish head	P20, P21, P1
A5	Angle of front triangle	P9, P15, P10
A6	Angle of whole fish	P16, P4, P17
A7	Angle of eye-end mouth	P1, P3, P5
A8	Angle of second caudal	P9, P4, P10
A9	Angle of rear triangle	P9, P16, P10

3.2.3.6 Color signature (Color Histogram)

Color histogram is the representation of color distribution in a picture. Color histogram is the quantity of pixels with colors in each of the specified color ranges that span the image's color space for digital color. Color histogram is based on some color space such as channels R, G and B. Histogram of intensity used in monochromatic pictures. The color histogram is N-dimension for multi-spectral pictures, where each pixel is represented by an arbitrary amount of measurements. [4]

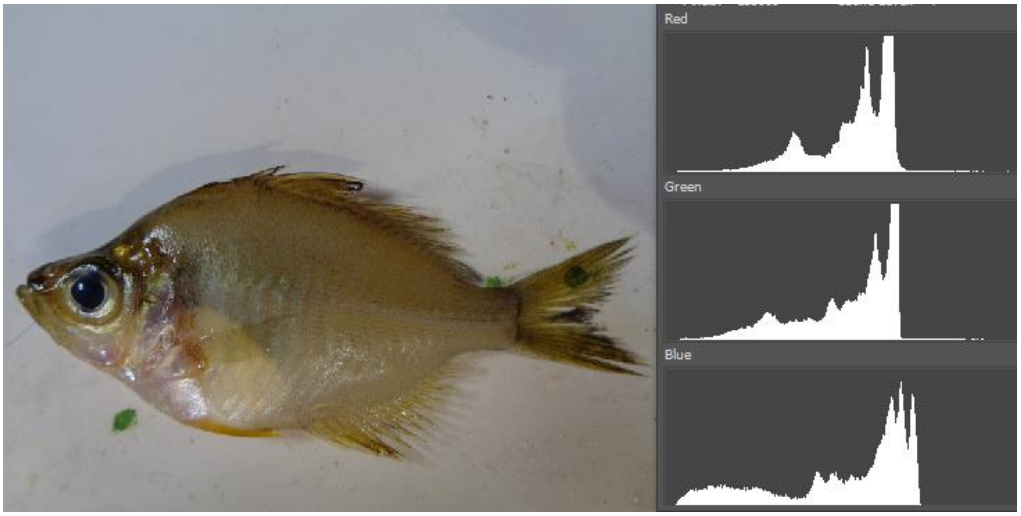


Figure 3.6 : Color histogram for a fish

Six separate steps happen in the processing of the color signature extraction. The acquisition of images. It is essential that pictures are of high quality color include in our research database. Then remove the crop from the ventral portion of the interest pattern (fish object). The third step is using RGB color space extract the color value from the captured crop. As shown in figure, the RGB matrix was divided into individual matrixes (Red, Green and Blue). 4^{th} step is R.G.B color matrix contains a color histogram. 5^{th} step is beginning with the color histogram method in which 3 features (Standard deviation, Homogeneity, Energy) are obtained using GLCM. Finally store the characteristics acquired.

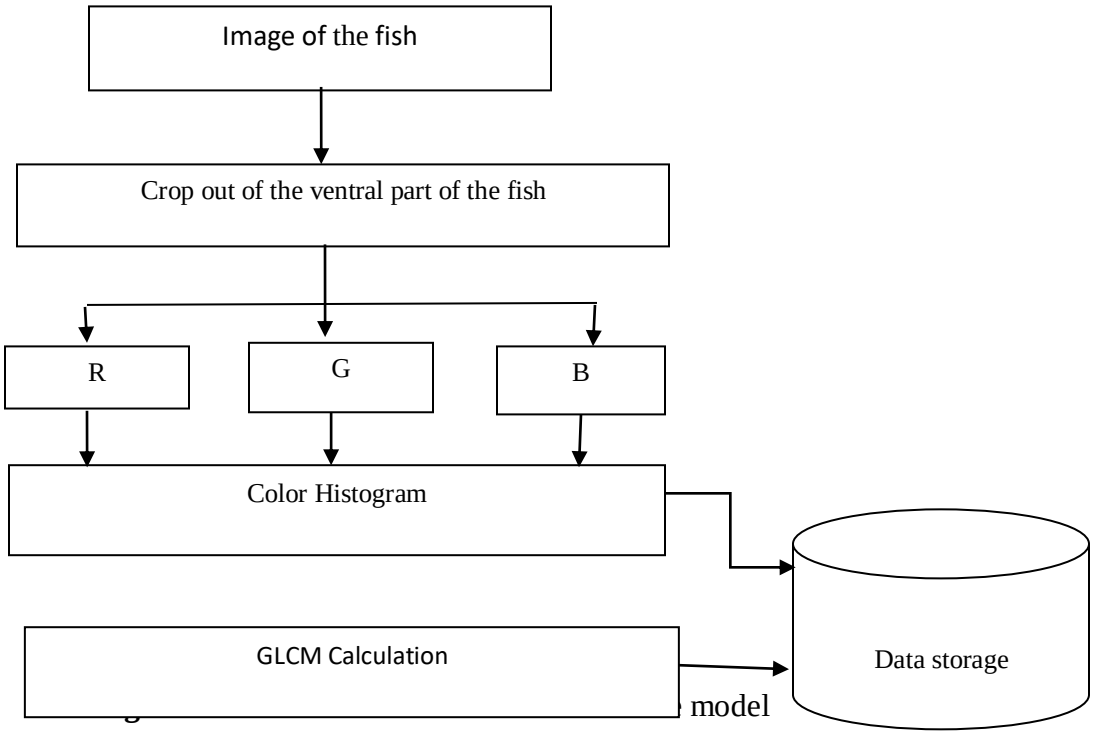


Figure 3.7: The extracted color-based feature model

3.2.3.7 Texture Measurement (Haralick Texture)

Texture is one of the significant feature used to identify objects or areas of concern in a picture. Used Haralick Texture to quantify a texture-based picture. The basic idea of Haralick Texture computing characteristics is the GLCM or Gray Level Co-occurrence Matrix. The fundamental concept of GLCM is to look for pairs of contiguous pixel values in a picture and keep recording it throughout the picture. [5]

3.2.4 Classification:

In this program we use Random Forest Classification. A random forest is a machine learning information structure that creates big numbers of random decision trees that analyze sets of variables. This sort of algorithm enables to improve how complicated information is analyzed by techniques.

Random Forest Ensembles are a dividing and conquering strategy used to enhance the efficiency of separately weak models of the Decision Tree [17] [18].

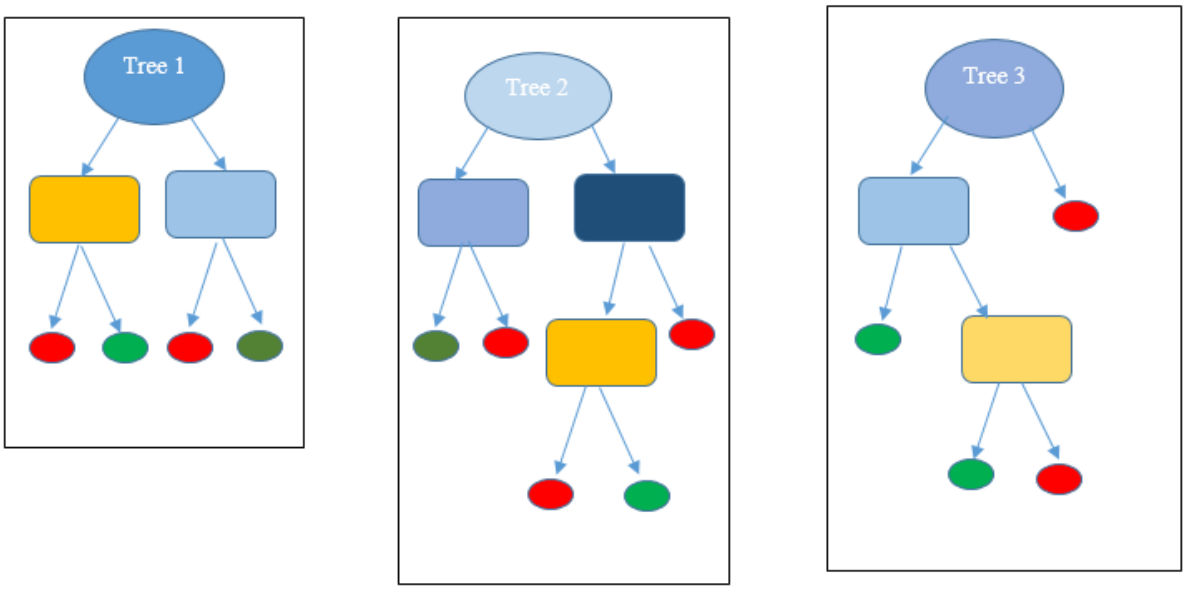


Figure 3.8: Random Forest Classifier Tree

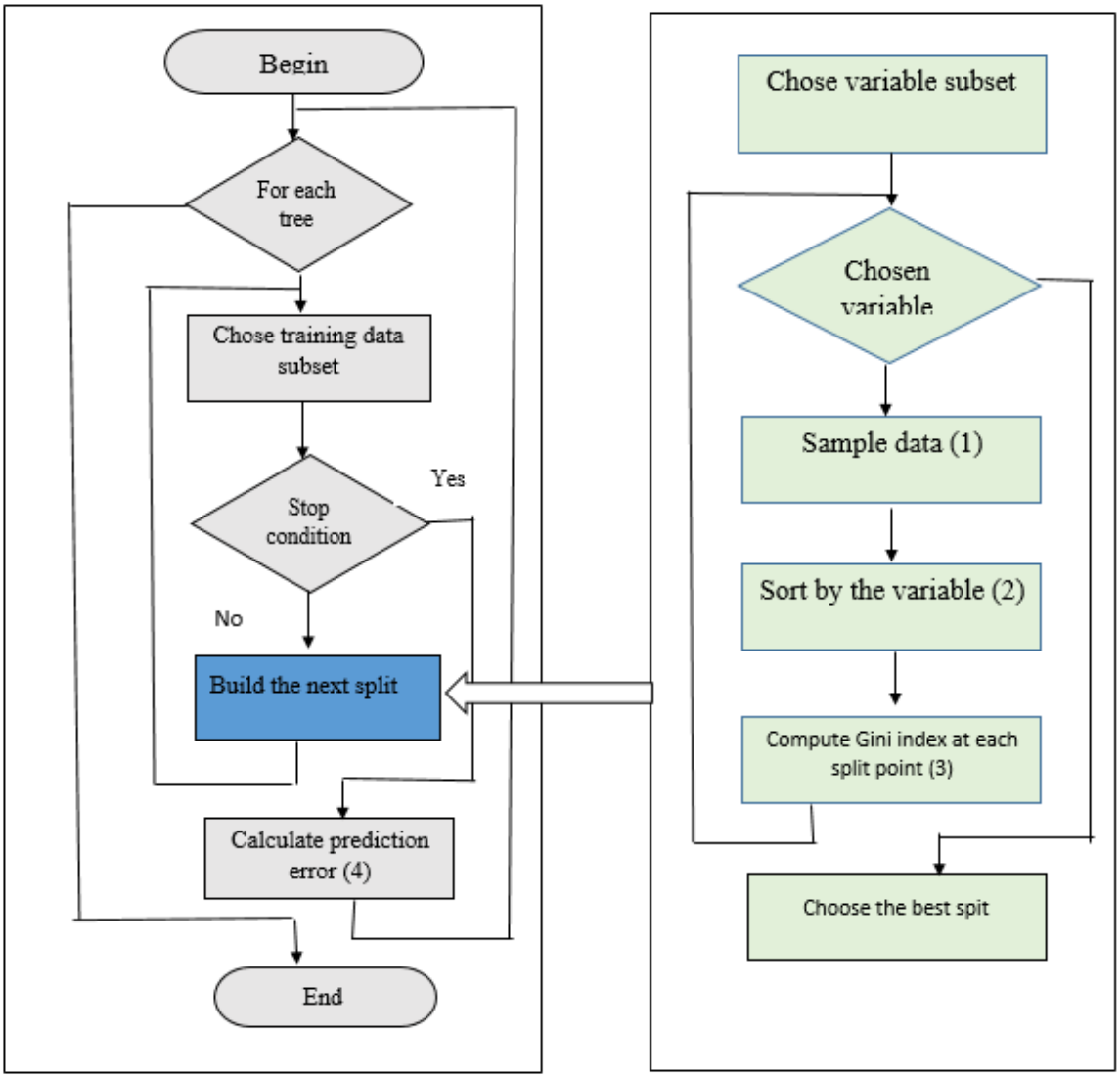


Figure 3.9: Algorithms of Random Forest Classifier

Margin Function: The classifier margin function is described as

$$mg(\underline{X}, Y) = \frac{\sum_{k=1}^{K} I(h_k(\underline{X}) = Y)}{K} - \max_{j \neq Y} \left[\frac{\sum_{k=1}^{K} I(h_k(\underline{X}) = j)}{K} \right] \quad \ldots\ldots\ldots \quad xiv$$

When mg(X, Y) greater than 0 then the classifiers sets votes for the right classification.

If mg(X, Y) less than 0 then the classifiers set votes for a classification that is wrong

The mistake in generalization is:

$$PE^* \to P_{\underline{X},Y}\left(P_{\underline{\theta}}(h(\underline{x},\underline{\theta}) = Y)) - \max_{j \neq Y} P_{\underline{\theta}}(h(\underline{x},\underline{\theta}) = j)) < 0\right) \quad \ldots\ldots\ldots\ldots \quad xv$$

3.3 Data collection procedure

We notice that in many local markets there are different kinds of local fish. We used to go to many local markets and persuaded the vegetable vendor to photography different local fishes and we said that we are working with local extinct fish. The row pictures differ in size, shape and category with distinct combinations of colors and textures. From occasionally we purchased fish from the market and took photos of fish. All pictures are prepossessed in order to obtain all of them in an appropriate type of information.

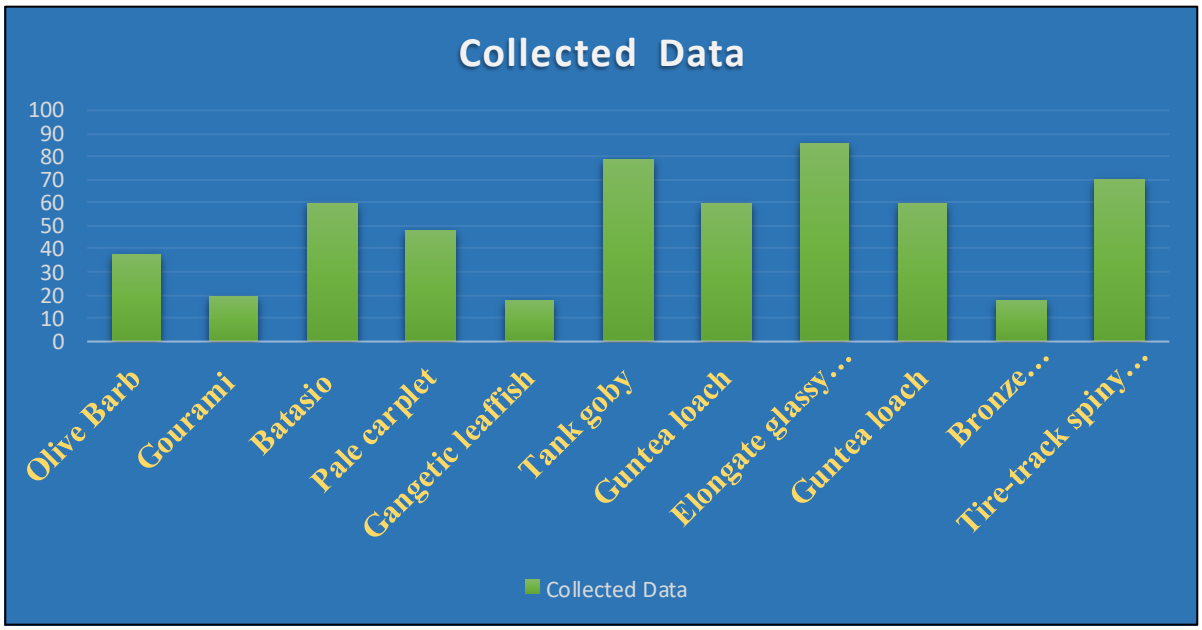


Figure 3.10 Collect data procedure

3.4 Statistical Analysis

According to the information assessment, 88 percent is accomplished in calculating the precision of the suggested scheme. Using the confusion matrix, some more characteristics are obtained to calculate this. For instance, the vegetable detection confusion matrix is provided below.

```
confussion matrix:
[[16  0  0  0  0  0  0  0  0  0]
 [ 0  6  0  0  0  0  0  2  0  0]
 [ 0  0 19  0  0  0  0  1  0  0]
 [ 0  0  0  5  0  0  0  0  0  0]
 [ 0  0  0  0  3  0  0  0  0  0]
 [ 0  0  0  0  0  8  1  0  1  0]
 [ 0  0  0  0  0  0  6  2  0  0]
 [ 0  0  0  0  0  0  0 17  0  0]
 [ 0  0  0  0  0  1  0  0 14  3]
 [ 0  0  0  0  0  1  1  0  2 16]]
```

Figure 3.11: Confusion Matrix

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

The raw pictures were first gathered and captured from varied farms, fish markets and some supper shops to get the final outcome. The primary trained dataset was prepared after data pre-processing. Using Random forest classifier, the test picture is finally likened to the data set and then final outcome is displayed.

4.2 Experimental Result

The detection in this article shows ten kinds of local fish.
The domain used is as follows:

1. Batasio ….……60
2. Bronze Featherback …….....17
3. Elongate Glassy Perchlet ….……96
4. Gangetic Leaffish ….……...……17
5. Gourami ……..19
6. Guntea Loach .....……60
7. Olive Barb .....……37
8. Pale Carplet ...……...48
9. Tank Goby …….……79
10. Tire-track Spiny Eel ……..70

About 500 pictures were used in our information set to identify local fish.

Figure 4.1: Images of Batasio Fish

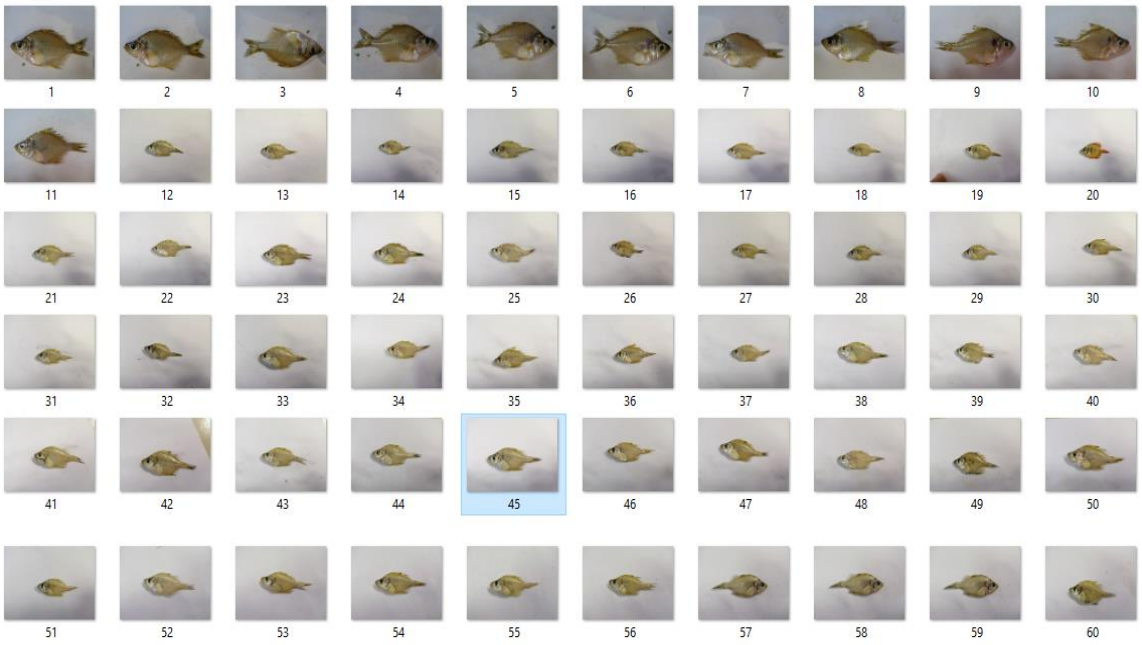


Figure 4.2: Images of Elongate glassy perchlet Fish

Figure 4.3: Images of Pale carplet Fish

Figure 4.4: Images of Tire-track spiny eel Fish

The pre-processing of these pictures was performed after picture acquisition. Then we train those data set and level them. Extraction of the function was performed after this. Now step by step here is a screen short of the project:

```
In [21]: len_images_level = LabelEncoder()
         encoder=len_images_level.fit(labels)
         encoded_labels=encoder.transform(labels)
         print("Number of the Image Level: ",len(set(labels)))

         Number of the Image Level:  10
```

Figure 4.5: Number of the images levels

```
In [276]: len_images_level = LabelEncoder()
          encoder=len_images_level.fit(labels)
          print(encoder.classes_)
          encoded_labels=encoder.transform(labels)

          ['Batasio' 'Bronze featherback' 'Elongate glassy perchlet'
           'Gangetic leaffish' 'Gourami' 'Guntea loach' 'Olive Barb' 'Pale carplet'
           'Tank goby' 'Tire-track spiny eel']
```

Figure 4.6: Images labels name

In [69]:
```
import cv2
import matplotlib.pyplot as plt
image = cv2.imread("1.jpg")
image1 = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
image2 = cv2.cvtColor(image, cv2.COLOR_BGR2HSV)

plt.subplot(1,2,1)
plt.imshow(image1)
plt.title('image')

plt.subplot(1,2,2)
plt.imshow(image2)
plt.title('image')

plt.show()
cv2.waitKey(0)
```

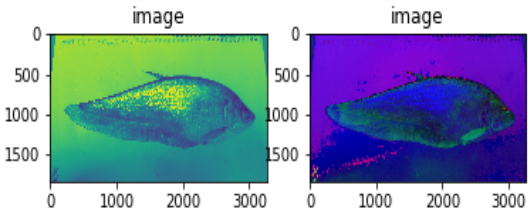


Figure 4.7: Image conversion BGR to GRAY and HSV

```
X_train, X_test, y_train, y_test = train_test_split(data, labels, test_size=0.25, random_state=42)
print('Total train data: ', len(X_train))
print('Total test data: ', len(X_test))
from sklearn.ensemble import RandomForestClassifier
clf = RandomForestClassifier(n_estimators=100, random_state=9)
clf.fit(X_train, y_train)
prediction=clf.predict(X_test)
```

```
Total train data:  375
Total test data:  125
```

Figure 4.8: Train and Test Data

```
In [35]: from sklearn.ensemble import RandomForestClassifier
         clf = RandomForestClassifier(n_estimators=100, random_state=9)
         print(clf.fit(X_train, y_train))

RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
            max_depth=None, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=100, n_jobs=None,
            oob_score=False, random_state=9, verbose=0, warm_start=False)
```

Figure 4.9: Using Random Forest Classifier

```
In [275]: from sklearn.metrics import precision_score, \
          recall_score, confusion_matrix, classification_report, \
          accuracy_score, f1_score
          print('Accuracy:', accuracy_score(y_test, prediction))
          print('F1 score:', f1_score(y_test, prediction,average='micro'))
          print('Recall:', recall_score(y_test, prediction,average='micro'))
          print('Precision:', precision_score(y_test, prediction,average='micro'))
          print('\n clasification report:\n', classification_report(y_test,prediction))
          print('\n confussion matrix:\n',confusion_matrix(y_test, prediction))

Accuracy: 0.88
F1 score: 0.88
Recall: 0.88
Precision: 0.88
```

Figure 4.10: Accuracy of local fish recognition

```
clasification report:
                          precision    recall  f1-score   support

                 Batasio       1.00      1.00      1.00        16
      Bronze featherback       1.00      0.75      0.86         8
Elongate glassy perchlet       1.00      0.95      0.97        20
       Gangetic leaffish       1.00      1.00      1.00         5
                 Gourami       1.00      1.00      1.00         3
            Guntea loach       0.80      0.80      0.80        10
              Olive Barb       0.75      0.75      0.75         8
            Pale carplet       0.77      1.00      0.87        17
               Tank goby       0.82      0.78      0.80        18
    Tire-track spiny eel       0.84      0.80      0.82        20

               micro avg       0.88      0.88      0.88       125
               macro avg       0.90      0.88      0.89       125
            weighted avg       0.89      0.88      0.88       125


confussion matrix:
[[16  0  0  0  0  0  0  0  0  0]
 [ 0  6  0  0  0  0  0  2  0  0]
 [ 0  0 19  0  0  0  0  1  0  0]
 [ 0  0  0  5  0  0  0  0  0  0]
 [ 0  0  0  0  3  0  0  0  0  0]
 [ 0  0  0  0  0  8  1  0  1  0]
 [ 0  0  0  0  0  0  6  2  0  0]
 [ 0  0  0  0  0  0  0 17  0  0]
 [ 0  0  0  0  0  1  0  0 14  3]
 [ 0  0  0  0  0  1  1  0  2 16]]
```

Figure 4.11: Final Result of the classifier

4.3 Performance Analysis

Here we use Random Forest Classification to identify the name of the local Fish. We also use methods of moment of HU, Color Histogram and Haralick Textures to extract data sets characteristics.

Finally, the local fish name anticipated will be printed on the screen. Here in our confusion matrix:

```
confussion matrix:
[[16  0  0  0  0  0  0  0  0  0]
 [ 0  6  0  0  0  0  0  2  0  0]
 [ 0  0 19  0  0  0  0  1  0  0]
 [ 0  0  0  5  0  0  0  0  0  0]
 [ 0  0  0  0  3  0  0  0  0  0]
 [ 0  0  0  0  0  8  1  0  1  0]
 [ 0  0  0  0  0  0  6  2  0  0]
 [ 0  0  0  0  0  0  0 17  0  0]
 [ 0  0  0  0  0  1  0  0 14  3]
 [ 0  0  0  0  0  1  1  0  2 16]]
```

Figure 4.12: Confusion Matrix

And Finnaly result:

```
clasification report:
                          precision    recall  f1-score   support

                 Batasio       1.00      1.00      1.00        16
      Bronze featherback       1.00      0.75      0.86         8
Elongate glassy perchlet       1.00      0.95      0.97        20
       Gangetic leaffish       1.00      1.00      1.00         5
                 Gourami       1.00      1.00      1.00         3
            Guntea loach       0.80      0.80      0.80        10
              Olive Barb       0.75      0.75      0.75         8
            Pale carplet       0.77      1.00      0.87        17
               Tank goby       0.82      0.78      0.80        18
    Tire-track spiny eel       0.84      0.80      0.82        20

               micro avg       0.88      0.88      0.88       125
               macro avg       0.90      0.88      0.89       125
            weighted avg       0.89      0.88      0.88       125
```

Figure 4.13: Classification Report

4.4 summary

Fisheries plays a significant role in Bangladesh's earning in food, nutrition, jobs and export. Approximately 5 percent of the total national manufacturing of the country is

obtained from this industry and around 14 million individuals are engaged in fishing all the time. Sometime when our fishermen go fishing, they don't recognize the fish.

By using our method, fishermen will simply apprehend the fish information. When people go to market for buy fishes for their families but they have no idea about local fish, it also help them to identify the fish. Knowing about local fishes will also be useful for the kids. Our proposed technique also helps create more applications like this by new developers.

CHAPTER 5

CONCLUSIONS AND FUTURE RESEARCH

5.1 Conclusions

In this paper, we present a methodology for local fish classification based on a combination between random forest classifier, global features extraction, and geometric parameters. 500 hundred fish pictures families are regarded for experimentation purposes, 375 fish pictures for training and the remainder 125 for testing. We focus our research to recognize and classify of local fish based on digital image and to development of a new prototype of local fish recognition using Random Forest Classifies, color, shape, texture features, geometric parameters (distance and angle measurement).Local fish are recognized in this work through of image processing methods that can be used in various apps to identify any fish that can open a door to help fishermen, people and kids. Although some obstacles existed while working, the background color and bad image quality may distract the application to produce more extract results.

The entire method is carried out with 89% accuracy using pictures from distinct angles of fishes gathered.

5.2 Implication of further study

Future research includes experimentation with many datasets, classifier and methods. In future we will develop a mobile application. In future we will work all kind of fish. The proposed mechanism demonstrates a way of involving the machine learning technique that can detect fish and also automatically recognize the fish. Sometimes the fishermen are not sure that what fish they have caught and not sure if can keep it, they needed help to identifying the species. They just take a photo with their phone or other, the app identify fish and show classify of this fish.

In Bangladesh, about 64 species of fish have been extinct. We want to contribute to bring the extinct fish and introduce them to the people. Our children eat fish but don’t know the

fish’s name or recognize fish. We will help the future generation and fishermen to recognize local fish.

Therefore, future study seeks to improve the rate of classification and to classify fish on bigger datasets and species.

REFERENCES

[1] A. M. K., Omar, K., Mohd Noah, S. A., & Almarashdeh, I. ,"Fish Recognition Based On The Combiation Between Robust Features Selection,Image Segmentation and Geometrical Parameters " ,Computer Science and Information Journal, Vol (6) No(2), pp.215-221, 2009.

[2] Usama A. Badawi & A. M. K., " A Hybrid Memetic Algorithm With Back-Propagation Classifier for Fish Recognition ", Computer Science International Journal, Vol. 10, Issue 2, No 1,pp.384-356, March 2013.

[3] Alsmadi, M. K., Omar, K., Mohd Noah, S. A., & Almarashdeh, I. ," Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements ", Computer Science Journal6 (10): 1059-1065, 2010.

[4] MK Alsmadi, Omar, K., Mohd Noah, S. A.," Fish Classification Based on Robust Features Extraction from Color Signature Using Back-Propagation Classifier ", Journal of Computer Science 7 (1): 52-58, 2011.

[5] S.O. Ogunlana, O. Olabode , S.A. A. Oluwadare & G. B. Iwasokun, " Fish Classification Using Support Vector Machine",African Journal of Computing & ICT, Vol 8, No. 2. Pp 75-82, June, 2015.

[6] Usama A. Badawi & Alsmadi, M. K., "A General Fish Classification Methodology Using Meta-Heuristic Algorithm With Back Propagation Classifier ", Journal of Theoretical and Applied Information Technology, Vol. 66 No.3, 803-812,31st August 2014.

[7] N J C Strachan," Recognition of fish speciesby colour and shape ",Image and Vision Computing ,Volume 11, Issue 1, Pages 2-10, January–February 1993.

[8] Kamal N. and Abdesslam B., "Shape-Based Fish Recognition via Shape Space ", pre European Signal Processing Conference, 145-149,23rd 2015.

[9] Alsmadi, M. K., Omar, K., Mohd Noah, S. A., & Almarashdeh, I. , "A hybrid Memetic algorithm with Back-propagation classifier for fish classification based on Robust Extraction from PLGF and Shape Measurements", Information Technology Journal10(5):944-954,2011.

[10] Phoenix X. H., Bastian J. B., Robert B. F., "Hierarchical Classification for Live Fish Recognition", 1-10, 2012.

[11] ROBERT M. HARALICK, K. SHANMUGAM, and ITS, HAK DINSTEIN, “Textural Features for Image Classification”IEEE TRANSACTIONS ON SYSTEM, MAN AND CYBERNETICS, Vol (6), pp. 610-621, 1973

[12]Local fish information, that available in <<https://en.wikipedia.org/wiki/Fishing_in_Bangladesh>>, we last accessed on 05-04-2019 at 2:00pm.

[13]Local Fish Bengali name, available in<<https://en.wikipedia.org/wiki/List_of_fishes_in_Bangladesh >>, we last accessed on 03-05-2019 at 11:00pm.

[14] Some fish information, available at << https://www.worldfishcenter.org/country-pages/bangladesh>>, we last accessed on 07-07-2019 at 12:50pm.

[15] Distance formula available in <<https://www.varsitytutors.com/hotmath/hotmath_help/topics/distance-formula >>, we last accessed on 05-07-2019, 8.00pm

[16] Angled triangle formula that available in <<https://www.mathsisfun.com/algebra/trig-finding-angle-right-triangle.html>>, we last accessed 07-07-2019, 12:50pm.

[17] Random Forest Algorithm, available in << https://dataaspirant.com/2017/05/22/random-forest-algorithm-machine-learing/>>, we last accessed on 07-07-2019 at 12:50pm.

APPENDICES

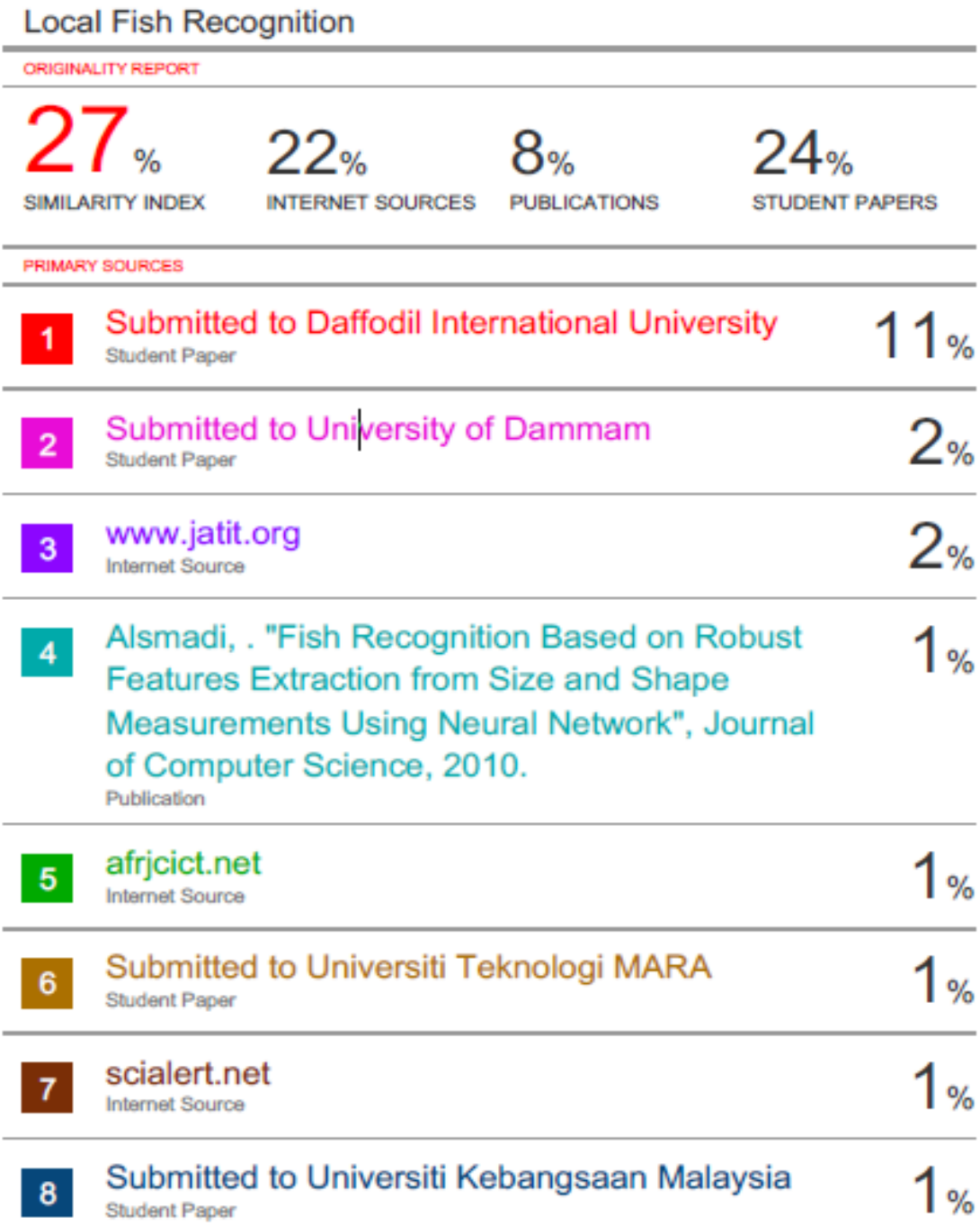
Appendix A: Research Reflection

This Appendix is intended to provide an introduction to the Refection of Research. The study project for the group was a difficult and pleasant experience that was typical of the entire course. We had little exposer at university to group work. So, being part of an efficient and vibrant team was a good change.

The experience has taught us that it takes longer to plan and develop reactions in team.

Appendix B: Related Issues

It was very hard to collect pictures from this kind of town area like Dhaka. We had to go to villages and markets to capture the pictures of local fish. We had to speak to the fisherman a while to let them know the study's issue and significance. They were very helpful to assist us. We had to study many new techniques, features and algorithms to execute our ideas and research work. Variation of picture backgrounds and image quality were difficult to alter and decrease the result changes hereby.



Local Fish Recognition

ORIGINALITY REPORT

27%	22%	8%	24%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	11%
2	Submitted to University of Dammam Student Paper	2%
3	www.jatit.org Internet Source	2%
4	Alsmadi, . "Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements Using Neural Network", Journal of Computer Science, 2010. Publication	1%
5	afrjcict.net Internet Source	1%
6	Submitted to Universiti Teknologi MARA Student Paper	1%
7	scialert.net Internet Source	1%
8	Submitted to Universiti Kebangsaan Malaysia Student Paper	1%

*****End****