# Species based Frozen/Fresh Fish Classification Using Texture Based Image Analysis


Project Team

| | |
|---|---|
| Babar Ali | p16-6087 |
| Faraz Naeem | p16-6094 |
| Muhammad Shamyl | p15-6407 |

Session 2016-2020

Supervised by

Dr. Hafeez ur Rehman

**Department of Computer Science**

**National University of Computer and Emerging Sciences**
**Peshawar, Pakistan**


**August, 2020**

# Student's Declaration

We declare that this project titled "*Species based Frozen/Fresh Fish Classification Using Texture Based Image Analysis*", submitted as requirement for the award of degree of Bachelors in Computer Science, does not contain any material previously submitted for a degree in any university; and that to the best of our knowledge, it does not contain any materials previously published or written by another person except where due reference is made in the text.

We understand that the management of Department of Computer Science, National University of Computer and Emerging Sciences, has a zero tolerance policy towards plagiarism. Therefore, We, as authors of the above-mentioned thesis, solemnly declare that no portion of our thesis has been plagiarized and any material used in the thesis from other sources is properly referenced.

We further understand that if we are found guilty of any form of plagiarism in the thesis work even after graduation, the University reserves the right to revoke our BS degree.

| | |
|---|---|
| Babar Ali | Signature: ________________ |
| Faraz Naeem | Signature: ________________ |
| Muhammad Shamyl | Signature: ________________ |

____________________________

Verified by Plagiarism Cell Officer
Dated:

# Certificate of Approval

The Department of Computer Science, National University of Computer and Emerging Sciences, accepts this thesis titled *Species based Frozen/Fresh Fish Classification Using Texture Based Image Analysis*, submitted by Babar Ali (p16-6087), Faraz Naeem (p16-6094), and Muhammad Shamyl (p15-6407), in its current form, and it is satisfying the dissertation requirements for the award of Bachelors Degree in Computer Science.

**Supervisor**

Dr. Hafeez ur Rehman Signature: __________________

______________________________

Shakir Ullah

FYP Coordinator
National University of Computer and Emerging Sciences, Peshawar

______________________________

Dr. Hafeez ur Rehman

HoD of Department of Computer Science
National University of Computer and Emerging Sciences

# Acknowledgements

We are very thankful to Almighty Allah who is most Merciful and most Beneficent for giving us strength and courage to choose such challenging project and for completing Final Year Project Report with great dignity and grace. However, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend our sincere thanks to all of them. We would like to express deep gratitude to our project supervisor Associate Professor Dr. Hafeez Ur Rehman, for his guidance, encouragement and useful suggestions of this research work. His guidance and experience proved very helpful to us in order to make progress in our project. We are highly indebted to FAST NUCES university Peshawar for providing us platform to showcase our work. We would like to thank and appreciate our colleagues for giving us useful tips and suggestions in developing the project and people who have willingly helped us out with their abilities.

Babar Ali

Faraz Naeem

Muhammad Shamyl

# Abstract

Fish is a widely consumed item around the globe but Frozen thawed fish being sold as fresh fish is becoming a common fraudulent which can cause many health problems. Unlike fresh fish, thawed fish requires undergoing a lengthy and costly process to reduce degenerative changes which could in turn cause severe diseases such as Food poisoning. This fraudulent of mislabelling frozen fish as fresh can be detected and eradicated with the help of Histology and Digital Image Processing Techniques. Following texture-based analysis, observation of protein to fat ratios and empty spaces within cells, Fresh fish can be distinguished from Frozen thawed. The main idea of this work is to observe fat to protein ratios in histology images which constitute the edges in the image. Images were exposed to cross-bilateral filter which was helpful in smoothing the image while edges in the images were preserved which showed good contrast in order to distinguish fat from muscle fiber. Glcm and Gabor filter were applied separately and combined on the images producing feature vectors which were then used in 5-fold cross validation to extract results. When Classifying intra fish specie into fresh and frozen, Glcm feature results showed up to 78% accuracy and Gabor filter showed 82% but when both Gabor and Glcm features were combined, it showed of accuracy up to 88%. This solution can be applied on Industrial level by applying on random samples on incoming batches via comparison with previously fed data set.

# Contents


**1 Preliminaries and Introduction** **1**

1.1 Introduction . . . . . . . . . . 1

1.2 Motivation . . . . . . . . . . 2

1.3 Problem Statement . . . . . . . . . . 2

1.4 Objectives . . . . . . . . . . 2

1.5 Project Scope . . . . . . . . . . 3

**2 Review of Literature** **4**

2.1 Some Preprocessing Techniques . . . . . . . . . . 5

2.1.1 Cross-bilateral filter . . . . . . . . . . 5

2.1.2 Sobel Operator . . . . . . . . . . 5

2.2 Some Feature Extraction Techniques . . . . . . . . . . 6

2.2.1 GLCM . . . . . . . . . . 6

2.2.2 Local Binary Patterns . . . . . . . . . . 6

2.2.3 Gabor Filter . . . . . . . . . . 6

2.2.4 CNN . . . . . . . . . . 7

2.3 Some Classifiers . . . . . . . . . . 8

2.3.1 SVM . . . . . . . . . . 8

2.3.2 KNN . . . . . . . . . . 8

**3 System Analysis and Design** **9**

3.1 Use Case Diagram . . . . . . . . . . 10

3.2 Sequence Diagram . . . . . . . . . . 11

**4 System Implementation** **12**

4.1 Dataset . . . . . . . . . . 13

4.2 PreProcessing . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
4.2.1 Data Augmentation . . . . . . . . . . . . . . . . . . . . . . . 15
4.2.1.1 Noise Injection . . . . . . . . . . . . . . . . . . . . . . 15
4.2.1.2 Shearing . . . . . . . . . . . . . . . . . . . . . . . . . 15
4.2.1.3 Rotation . . . . . . . . . . . . . . . . . . . . . . . . . 16
4.2.1.4 Flip . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
4.2.1.5 Brightness . . . . . . . . . . . . . . . . . . . . . . . . 16
4.2.2 Greyscale Conversion . . . . . . . . . . . . . . . . . . . . . . 17
4.2.3 Cross Bilateral Filter . . . . . . . . . . . . . . . . . . . . . . 18
4.3 Feature Extraction . . . . . . . . . . . . . . . . . . . . . . . . . . 19
4.3.1 Grey Level Co Occurence Matrix . . . . . . . . . . . . . . . . 19
4.3.2 Gabor filter . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
4.4 Classification . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
4.4.1 Support Vector Machine . . . . . . . . . . . . . . . . . . . . . 24
4.5 Graphical User Interface . . . . . . . . . . . . . . . . . . . . . . . 25

**5 Results** **26**

**6 Discussion** **29**

**References** **32**

# List of Figures

3.1 Sequence Diagram . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

4.1 Fresh and Frozen Orata samples. . . . . . . . . . . . . . . . . . . . . . . . 13

4.2 Fresh and Frozen Orata samples. . . . . . . . . . . . . . . . . . . . . . . . 14

4.3 Image after Grey Scale Conversion . . . . . . . . . . . . . . . . . . . . . . 17

4.4 Image after cross Bilateral Filter . . . . . . . . . . . . . . . . . . . . . . 18

4.5 Graphical User Interface of Software . . . . . . . . . . . . . . . . . . . . . 25

5.1 Classification results using 70/30 split. . . . . . . . . . . . . . . . . . . . 27

5.2 Classification results using 5-fold cross validation. . . . . . . . . . . . . . 27

# List of Tables

# Chapter 1

# Preliminaries and Introduction

## 1.1 Introduction

Fish is a product that has been widely consumed all around the globe. It has grown upon consumers for consumption in both uncooked and cooked form. Due to its vast consumption in raw form, demand of fresh fish has greatly increased in seafood retailing. According to a study by Oceana which reveals that 25 to 70 percent of sea food in America only is mislabelled that mainly includes Salmon and Atlantic Cod [15]. The life expectancy of fresh fish is very less as degenerative changes begin to occur soon after catching fish. Fish for raw consumption has to undergo lengthy and time-consuming procedure of initially freezing at -20C for at least 24h to kill parasites. Therefore, fraudulent of selling thawed fish as fresh has been observed as a common practice to meet both demand and reduce provision cost. Selling defrosted fish with a label of fresh fish product comprises qualitative fraud, due to the physical changes in products produced happened because of defrosting procedures which can affect the nutritional value as well as it becoming unfit for use. More importantly, this fraudulent results in incorrect shelf life expectancy leading to serious diseases upon consumption. Data on techniques for preservation of fish items passed on through exact naming is critical for the shopper's view of item quality. Fish populations can differ in freezing quality although belonging to same fish species. The frozen condition of fish affects microorganisms and sensory attributes can also be altered by freezing. Histology has been considered as a solution to the said problem. It can be applied on Industrial level for distinguishing of fresh state from frozen by texture-based analysis following protein to fat ratios and observation of tissue structure.

## 1.2 Motivation

Quality of food has raised a lot of concerns in the market ever since discovery of its association to different diseases. Fish is a widely consumed product famous for being rich in protiens. We already know what bad food can do as there is a wide concern that Corona disease also originated due to consumption of food item. Fraudulent of selling frozen thawed as fresh fish will result in estimate of incorrect shelf life expectancy. This would in turn lead to degenerative changes which can turn out to be seriously harmful and can lead to serious diseases. Hence an application which could tell the state of fish from a batch in consideration ensure on an industrial scale that consumer end is served fresh and not frozen product like desired. This software upon being trained on dataset could in future be also trained for classification of similar materials.

## 1.3 Problem Statement

Fish being a widely consumed item has a huge chance of being mislabelled, hence there is a need to detect and classify different fish species and also if they are fresh or frozen. Main difference between fresh and frozen thawed fish is protein and fat ratio in their muscle fiber as freezing influences fat decomposition. Inaccuracy in classification can occue because of incorrect detection due to merged or really small muscle fibers.

## 1.4 Objectives

- Distinguishing fresh fish from frozen thawed.
- To differentiate the individual species with the help of histology images.
- Extracting features based on the texture of fish histology images.

- Classification using texture features into fresh and frozen.
- Classification using texture features into different fish species.

## 1.5 Project Scope

The range of project is to classify fresh fish from frozen thawed based on the features calculated from histology image dataset of different fish species. The project is going to be helpful for health authorities as well as fisheries department to apply check and balance in order to decrease the fraud sales.The scope of the project can be extended from laboratory use like research purposes up to industrial scale.

# Chapter 2

# Review of Literature

This chapter provides detailed survey of the work done in similar domain. In this chapter, we study the different techniques applied previously and their affects and shortcoming. Histology is known as a simple, rapid and highly accurate method for differentiating texture based images[5]. In literature, different edge detecting and preserving methods like Soble edge detector and cross bilateral filter which help in smoothing the image while preserving edges[17] have been used for preprocessing of image. Many of researchers have used Local binary patterns, Glcm, Gabor filter, Convolutional neural network and wavelet-based approaches[7]. Mostly Support vector machine and KNN were used in researches for the texture based classification which provided results of up to 93% of accuracy[3]. Following are some the methods studied in detail.

## 2.1 Some Preprocessing Techniques

Image preprocessing is done to improve the image such that key image features are enhanced for further processing while suppressing the undesired distortion or noise hence decreasing the degree of unwanted results[13].

### 2.1.1 Cross-bilateral filter

The cross-bilateral filter is a technique that is used to smooth an image while preserving its edges hence enhancing the features present on edges[9]. In our work, as fat and protein on muscle fibers constitute edges or borders hence this technique is useful.

### 2.1.2 Sobel Operator

Sobel operator is used for edge detection where it performs gradient measurement on an image such that areas with high frequency which can correspond to as edges are emphasized[4]. It is slower to compute but large convolution kernel makes it less sensitive to distortion.

## 2.2 Some Feature Extraction Techniques

Feature extraction is used to extract features from a preprocessed histology image of fish. Haralicks texture features are the prominent used features in literature which mainly consists of energy, entropy, homogeneity, contrast etc.[10] There are many methods proposed in different studies for texture based feature extraction, most of which are stated below[2].

### 2.2.1 GLCM

Grey level co-occurrence matrix is one the most widely used method to deal with Haralick texture features. It is a second order statistical method which deals with pixel relationship at certain direction and distance. It gives a measure of the variation in intensity at the pixel of interest[14]. However, strict measures must be practiced in order to obtain high accuracy such as parameters have to be carefully decided otherwise it can lead to false positive and false negative results. This method shows sensitivity to noise and doesn't perform well in border regions of image, also it takes too much time processing an image[12].

### 2.2.2 Local Binary Patterns

Local binary pattern has become one of the widely used textural image descriptor which defines the contrast and spatial structure of image. This method has a high computational complexity but it is insensitive to rotation making it sustainable to the uniform change in grey scale. [2][8].

### 2.2.3 Gabor Filter

Gabor filter is one of the most used methods in texture analysis domain. It can be used to reveal edges and lines in an image as it has showed great accuracy in distinguishing the features and edges of an image specially around a certain direction and frequency[2][16]. It can be used with glcm to minimize the inaccuracy of glcm at edges[12].

### 2.2.4 CNN

Convolution neural network is a deep learning model which has shown great potential in recognizing different objects like edges from multiple images due to its feature learning power. In the analysis of repetitive texture patterns, filter banks make a great tool to accurately distinguish features showing up to 93% accuracy[1].

## 2.3 Some Classifiers

Some known classifiers in the literature working in texture analysis domain are support vector machine and K- nearest neighbor, where in order to train the model, mostly supervised learning is used[2].

### 2.3.1 SVM

Support vector machine is a classifier that can solve linear and non-linear problems by creating a hyperplane which in turn separates the data into classes. It does not perform well with noisy data hence data shall be preprocessed first. It is known to give good results in texture analysis problems where similar type of harralick features are used[11].

### 2.3.2 KNN

The K-nearest neighbor is a supervised learning algorithm which classify new data based on class of previously categorized data. It can be very expensive computationally and also sensitive to noise and outliers. Applying KNN on similar type of problem where features are extracted from Glcm and Gabor filter showed up to 88% of accuracy in results[6].

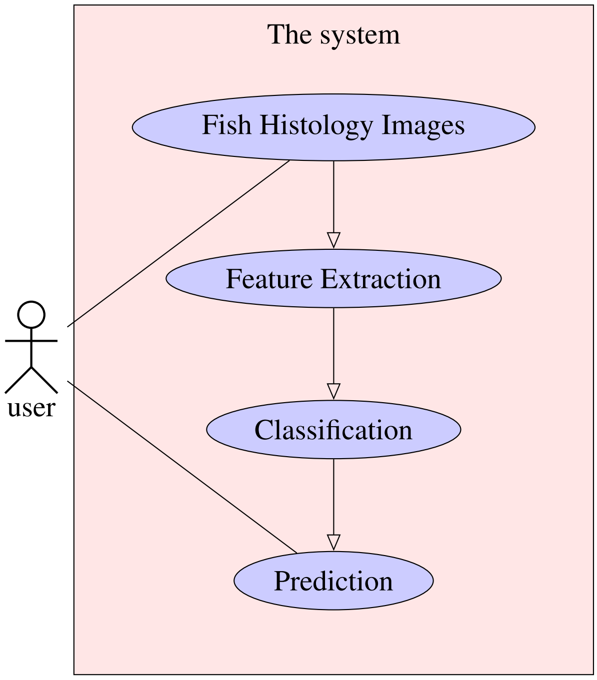
# Chapter 3

# System Analysis and Design

Analysis and design is one of most important phase in system development. In this chapter, we deal with how requirements are gathered and system is designed to convert the requirements into different usecase and sequence diagrams that helps in development phase.Coverting requirements directly into code can cause problems but having turned them into a diagrams and schmeas such that it is helpful in implementation is the major role of system analysis and design. Following usecase and sequence diagram shows system's insight on how it works on abstract level.

## 3.1 Use Case Diagram

The following diagram provides an abstract view of the system.



## 3.2 Sequence Diagram

Sequence diagram describes interaction among classes to carry out the tasks. Main class will call the Dataset to obtain the image dataset, which will be given to Preprocessing class in order to preprocess the images. the preprocessed images will be used by Feature-Extraction class which will use Glcm and Gabor to extract the features and return them. Then, Main class will classify the query image based on features extracted.

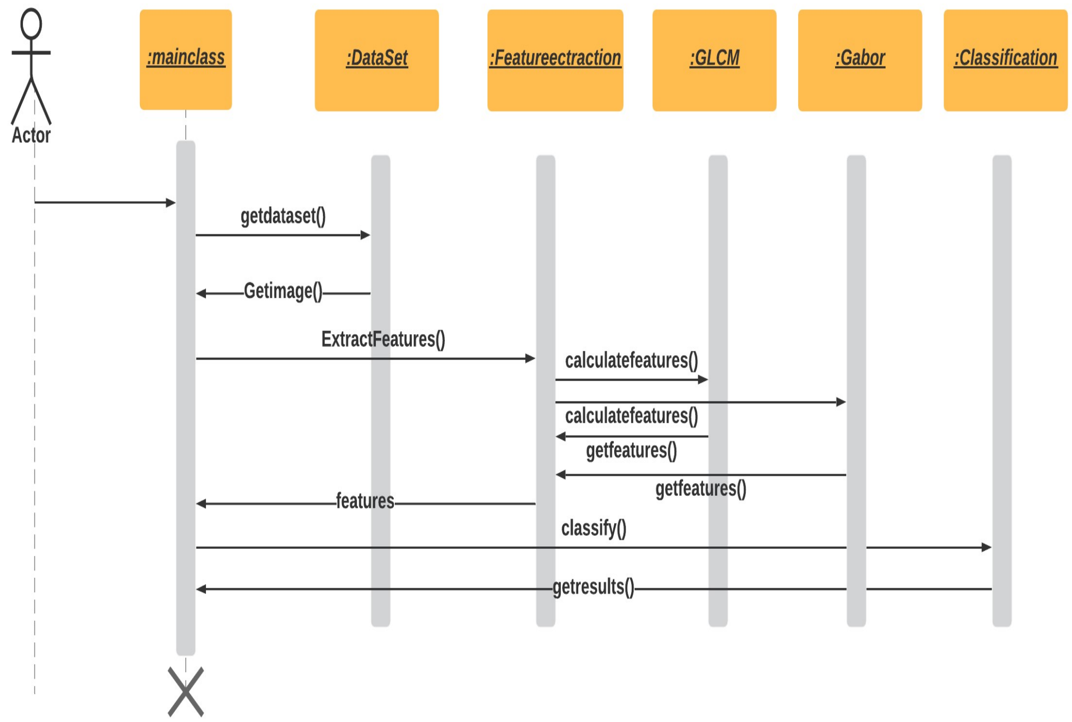



Figure 3.1: Sequence Diagram

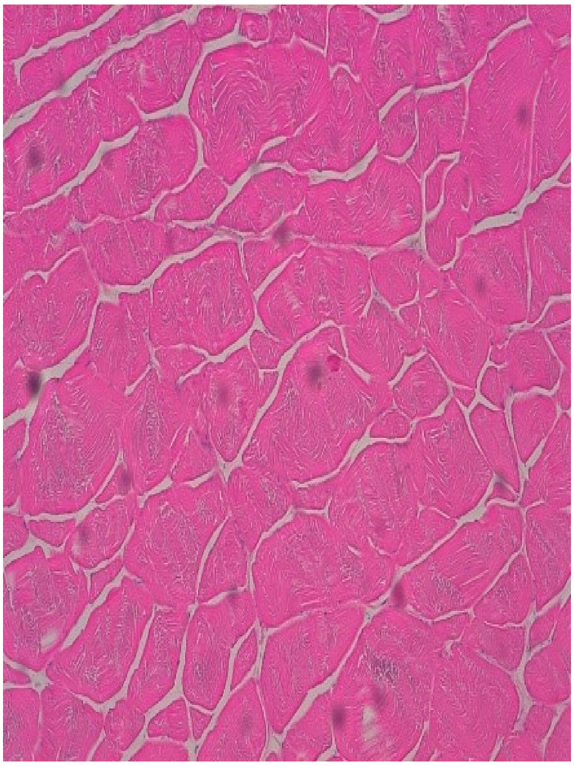
# Chapter 4

# System Implementation

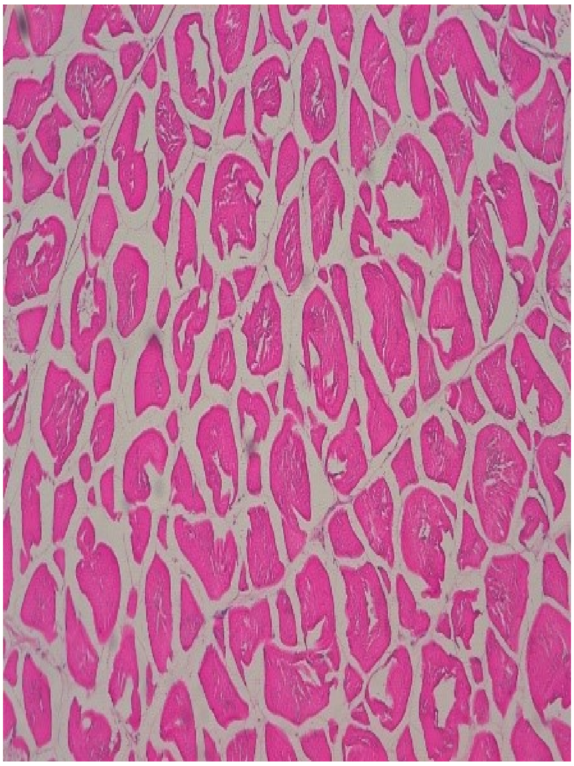
The system is implemented according to the designed flow and diagrams. Firstly, the data set of fish histology images was acquired then some preprocessing techniques were applied to both augment and standardized the image dataset such that it gives better results. Feature extraction methods were chosed to extract texture based features according to review of literature and whatmethods provided better results in past studies. Extracted features were then fed to a classifier for training and testing such that it classify the query image into right class.

## 4.1 Dataset

Image dataset consisted of Orata and Triglia fish specie was acquired which was further divided into fresh and frozen classes. Below are the sample images of both species.

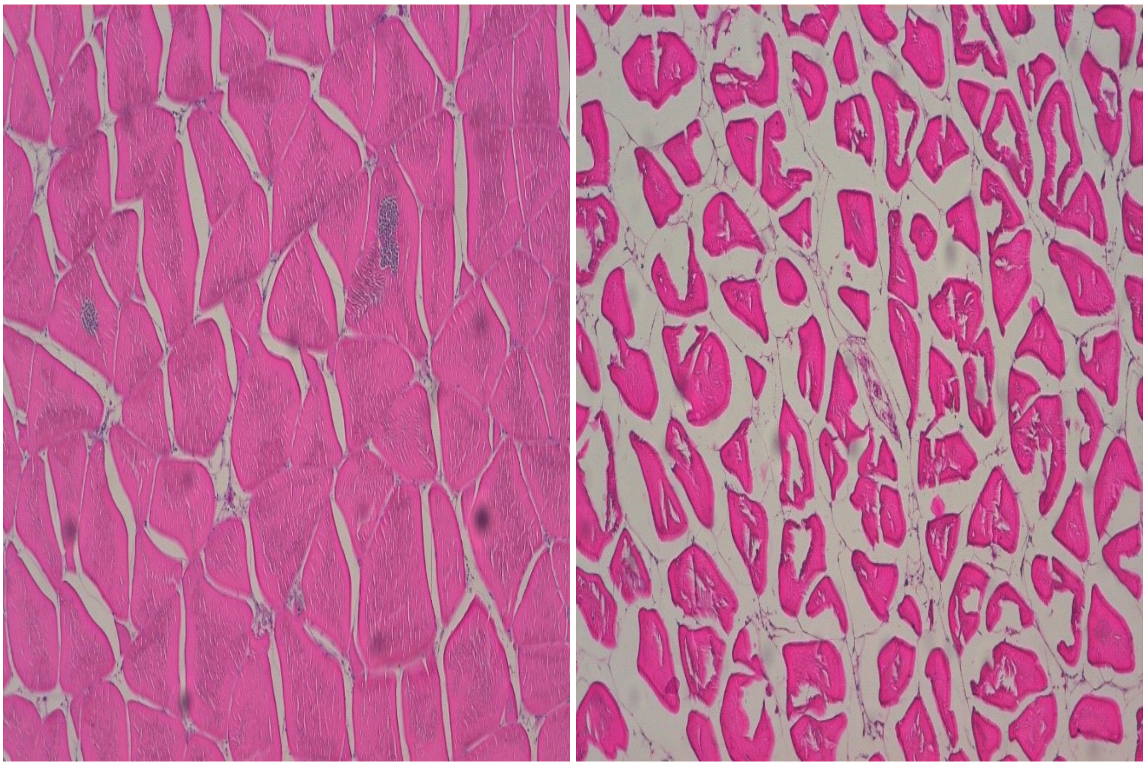


(a) Fresh Orata sample



(b) Frozen Orata sample

Figure 4.1: Fresh and Frozen Orata samples.



(a) Fresh Triglia sample (b) Frozen Triglia sample

Figure 4.2: Fresh and Frozen Orata samples.

## 4.2 PreProcessing

Image dataset is usually not in the standard form to get processed by a method and draw good results, there can be many factors including noise, difference in dimensions, some methods doesn't do well with very high resolution images hence there is a need to convert an image such that it can be processed easily and give good results. Preprocessing is done to standardized the image dataset such that Histology images from provided dataset undergo slicing. These images are then set at fixed resolution 2048x1536.

### 4.2.1 Data Augmentation

Image augmentation is used to create various new data images by applying various techniques on original image in order to increase the dataset for training and testing. Initially the dataset of Fish histology images was only of limited number hence, many augmentation techniques were used to expand the dataset.

#### 4.2.1.1 Noise Injection

Noise is deliberately altering pixels to be different than what they may should have represented. Gaussian noise which is a statistical noise having PDF equal to normal distribution was added to images producing new dataset with images containing noisy data.

#### 4.2.1.2 Shearing

Shearing creates a sort of stretch in the original image. In this technique,image is streched at a certain angle while keeping one axis fixed hence creating sort of a stretch at certain angle giving new image. Various images can be created using one image and multiple shearing angles.

#### 4.2.1.3 Rotation

Random Rotation is an augmentation technique in which image is rotated at certain angles to get new image. By doing so it is also possible that image can lose its shape such that some parts of original image are absent while showing only background at those points. Again it can give various new images based on degree of rotation for one input image.

#### 4.2.1.4 Flip

Flipping is a part of image rotation where it is only flipped across certain axis. Flipping an image (and its annotations) is a deceivingly simple technique that can be used for creating several versions of our images in various orientations. Images were flipped horizontally and vertically such that resolution of newly created images shall remain similar.

#### 4.2.1.5 Brightness

Adjusting image brightness to be randomly brighter and darker can give us a new image. The brightness of the image can be augmented by either randomly darkening images, brightening images, or both.

### 4.2.2 Greyscale Conversion

The images with new resolution undergo Greyscale conversion to obtain grey level intensity values. This done so because pixel values in greyscale range from 0-255 providing vast variety and hence contrast for comparison and processing. In some applications, work can be done more instantly by using only grey level instead of full color as it can cost extra time and space in memory to store and process more pixels other than grey scale. Hence,it is better to use grey level images unless color images are necessary for the work to be done.

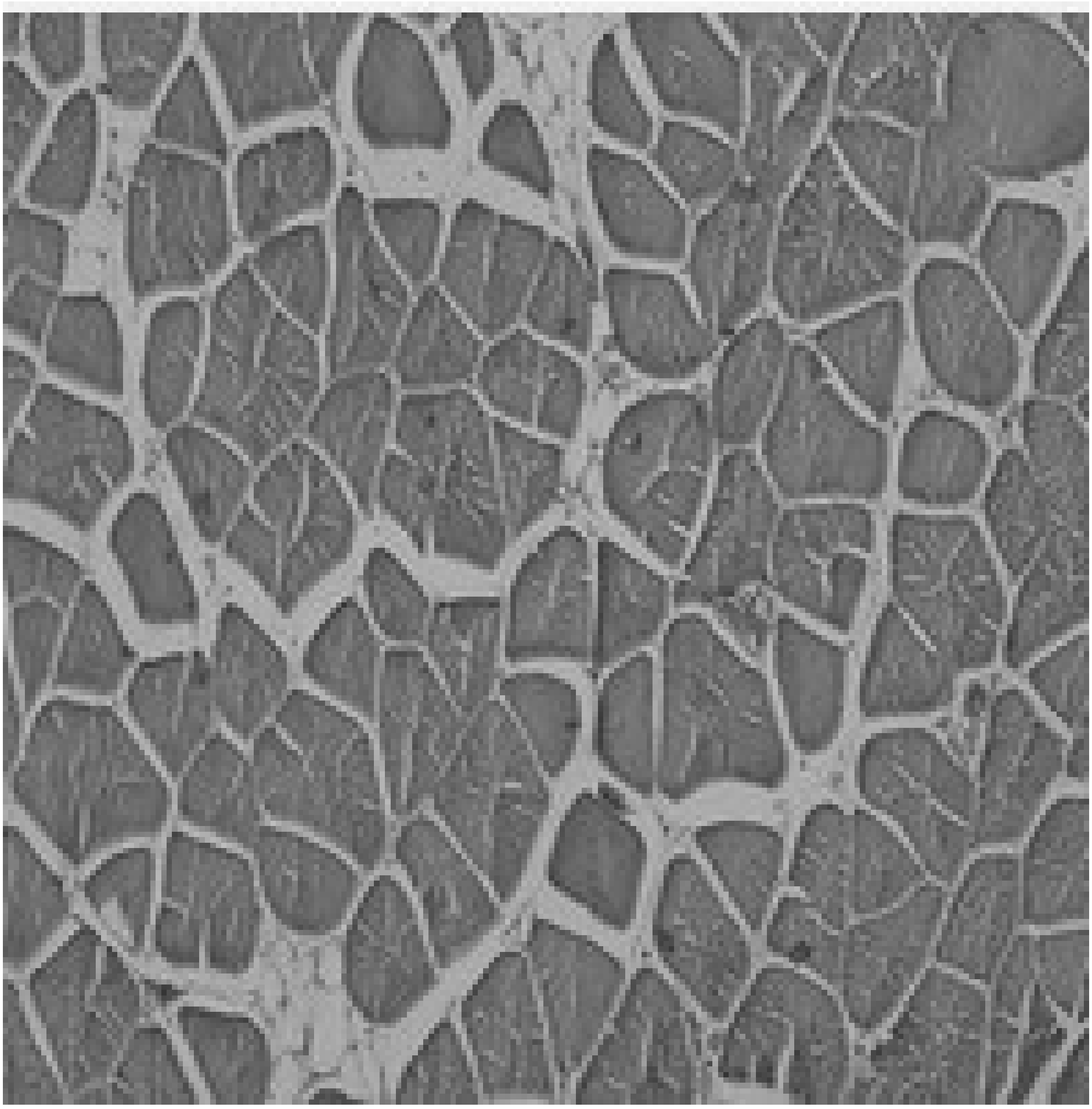


Figure 4.3: Image after Grey Scale Conversion

### 4.2.3 Cross Bilateral Filter

The result of greyscale conversion will undergo cross bilateral filter. It is a non-linear, smoothing filter which replaces the intensity value of each pixel with an assigned average weightage w.r.t to nearby pixels.It smooths the whole image by preserving edges of high intensity.

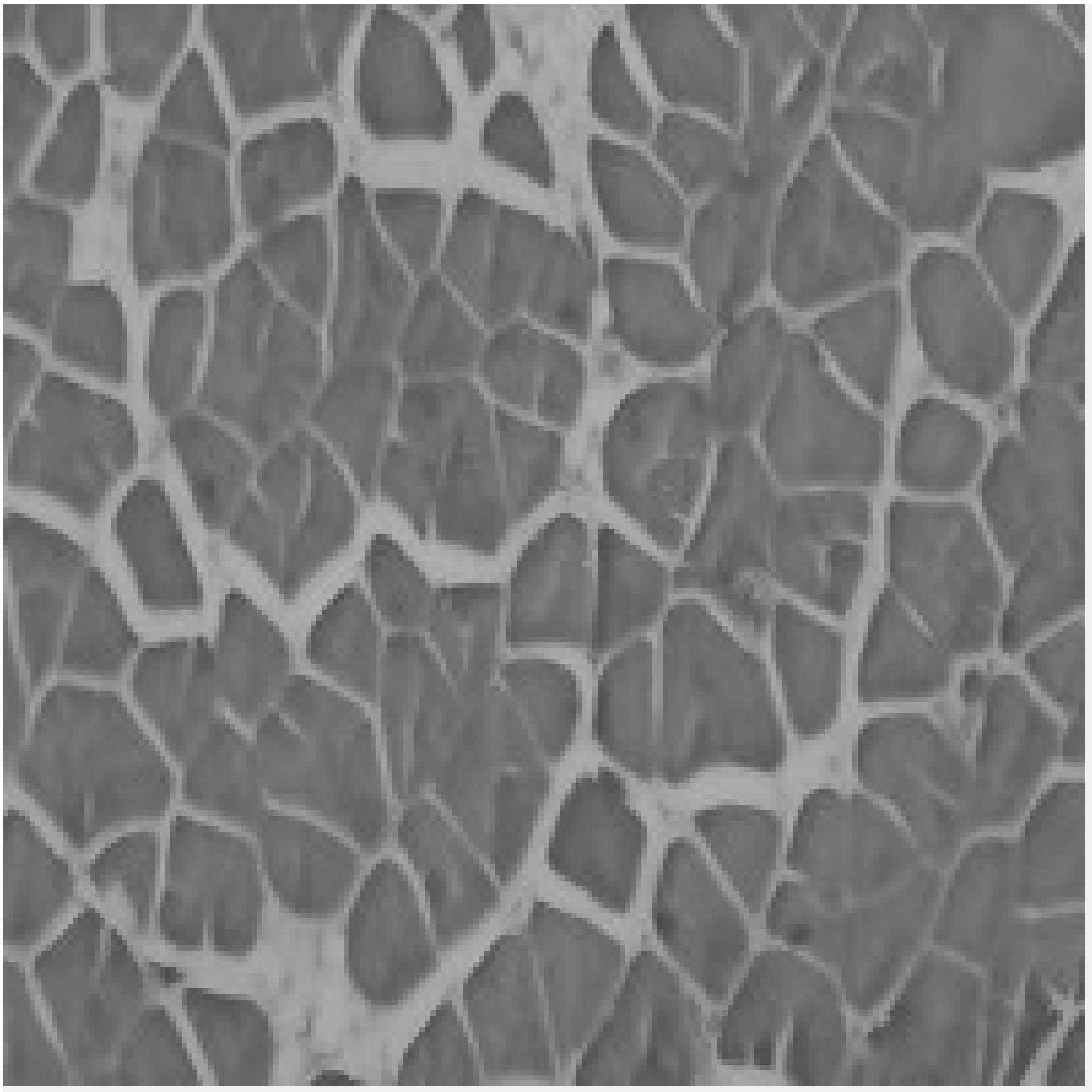


Figure 4.4: Image after cross Bilateral Filter

## 4.3 Feature Extraction

A feature is a distinctive attribute or some quantifiable bit of information in your image which is not present in other objects. It can be anything, a distinct color, some edge, or a line. Features should be distinctive enough to differentiate between different objects. As from the literature review, Harralick features work best for texture based analysis and as Glcm and Gabor filter are known for their accuracy in this domain, hence Grey level co-occurrence matrix and Gabor filter were used to extract features from image dataset.

### 4.3.1 Grey Level Co Occurence Matrix

GLCM is a second order statistical technique used for image texture analysis. It is a statistical method applied on images to examine textures in image and extract features which are later compared for comparison and distinguishing. In this method, an image having different pixel values will produce a co-occurence matrix based on the number of times somepixel values occured together at certain direction and distance that is values in Glcm matrix are count of frequencies of the neighboring pairs of image pixel values. First of all intensity or pixel values of image will be quantized into specific grey levels, such as if we are using 5 bits, 32 grey levels will be used or if we are using 8 bits for quantization then 256 grey levels will be used,It can be simplified as we'll be needing two power bits of grey level using any number of bits. Next on the basis of predefined direction and distance where distance shows how adjacent the pixels will be such that if distance is one, adjacent pixels will be considered which shows neighbor connectivity on a shorter range, similarly if distance is two, pixel after the adjacent one will be considered. We only take four directions that are (x+1, y) or 0 degrees, (x+1, y+1) or 45 degrees, (x, y+1) or 90 degrees, (x-1, y+1) or 135 degrees as other four degrees used after the matrix is made symmetrical. A new square matrix of M*M will be made based on the pixel relationship values where M is the total unique pixel values in the original image. Data will be populated in Glcm matrix based on how frequently two pixels occured together. after populating the Glcm matrix, it's transpose will be added to itself to make it symmetrical. After the Glcm matrix is made symmetrical, it will be normalized by dividing every entry in matrix by the total sum of all entries.

**Grey Level Co Occurrence Features:**

Grey level co-occurence matrix is mostly used to extract Haralick texture features which describes the correlation in intensity or grey level of pixels that are on a certain distance and direction mainly consisting of energy, homogeneity, entropy and contrast etc.

**Contrast:**

Contrast is the measure of local variations in an image where it will be greater as the variation or difference between pixel values increases. From the following formula of contrast, it is obvious that if both pixels i and j are equal that is on the main diagonalline then the result will be zero that is not much contrast or we can say that pixels are identical.

$$\sum_{i,j} |i-j|^2 P(i,j)$$

**Energy:**

Measures uniformity of grey level distribution. It shows how close the values are to the diagonal, as from the following formula value will be greater if pixels are near to the main diagonal.

$$\sum_{i,j} P(i,j)^2$$

**Entropy:**

Entropy measure complexity of an image.Where from the following formula, It is inversely proportional to GLCM energy.It will be very large when pixel relationships will have very small values showing that texture of image is not uniform.

$$-\sum_i \sum_j P(i,j) log P(i,j)$$

**Homogeneity:** Homogeniety unlike contrast shows how simliar the grey level pixel values in an image are or how smooth it is. More smooth the image is,higher the homogeneity will be.

$$\sum_{i,j} \left( \frac{P(i,j)}{1+|i-j|} \right)$$

Following are some example features we extracted from two different classes Congelata and Fresca of Fish. where theta(direction)=45 and d=1;

| Type | Orata Congelata | Orata Fresca |
| --- | --- | --- |
| **Energy** | 0.0023924334534 | 0.01455908799 |
| **Entropy** | 3.0401485829122 | 2.17304160119 |
| **Contrast** | 128.18430875305 | 13.5353051137 |
| **Homogeneity** | 0.2835542218941 | 0.370364921947 |

### 4.3.2 Gabor filter

Gabor filter has emerged as one of the most prominent texture feature extraction method globally. Gabor channels are commonly utilized in surface examination, feature extraction as they allow certain frequencies to pass through while rejecting others making them a band pass filter. Gabor filter is known to give good results around border areas of image hence can be useful in our domain where images are consisted of fat and muscle layers making them edgy. It decomposes an image into various channels which in turn gives different features based on the orientation, filtering criteria, variance and kernel size. It picks the features in an orientation defined such as when on edges, the features get enhanced due to difference in pixel values thus making it more suitable for working on bordered region of images. This methodology breaks down a picture into some responses set given by a Gabor channels bank, that almost covers in a consistently way the spatial-recurrence space. It depends on the picture measurements, and the quantity of portions in a bank of Gabor channels shifts as indicated by the quantity of mixes among frequencies and directions. Gabor channels are applied to pictures essentially similar path as are traditional channels. Gabor filters are designed to extract features at different orientations and scales. This multiresolution sensitivity of Gabor filters may be helpful for extracting useful meaningful features that can characterize underlying texture. In gabor filtering, a convolution part or 2D array of pixel values while each pixel value is assigned a weight is used as a filter. This mask or array is then slide over each pixel of image to perform the convolution operation. Following is the general formula for gabor filter.

$$g(x,y;\lambda,\theta,\sigma,\phi,\gamma) = exp\left[-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right] cos\left[2\pi\frac{x'}{\lambda}+\phi\right]$$

The ouput of gabor filter depends on certain parameters and how they are treated.

- Gabor kernel size is represented as ksize. We have a gabor kernel of size x*y if ksize= (x, y). On varying this parameter, the size of the convolution kernel varies.
- sigma $\sigma$ represents the standard deviation of a certain Gaussian function that is used

by the Gabor filter. This parameter controls the overall size of Gaussian envelope being used in the Gabor kernel.

- Theta $\theta$ is used to control the orientation of the Gabor function. It is one of the most important parameter which shows features in certain orientations such that If theta is 90 degrees it will repond to horizontal postion of gabor function, similarly it will respond to vertical position if changed to zero degree.

- lambda $\lambda$ represents the wavelength of the sinusoidal factor mainly cosine factor of the Gabor kernel. It governs the width of the Gabor function's stripes. Stripes will be thicker if lambda is increased, similarly thinner stripes will be produced by decreaing its value. It is adviced to keep the lambda value smaller in order to overcome undesired impacts.

- gamma $\gamma$ is the spatial aspect ratio. The support will be circular for gamma value equal to one, similarly for the gamma values less than one the support will be elongated in the direction of the parallel function stripes.It is known to control the height of Gabor function.

- Phase offset $\phi$ of the sinusoid that is specified in degrees is the argument in Gabor function. Phase offset's valid values range in real numbers between -180 and 180.

## 4.4 Classification

Image classification organizes data in categories based on various features obtained in feature extraction. Classification as a whole can be divide into training the dataset based on the features and then testing the data or image on the trained model.

### 4.4.1 Support Vector Machine

A common approach used for classification via extracted features in application of Support Vector Machine. This classifier is chosen for our study because in similar type of works, it has shown very good results[11]. It is a supervised learning model that analyze data and can solve linear and non-linear problems. It trains on the provided training set and later distinguishes and categorizes new data by creating a hyperplane on which all the data is mapped. SVM then compares new data with old and categorizes. It is used for multi-label problems and works good on numerical features which are used in our work. We used 5-fold cross validation method as to validate the the model, where using only glcm features, only gabor features and mixed glcm and gabor features.

## 4.5 Graphical User Interface

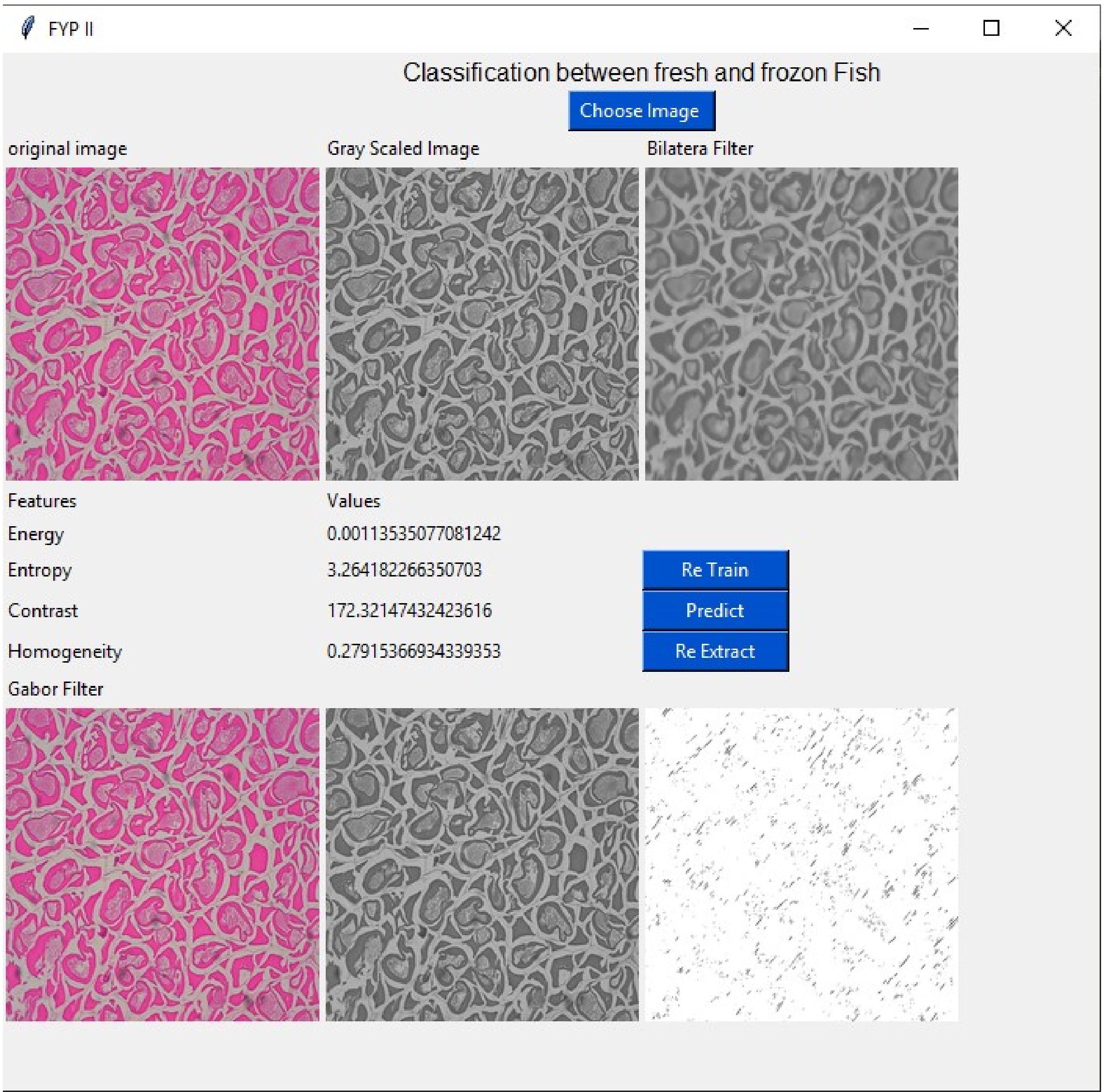



Figure 4.5: Graphical User Interface of Software

# Chapter 5

# Results

The initial dataset was consisted of 300 images which was later augmented to 2400 images using Noise, Brightness, Rotation and shear. Harralick texture features were extracted using Glcm and Gabor filters. We have tested our dataset in two ways, first using a 70/30 split and other is 5-fold cross validation was used on the features extracted using GLCM, Gabor and mixed for training and testing. In 70/30 split, we split the dataset into 70 percent training and 30 percent testing which also produced many false positive results. 5- fold cross validation was used to randomly train and test the dataset divided into 5 groups. First classification was done only on GLCM features but as for different species, different fat to protein ratios were observed which led to false positive and false negative results in some cases which resulted in limiting the overall accuracy of model to 78% as don't give better results on border regions. Similarly, Gabor filter gave 82% accuracy but when we used features from both methods combined, accuracy increased up to 88%. Contrast feature produced differentiating values as the difference between contrast of fresh fish and frozen was greater such as 128.184 for frozen and 13.535 for fresh fish specie. Classification results using different methods are given below.

Following graphs shows the intra specie classification of fish based on 70/30 split and 5-fold cross validation.

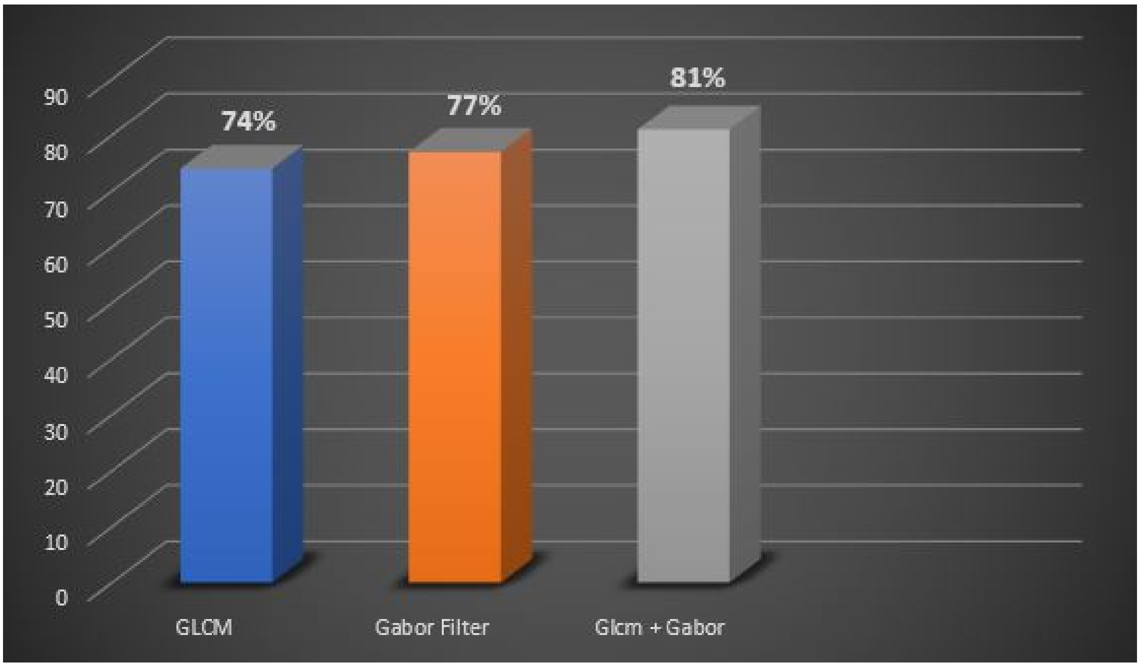



Figure 5.1: Classification results using 70/30 split.

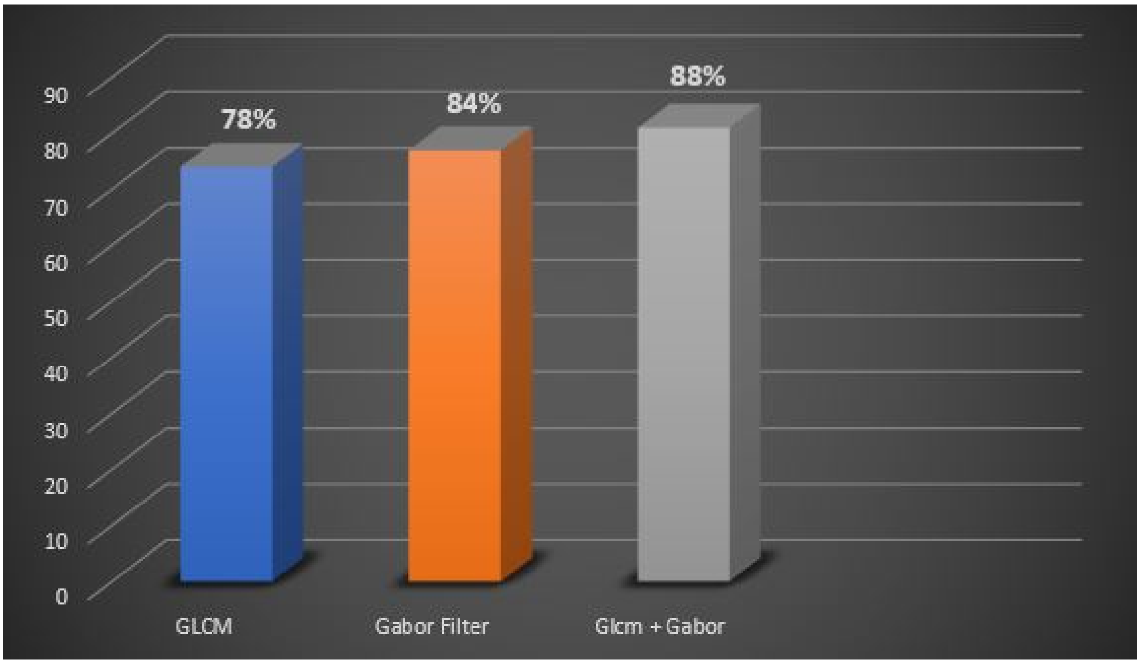



Figure 5.2: Classification results using 5-fold cross validation.

Following are the detailed classification results based on Glcm, Gabor feature extraction and using both 70/30 split and 5-fold cross validation.

| Feature Extractors | Classification | Precision% | Recall% | F1-Score% | Accuracy% |
|---|---|---|---|---|---|
| Glcm | 70/30 Split | 77 | 90 | 83 | 74 |
| Gabor Filter | | 81 | 72 | 88 | 77 |
| Glcm + Gabor Filter | | 84 | 74 | 88 | 81 |
| Glcm | 5-fold cross validation | 81 | 69 | 71 | 78 |
| Gabor Filter | | 89 | 89 | 73 | 84 |
| Glcm + Gabor Filter | | 92 | 89 | 90 | 88 |

# Chapter 6

# Discussion

The results indicate that texture based features work great on histology images of muscle fiber. With the observation of protein to fat ratios and empty spaces within cells, Fresh fish can be distinguished from Frozen thawed. The study correlates the realtionship between fat and protein on muscle fiber to the edges present in histology image, hence constituting that edge enhancing and preserving methods should work fine in this domain. The texture based features such as contrast played a great role in distinguishing fresh fish from frozen based on the local variation in the image. Thses results build on existing evidence of how texture based features using grey level co-occurence matrix and gabor filter can draw good results. Though time and space complexity using grey level co-occurence matrix and getting false positive results did limit the results based on how time consuming the feature extraction using glcm was. Future studies should take into account the time complexity of this method and also the someprocessing technique to help clean the image dataset such that false negative and false positive results shall be avoided.

# Bibliography


[1] Chapter 4 - deep learning in texture analysis and its application to tissue image classification. In Adrien Depeursinge, Omar S. Al-Kadi, and J.Ross Mitchell, editors, *Biomedical Texture Analysis*, pages 95 – 129. Academic Press, 2017.

[2] Laleh Armi and Shervan Fekri Ershad. Texture image analysis and texture classification methods - a review. 2:1–29, 01 2019.

[3] Lucia Ballerini, Robert B. Fisher, Ben Aldridge, and Jonathan Rees. *A Color and Texture Based Hierarchical K-NN Approach to the Classification of Non-melanoma Skin Lesions*, pages 63–86. Springer Netherlands, Dordrecht, 2013.

[4] S. Beeran Kutty, S. Saaidin, P. N. A. Megat Yunus, and S. Abu Hassan. Evaluation of canny and sobel operator for logo edge detection. In *2014 International Symposium on Technology Management and Emerging Technologies*, pages 153–156, 2014.

[5] E. BOZZETTA, M. PEZZOLATO, E. CENCETTI, K. VARELLO, F. ABRAMO, F. MUTINELLI, F. INGRAVALLE, and E. TENEGGI. Histology as a Valid and Reliable Tool To Differentiate Fresh from Frozen-Thawed Fish. *Journal of Food Protection*, 75(8):1536–1541, 08 2012.

[6] Devanur Guru, Y H Kumar, and Manjunath Shantharamu. Texture features and knn in classification of flower images. *International Journal of Computer Applications,Special Issue on RTIPPR*, 1:21 – 29, 01 2010.

[7] A. Humeau-Heurtier. Texture feature extraction methods: A survey. *IEEE Access*, 7:8975–9000, 2019.

[8] Simily Joseph and Kannan Balakrishnan. Local binary patterns, haar wavelet features and haralick texture features for mammogram image classification using artificial neural networks. In David C. Wyld, Michal Wozniak, Nabendu Chaki, Natarajan Meghanathan, and Dhinaharan Nagamalai, editors, *Advances in Computing and Information Technology*, pages 107–114, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[9] Pierre Kornprobst, Jack Tumblin, and Frédo Durand. Bilateral filtering: Theory and applications. *Foundations and Trends in Computer Graphics and Vision*, 4:1–74, 01 2009.

[10] Tommy Löfstedt, Patrik Brynolfsson, Thomas Asklund, Tufve Nyholm, and Anders Garpebring. Gray-level invariant haralick texture features. *PLOS ONE*, 14:e0212110, 02 2019.

[11] Kuan Liu, Bin Lu, and Yaxun Wei. Better image texture recognition based on SVM classification. In Zhiguo Cao, editor, *MIPPR 2013: Pattern Recognition and Computer Vision*, volume 8919, pages 211 – 216. International Society for Optics and Photonics, SPIE, 2013.

[12] Fardin Mirzapour and Hassan Ghassemian. Fast glcm and gabor filters for texture classification of very high resolution remote sensing images. *IJICTR, International Journal of Information and Communication Technology Research*, 7:21–30, 09 2015.

[13] Mark S. Nixon and Alberto S. Aguado. Chapter 4 - low-level feature extraction (including edge detection). In Mark S. Nixon and Alberto S. Aguado, editors, *Feature Extraction and Image Processing for Computer Vision (Third Edition)*, pages 137 – 216. Academic Press, Oxford, third edition edition, 2012.

[14] Annamalai Suresh and K. Shunmuganathan. Image texture classification using gray level co-occurrence matrix based statistical features. *European Journal of Scientific Research*, 75, 04 2012.

[15] Kimberly Warner. Oceana study reveals seafood fraud nationwide. *Oceana*, 11, 01 2013.

[16] Khin Win, Sung Baik, Ran Baik, Sung Ahn, Sang Kim, and Yung Jo. Texture feature extraction and selection for classification of images in a sequence. In Reinhard Klette and Joviša Žunić, editors, *Combinatorial Image Analysis*, pages 750–757, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[17] Y. Zhang, X. Wang, and L. Peng. Text based bilateral filter for color image edge preserving. In *2013 5th IEEE International Conference on Broadband Network Multimedia Technology*, pages 49–52, 2013.