Republic of Iraq

Ministry of Higher Education and Scientific Research

University of Diyala

College of Science

Intelligent Human Age Estimation System Based on Hybrid Techniques

A Dissertation
Submitted to the Department of Computer Science\ College of Science\ University of Diyala in a Partial Fulfillment of the Requirements for the Degree of Master in Computer Science

By
Dalal Adnan Abbas

Supervised By
Dr. Jamal Mustafa Abbas
Assistant Professor

2020 A.C. **1441 A.H.**

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

﴿الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ ۚ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ﴾

صَدقَ ٱللَّـهُ العَظيم

سورة المجادلة

الآية (11) الجزء (28)

Dedication

To…

My dear parents

My all family

My friends

Everyone who loves me sincerely

I produce this work with all my love….

Dalal Adnan

Acknowledgment

First of all, praise is to GOD, the lord of the whole creation, on all the blessing was the help in achieving this thesis to its end.

I wish to express my thanks to my college (college of science), my supervisor, Ass- Prof. Dr. Jamal Mustafa Al-Tuwaijari for supervising this research and for the generosity, patience and continuous guidance throughout the work. It has been my good fortune to have the advice and guidance from him. My thanks to the academic and administrative staff at the Department of the computer sciences.

I would like to thank my family who has supported me throughout and encouraging me to further my studies and help to complete this project, and I wish to express my thanks to my friends.

Dalal Adnan

Linguistic Certification

This is to certify that this thesis entitled **“Intelligent Human Age Estimation System Based on Hybrid Techniques”** *was prepared under my linguistic supervision. It was amended to meet the style of English language.*

Signature :

Name **: Asst. Dr. Ghazwan Mohammad Jaafar**

Date : / /2020

Supervisor's Certification

We certify that this thesis entitled **"Intelligent Human Age Estimation System Based on Hybrid Techniques"** was prepared by **Dalal Adnan Abbas** under my supervisions at the University of Diyala Faculty of Science Department of Computer Science, as a partial fulfillment of the requirements needed to award the degree of **Master of Science in Computer Science**.

(Supervisor)

Signature :

Name : **Asst. Prof. Dr. Jamal Mustafa Abbas**

Date : / / 2020

Approved by University of a Diyala Faculty of Science Department of Computer Science

Signature :

Name : **Asst. Prof. Dr. Taha Mohammed Hassan**

Date : / / 2020

Examination Committee Certification

We certify that we have read the thesis entitled **"Intelligent Human Age Estimation System Based on Hybrid Techniques"** *and as examination committee, examined the student* ***"Dalal Adnan Abbas"*** *in the thesis content and that in our opinion, it is adequate as fulfill the requirement for the Degree of Master in Computer Science at the Computer Science Department, University of Diyala.*

Signature:
Name: Asst. Prof. Dr. Taha Mohammed Hassan **(Chairman)**
Date: / / 2020

Signature:
Name: Asst. Prof. Dr. Methaq Talib Gaata **(Member)**
Date: / / 2020

Signature:
Name: Asst. Prof. Dr. Bashar Talib Hamid **(Member)**
Date: / / 2020

Signature:
Name:Asst. Prof. Dr. Jamal Mustafa Abbas **(Supervisor)**
Date: / / 2020

Approved by the Dean of College Science, University of Diyala

Signature:
Name: Asst. Prof. Dr. Tahseen Hussain Mubarak **(The Dean)**
Date: / / 2020

Abstract

In the past few decades, the age estimation of humans from visible data like video data and still images is one of interesting research because of many prospective applications like the development of intelligent human-machine interfaces and refinement of safety and protection in various sections as convey, security, and medicine. On the other hand, age estimation is still very challenging matter because of aging is influenced by several factors like accidents, facial hair, wrinkle, and skin condition. Lately, many machine learning techniques are used to estimate ages.

In this thesis, a new intelligent system for human age estimation from facial images by using machine learning techniques and two swarm intelligence based algorithms was proposed. The typical structure for the proposed system consists of several phases: pre-processing phase, face detection phase, feature extraction phase, feature selection phase for eliminating redundant features, and classification phase. The proposed system experimented using the aging FG-NET database. The database divided into seven classes which have been turning out that some classes have the same number of features. Therefore, these seven classes combined into three classes based on the number of features it contains to minimize the mean absolute error and increase precision. The face detection phase is performed using the viola – jones algorithm. Feature extraction is utilized using linear discriminate analysis LDA, and feature selection is implemented by using two swarm intelligence based algorithms, firefly and bat algorithms. Finally, two machine learning based algorithms, J48 classifier and decision table are utilized for age estimation.

The obtained results showed that the proposed system has given the precision of 88.77% when using Firefly with the J48 classifier, while achieved a precision of 86.79% when using Firefly with a decision table classifier. Furthermore, a hybrid approach of Firefly and Bat algorithms with a J48 classifier was applied and has given the precision of 90.45%, while the same hybrid approach achieved a precision of 87.05% when use a decision table classifier. Finally, a best result of mean absolute error of 1.14, was obtained by the hybrid approach of Firefly and Bat algorithm with the J48 classifier.

List of Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

Abbreviations	Meaning
2D	Two Dimension
3D	Three Dimension
AAM	Active Appearance Models
AdaBoost	Adaptive Boosting
ANN	Artificial neural networks
BIF	Biologically Inspired Features
CBIR	Content Based Image Retrieval
CNN	Convolutional Neural Network
CSC	Convolutional Sparse Coding
FG-NET	Face and Gesture recognition Research Network
FNR	False Negative Rate
FPR	False Positive Rate
GW	Gabor wavelets
KFA	Kernel Fisher Discriminant Analysis
HCI	Human Computer Interaction
HE	Histogram Equalization
HOG	Histogram of Oriented Gradients
ID-3	Iterative Dichotomiser
LBP	Local Binary Patterns
LDA	Linear Discriminate Analysis
LPQ	Local Phase Quantization
ML	Machine Learning
MRELBP	Median Robust Extended LBP
PCA	Principle Component Analysis
ROI	Region Of Interest
SCM	Skin Color Modeling
SDM	Strategic Decision Making
SDM-LDL	Strategic Decision-Making Label Distribution Learning
STD	Standard Deviation
SVMs	Support Vector Machines
SVR	Support Vector Regression

Chapter One

Chapter One

General Introduction

1.1 Introduction

In the past few years, with the increased focus on video-based security systems and social robotics, researches about the numerical analysis of faces (which includes face detection and identification, face expression recognition, and gender classification) drew the attention in computer vision and pattern identification communities [1]. About those researches, the estimation of a person's age by numerically analyzing their facial image is quite a new subject. The estimation of age by using numerical facial image analysis may be implemented in a wide range of potential implementations, for example, developing intelligent human-machine interfaces and improving the protection and safety of young people in a variety of certain sectors (such as medicine, transport, business, banks, and so on) [2]. It may be highly beneficial for developed video monitoring, business intelligence, demographic statistics gathering, enhancement of search in large data-bases, and patron profiling. The attribute of the age may be utilized in verifying faces and enriching tools which are utilized in the investigations of the police. Generally, automatic age estimate by the machine is beneficial in applications in which the goal is determining a person's are without recognizing him [3].

1.2 Age Estimation

The estimator of age could be utilize a method based on machine learning (ML) for training a framework. It use to obtain features and making predictions of

age for query faces with trained model. In general, estimation of the age may be taken under consideration as an issue of multi-class classification, an issue of regression or both combined [4].

The estimate of age from face images needs a considerable amount of information from input image. Extracting those characteristics is highly important due to the fact that the efficiency of a system of age estimation will be highly dependent on the obtained features quality. A considerable amount of studying the estimation of age was proceeding in the direction of extracting aging features [5]. Methods of classification appear to consider optimal performances in demographic classifications such as gender, age, and ethnicity. Nonetheless, this supposition can't constantly be true. However, this classification could give some sufficient results in a single database [6].

1.2.1 Feature Extraction

The aim of determining face detection is given a spot image. Thus, if there are faces in the image, in which case, return the location of the image and every face's extent. Some of face detection issues are the following:

• **Pose**: A person is not looking directly at the camera every time a picture is taken of him/her. Therefore, more often than not, a human's face will be at an angle with the camera and thus the ratios of the face will be altered.

• **Face expression**: The look of a face is influenced by the face expression of an individual.

• **Occlusions**: Occlusion could be the most complicated issue in facial detection and identification. It indicates the fact that some human face parts be partly occluded by other facial images [7].

• **Presence or absence of features**: Face features like mustaches, spectacles, and beards, may be present or otherwise, and there's large variability of those features in bulk, format, and color.

• **Orientation of the image**: images of faces are different for various rotations around the optical axis of the camera.

• **Image condition**: Image conditions such: distribution of lighting, background, the distance between camera and person, and resolution of the image are important factors. Characteristics of image capturing devices effect images and the appearance of faces [8].

There are numerous matters that are closely related with the detection of faces. A face localization has the aim of determining image index of a single face. This is a simplified issue of detection assuming that an image includes a single face only. The target of feature detection of faces is the detection of the existence and position of some face characteristics such as: nose, eyes, lips, eyebrows, mouth, etc. Also, the proposing of there is merely a single face present in an image. Face recognition performs a comparison of an input image i.e. a probe versus a database i.e. a gallery and returns a match if found. In addition, facial expression recognition, face tracking, and face authentication are worth invoking as common problems. Detection of faces is an initial step in automated systems that is capable of solving the abovementioned issues [9].

1.2.2 Classification

Classification is one of the Data Mining techniques that is using to analyze a specific dataset and takes each instance of it and assigns this instance to a certain class. The classification is using to extract models that accurately define important data classes within the specific dataset. So classification is the process to assign a

class label from dataset whose class label is unknown. There are two classification kinds, which are supervised and the unsupervised [10].

1.2.2.1 Supervised Classification

Supervised classification utilizes spectral signatures that have been extracted from the training samples for classifying the image. Utilizing the toolbar of Image Classification becomes easily creating training samples for the representation of classes that we need to be extracted. This system is also capable of easily creating a file of signatures with the use of the training samples that is afterwards utilized by tools of multi-variate classification for the classification of images [11].

1.2.2.2 Unsupervised Classification

Unsupervised classification finds spectral classes i.e. clusters in multi-band images with no interference from an analyst. This classification type is accomplished via granting access to the tools for the sake of creating clusters, the capability of cluster quality analysis, and access to tools of classification [12].

1.3 Related Works

In the past few years, there are many works have been in the human age estimation area this thesis presents some of them:

- Jhony K. Pontes, Alceu S. Britto, and et al (2015) [13]: They proposed an innovative model for integrating Local Binary Patterns (LBP), Active Appearance Models (AAM), Local Phase Quantization (LPQ), and Gabor wavelets (GW) for the sake of obtaining a feature representation which is highly discriminative and is capable of modeling appearance, shape, skin spots, and wrinkles. Moreover, they have proposed an innovative flexible

method for hierarchical estimation of age, which consists of multi-class Support Vector Machines (SVMs) for the classification of subjects to age groups followed by an Support Vector Regression (SVR) for the estimation of a particular age.

Errors which could occur in the step of classification, as a result of hard borders amongst the classes of ages, have been compensated in estimating particular age via a flexible age ranges overlapping. The efficiency of the suggested method has been measured on FG-NET Aging database and a value of mean absolute error (MAE) = 4.50 was achieved.

- Pablo Pardo García, and Sergio Escalera Guerrero, et al (2015) [14]: A new age estimation facial image database is presented, a study comparing the estimation to both kinds of ages was done by suggesting two age estimation methods use the state of art approaches, one depended on Biologically Inspired Features (BIF) and the second based on Convolutional NN (CNN). The results of the two methods executed are show the superiority of the CNN over the BIF and they too show the hard of the age estimation problem with face images captured in an uncontrolled environment. The method achieves MAE of 7.99.
- Wei Zhao, and Han Wang (2016) [15]: They proposed an Strategic Decision Making Label Distribution Learning algorithm (SDM-LDL) with a set of methods that are specified for various age label distribution types. The experimental results which have been performed on the most common data-base of aging faces, FG NET, showed that it was superior to and more valid compared to all suggested SDM learning algorithms over existing label distribution learning and other single-label learning algorithms for the

estimation of face age. The inner characteristics of SDM-LDL have been additionally investigated with more benefits, the MAE its value was 5.07.

- Xiaolong Wang, Robert Li, and et al (2017) [16]: The authors have adopted a model of convolution map, and proposed a new method for feature learning which has been based on Convolutional Sparse Coding (CSC) that is capable of automatically learning the characterization of aging signatures. In comparison with other common methods of feature learning such as deep convolutional NNs (CNNs), they have verified that the learning method is capable of extracting localized subtle features of aging such as the CNNs, and in addition to that, considerably decrease the size of the model. In addition to that, they have employed the pooling of standard deviation (STD) for summarizing the feature of aging. Ultimately, the obtained features have been fed to a discriminative manifold framework of learning for obtaining low-dimensional representations which are more discriminative and additionally enhance computational efficiency. The method uses FG-NET database and the MAE value was 4.01.
- Soumaya Zaghbani, and et al (2018) [17]: They have presented a new age estimation approach from face images according to auto-encoders. An auto-encoder can be defined as an Artificial neural networks (ANN) which is utilized in unsupervised learning of sufficient coding. It has the aim of learning the representation for a dataset. The aim of their study was exploiting the efficiency of auto-encoders in learning features using a supervised way for the estimation of the age of a user. They have utilized FG-NET database for testing the efficiency of their presented approach. The results of experimentations have shown that the presented approach is robust

and effective via the mean absolute error criteria which showed a value which is equal to 3.75 for FG-NET.

- Shahram Taheri and Önsen Toygar (2019) [18]: have applied an integration of various feature extraction algorithm types to face images for precise estimation of age. Their integration has been accomplished with the use of a 2-level fusion of scores and features by using score-level and feature-level methods of fusion. In this approach, the benefit of the utilization of various feature types like appearance-based, biologically inspired, and texture-based features have been utilized. The feature-level fusion of texture-based and biologically inspired approaches has been integrated to the suggested approach and their combination has been merged with the appearance-based approach with the use of score-level fusion. Experimental results that have been obtained from implementing the method on publicly available FG-NET database have proven the efficiency of the presented approach and the value of the mean absolute error was equal to 4.06.

1.4 Problem Statements

Face features play main role in many applications for civilian and military application. The appearance of a human face is affected significantly by aging. Facial aging affects are mainly assigned to bone growth and movement and skin-related deformities associated with reduction of muscle strength and the presenting of wrinkles. Design and implementation of a system that can predict human age based on facile image by using a computer facility in fast and high degree of accuracy is not easy manner. Age estimation encountered many problems related to typical face image interpretation such as face detection, face recognition, expression and gender recognition. However, the problem of age estimation shows further challenges include: in such cases variances in appearance between

neighboring age groups are trivial, producing difficulties in the age estimation process. In addition, many factors could influence the process of aging comprise, race, genetic traits and gender. Therefore, various approaches are required for age estimation for various groups of subjects. Moreover, accurate systems for age estimation needs appropriate datasets that suitable for training and testing, and include various images showing different ages and covering a wide range of age. These datasets require images captured in the past.

1.5 The aim of Thesis

This thesis is aimed to design and implement a system for human age estimation from face image with high degree of accuracy based on swarm intelligence algorithms (Firefly and Bat) and machine learning techniques by using J48 and Decision Table classification algorithms. The system can accept human face image as an input and extract its feature to classify it for estimate the age based on these features, which are obtained by applying such face detection techniques and algorithms.

1.6 Thesis Layout

In addition to chapter one this thesis consist of four other chapters and as follow is organized of these chapters

Chapter two: this chapter present the theoretical background of the general algorithm and techniques that used in this thesis.

Chapter three: this chapter contains the details of proposed designed system and its related algorithm and explained its steps.

Chapter four: this chapter includes the obtained results after applying the proposed system on used data set and discussion of the obtained results.

Chapter five: this chapter gives the conclusions of this work and the recommendation for future work.

Chapter Two

Chapter Two

Theoretical Background

2.1 Introduction

Estimating the age of human is of high importance for face images classification. For research purposes, definitions of fundamental terms are given. Concerning the present study, the estimation of age can be defined as the individual's age according to the biometric features, particularly according to the 2D images of the face. The face characteristic points can be specified as the standard reference point on the individual's face used via scientists for the purpose of recognizing the face of an individual, estimating the person's age, or in such situation. Morphology alone is considered to be a format's study. Variations in the face texture are specified as the face variations associated with muscle as well as skin flexibility. Thus, studying the shapes of the skull and the face is defined as the craniofacial morphology. The appearance and the structure of individuals are impacted in a various way due to the aging operation. The occurred changes are associated to face texture and craniofacial morphology. Specific craniofacial morphology features are seen in individuals of a specific age and change over the operation of aging. The changes in the texture of the skin are typically happen in puberty [19].

There are some obstacles related to estimating the age. One of the major obstacles is the age estimation morals, especially in estimating the age of children. The estimation of age can be specified as the fix of age groups or the age of individuals. There are many ways to set the age of individuals, yet the presented work is focused on estimating the age according to 2D images of individuals [20].

Many age classes are exist as illustrated in the following example:

- Estimated age can be specified as the age that is defined through the computer according to the appearance of individuals.
- Chronological age can be specified as how many years the person has lived.
- Perceived age can be specified through other individuals who define the age according to the appearance of the individual.
- Appearance can be specified as the information that is associated to age, described via the appearance of individuals.

Appearance age is usually very close to the actual or chronological age. The objective of age estimation is that estimated age is as close to appearance age as possible [21].

2.2 Image Preprocessing

Preprocessing of images can be defined as the image operations at the lowest abstraction levels, the output and the input are both intensity images. Such images can be considered as the same type as original data which has been captured through the use of a camera. The intensity images typically represented through the matrix of image function values, brightness. The main goal of preprocessing is enhancing the image data which overcomes the reluctant distortions or improving certain features of the image. They are high importance for additional processing, even though that the image geometric transformations (translation, scaling, and rotation) are categorized among preprocessing approaches here. That because comparable methods are applied. Many approaches are used for image data preprocessing. In this thesis, we use two methods for purpose RGB to gray converter and histogram equalization [22].

2.2.1 Converting RGB to Gray-Scale

Transmutation of the colored images to gray-scale images demand extra awareness on the color images. The pixel in colored images is defined as a mix of 3 colors (Red, Green, and Blue), which is abbreviated to RGB. Representing the color values of RGB is carried out in 3D (XYZ) axis, state via features of lightness, chroma, and hue. The quality of the colored images depends on the color that is delineated via the number of bits that digital devices can support. Usually, colored images are represented through 24 bits, representing 8 bit for every from red, green and blue. A grey-scale image is representing by using 8 bit. The gray-scale image lightness pixel value extent from 0 to 255 [23].

Given that we have a colored image with RGB components. The provided approaches to convert colored images to gray-scale ones, simple averaging, de-saturation, decomposition, single color channel. In this thesis, weighted average method used as illustrated in the following equation [24].

$$Gray = 0.2989\,R + 0.5870G + 0.1140B \qquad (2.1)$$

2.2.2 Histogram Equalization (HE)

Histogram Equalization can be defined as an extremely common process for modifying the image contrast. The main idea is to map the gray levels according to the probability distribution that are related to disparity gray levels. The dynamic domain related to the histogram of the image is flattened and dilated through HE which lead to total contrast improvement [25], which will allow the regions related to low local contrasts to gain high contrasts. HE applies this through efficient pervasion out the more frequently intensity values. Images with foregrounds and backgrounds which both luminous and deep are useful in this process. Especially, such a process could lead to best views of the bone structures in x-rays images, and

efficiently describing images which are down or up uncovered. The major benefit of the process is fully simple to grasp the approach and invertible operator [26]. Figure (2.1) showes gray-scale image and its gray histogram then shows the equalized image and its histogram of the equalized image. Consider a discrete grayscale image {x} and let ni be the number of occurrences of gray level i. The cumulative distribution function is illustrated in equation (2.2).

$$\text{Cdf}(x)=\sum_{i=1}^{x} \text{px}(i) \tag{2.2}$$

Where x: gray value, and

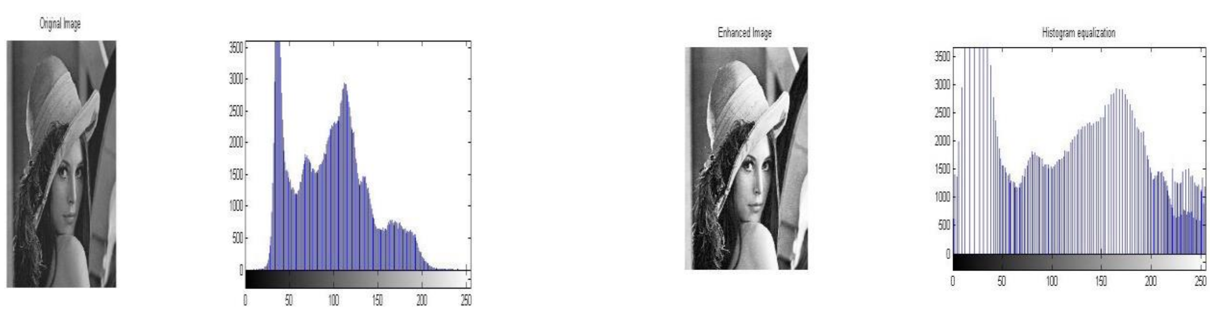
p: the image's histogram for pixel value i.

The general histogram equalization formula is illustrated in equation (2.3):

$$\text{Eh (j)} = round\left(\left(\frac{cdf(j)- cdf(x)_{min}}{E*F-cdf(x)_{min}}\right) * (\text{L} - 1)\right) \tag{2.3}$$

Where $cdf(x)_{min}$ can be defined as the smallest value related to cumulative distribution function is 1. And E* F: can be defined as the numbers of columns and rows in an image. Also

L is the applied Gray levels = 256.



(a) (b)

Figure (2.1): (a): Gray Scale Image (original) and Its Gray Histogram

(b): Equalized Image and Histogram of Equalized Image [27].

The histogram equalization process was implementation by applying histogram equalization algorithm (2.1).

Algorithm (2.1): Histogram Equalization Algorithm [28]
Input: image, //number of gray level from 0 to 255.
Output: enhance image.
Begin Step1: Get size of the gray image. Size of image: E x F, gray level from 0 to 255, build a matrix **G** of size 256 and start it with 0. Step2: Build image histogram by scanning every pixel of image and increasing the related element of matrix. G [gray-val (pix)] = G [gray-val (pix)] + 1 Step3: Get accumulative distribution function (cdf (x)) by using equation (2.2) CH[0] = H[0] CH[i] = CH [i-1] + H[i], i=1, 2, 3 ...255. Step4: Calculate new value by general histogram equalization formula. By using equation (2.3). Step5: Create new image by replacing original gray values with the new gray values. NewImg[x][y] = T[OldImg[x][y]] End

2.3 Face Detection

Human face detection is high importance research field in various applications, like automatic authorization, Human Computer Interaction (HCI),

video conference, content-based image retrieval (CBIR), and etc. Face detection problem could be defined as specifying if the image has a human face. In the case of the existence of faces in the image, it will return the locations of these faces in the image, in spite of the conditions of lighting and the position [29].

Recently, the field of face recognition has gained much interest. Face recognition has a lot of applications in regular access control systems and computer vision communications. One of the major steps in face recognition is the face detection. Yet, the face detection cannot be considered as obvious-cut because it has a lot of differences regarding the image look, like facial appearances, the orientation of the image, pose variations (non-front, front), occlusion, and illuminating situations [30].

Face detection can be considered as the midst of all the facial analysis, such as facial expression recognition, face feature detection, face localization, face verification, and face recognition. Furthermore, it can be considered as a major approach for other applications like HCI, video conference, and the CBIR.

The major goal of face detection is identifying the existence of faces in images. If such faces exists in the image, it will return the extent as well as the location for each face in the image [31]. The face detection methods as shown in Figure (2.2).

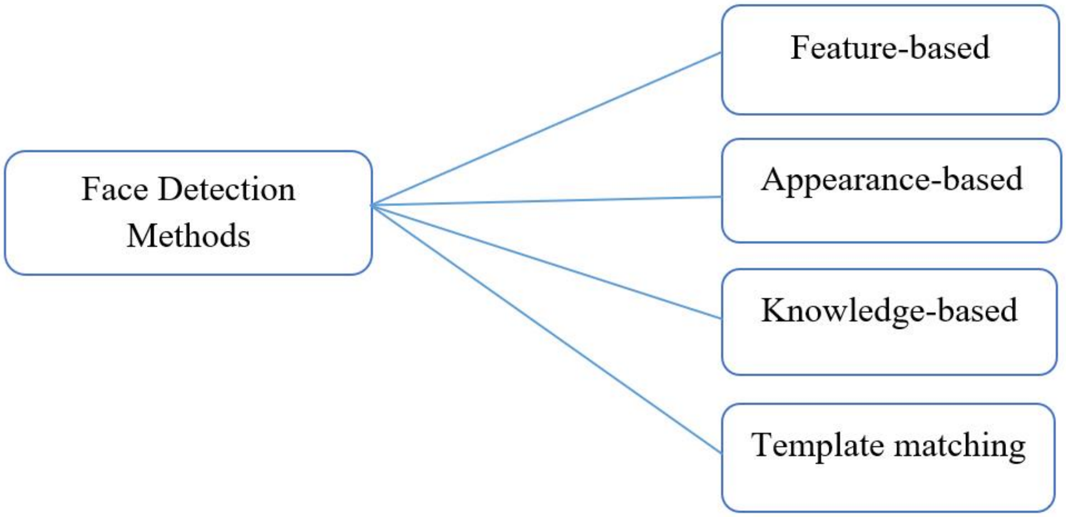


Figure (2.2): General Face Detection Methods [32]

2.3.1 Feature Based Face Detection

Estimating the feature's detection area like mouth, nose, and eyes has significantly improved the precision of detection. This method, identify the faces through the use of set of Haar-like features as in Viola-Jones boosting algorithm. Eyes appearance for the region will be reduced via supposing the expected position of the eye to be at the face's upper part. Eye detection is applied through the use of Haar-like features cascade. It does specify rectangular regions that contain the eyes. Assuming the ROI of the eyes, an algorithm will be created for the purpose of specifying the pupil of the eye through taking the hue information related to the image of the eye [33].

Hue image is threshold and contour is detecting in threshold image. The contour centroid will be detected as the pupil of the eye. After that, Haar-like features are utilized for detecting the nose. After identifying the center of the eyes, in addition to the nose position, a method depends on facial geometry for mouth position estimation. Algorithm will be created for specifying the corner points of the

lips, that are defined as efficient features to track the movement of the lips. Finally, the nostrils will be detected from the ROI of nose through obtaining the threshold of gray nose image and acquiring the threshold image contours [34].

2.3.2 Appearance Based Face Detection

It depends on the collection of delegate training face images for knowing face models. Generally, appearance-based approaches have displayed optimum implementation in comparison to other approaches.

Such approach specifies the search space and improve the rate of face detection. One of the major face detection methods used for videos and images is the Skin Color Modeling (SCM). However, feature selection can be defined as an approach of high importance for even more effective template matching performance with regard to time and detection rate [35].

2.3.3 Knowledge Based Face Detection

According to rules collection, on the basis of human knowledge, and for detecting faces, for example, there will be two eyes in the face, a mouth and a nose within specific position and distance relative to each other.

There are many common principles in the individual intelligence for face specifying like detecting face features containing two uniform eyes, ears, nose and mouth. These operations are optimized according to the face specifying rule in the individual brain. These are a portion of rule-based operations. The relationships among features are specified by their space and proportional positions. Optimizing these ways in different conditions seems hard because not all conditions are able to be accounted [36].

2.3.4 Template Matching

Applying parameterized or pre-defined templates of the face for specifying and detecting faces, through estimating correlation values between input image and the template.

Some patterns of the face are stored in different positions and the connection of input images. These patterns are applied as a standard for face confirmation. For a certain input image, the connection with popular templates is calculated for face outline, eyes, nose, and mouth individually. Yet, multiple scales, multiple resolutions and deformable form and sub forms have been submitted to cope with deformability and scale invariance [37].

2.4 Viola-Jones Face Detection

In 2001, Viola and Jones developed this algorithm [38], since that day it has been very important approach in face detection. This method improved the precision, and majorly the speed, letting real-time applications which have the ability of processing fifteen images in each second. Training method referred as boosting. It will be trained by using a big set of simple, computationally efficient features for the purpose of just the optimum candidates have the ability to pass through the whole cascade of classifiers. Since its development, there has been a lot of advancements, especially in state of the un-constrained face detection [38].

This algorithm has four steps, which are:

1. Haar Feature Selection.

2. Integral Image.

3. Adaboost Training.

4. Cascading Classifiers.

A simple illustration of the above stages

2.4.1 Haar Feature Selection

There are certain similar features related to the faces of individuals. Through applying Haar Features, such regularities can be matched.

Certain features which are common in the faces of individuals are:

• The area of the upper-cheeks is lighter than the area of the eyes.

• Nose bridge area has more brightness than the eyes.

Composition of properties forming matchable facial features:

• Sizes and location: mouth, eyes, bridge of nose.

• Value: The oriented gradients of the pixel intensities.

Rectangle features:

• Value can be defined as the difference of the sum of (pixels in black areas) and the sum of (pixels in the white areas).

• Three types: two-, three-, four- rectangles, Viola-Jones applied 2-rectangle features. See the Figure (2.3) that shown Haar-like features in Viola and Jones.

• For instance: brightness difference between black and white rectangles over certain region [39].

Each Haar-like feature can be calculated as the addition of six values from I:

$$f = -I(y_1, z_1) + I(y_2, z_2) + 2I(y_3, z_3) - 2I(y_4, z_4) - I(y_5, z_5) + I(y_6, z_6) \quad (2.4)$$

f is pixel value of area, I_1 - I_6 points in the integral image points 1 - 6 at the value.

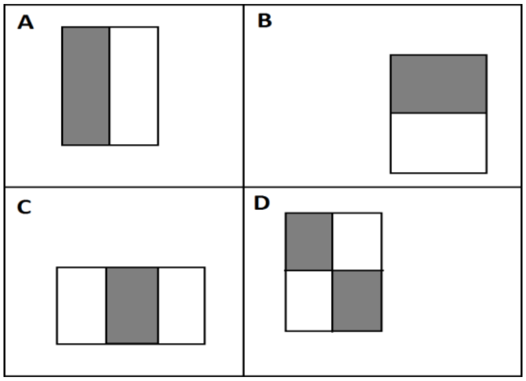


Figure (2.3): Haar Like Features in Viola and Jones [40]

2.4.2 Integral Image

Representation of the image is referred to as the integral image evaluate the rectangular features in a constant time, that is provided them with significant advantage of speed over more advanced alternative features. Since each rectangular area of the features always have adjacency to minimum one other rectangle, it will follow that any 2-rectangle feature could be evaluated in 6 array references, any 3-rectangle feature in 8, and any 4-rectangle feature in 9.

Integral image can be considered as double cumulative sum related to certain image, along column and row dimensions. Each one of the points in integral image are corresponding to the sum of original image upto that point [41], show in Figure (2.4). The integral image at location x, y includes the addition of the pixels at top and left of $x^{\prime}$,$y^{\prime}$ consist o

$$ii(x, y) = \sum_{\acute{x} \leq x, \acute{y} \leq y} i(\acute{x}, \acute{y}) \qquad (2.5)$$

Where *ii (x, y)* is the integral image at location x, y and $i(\acute{x}, \acute{y})$ is the pixel value of original image.

The addition of pixels in rectangle ABCD in Figure (2.4) can be computed with only four values from the integral image.

$$\sum_{(x,y)\in ABCD} i(x,y) = ii(D) + ii(A) - ii(B) - ii(C) \tag{2.6}$$

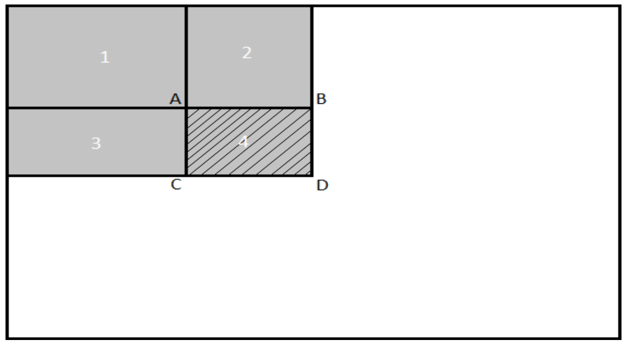


Figure (2.4): Summation of Gray Values [41]

2.4.3 Adaboost Training

Speed in which the features could be estimated doesn't sufficiently compensate for their number. For instance, in standard 24 x 24 pixel sub window. They will be overall (M = 162,336) likely features. It might be excessively costly to estimate all of them in the case of image testing. Therefore, the system of object detection applies different learning algorithm adaptive boosting to both chose the optimum features and for training the classifiers which apply them. Such algorithm construct "strong" classifier as linear combination regarding the weighted simple "weak" classifier [42]. A weak classifier is defined as

$$h(x,f,p,\theta) = \begin{cases} 1 & p * f(x) < T_\theta \\ 0 & otherwise \end{cases} \tag{2.7}$$

Where

x: is a sub window of a base precision of 24*24 pixels.

P: is the parity

f: Denotes the feature value

T_θ: is the threshold

Threshold is decided a face or a not-face.

$$H(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t * h_t(x) \geq \gamma_t \\ 0 & otherwise \end{cases} \quad (2.8)$$

Where H( ): is strong classifier

x: is a sub window of a base precision of 24*24 pixels

α_t: $\log \frac{1}{\beta_t}$

γ_t: is chosen such that all positive training samples are correctly classified.

2.4.4 Cascading Classifiers

• Usually, just 0.01 percent related to all the sub-windows are defined as positive faces.

• There is equal computation time that is spent on all of the sub-windows

• Most time should be spent just on the possibly positive sub-windows.

• Simple two-feature classifier have the ability of achieving nearly 100 percent detection rate with fifty percent false positive rate.

• Classifiers have the ability of acting as first layer of series for the purpose of filtering out the majority of the negative windows.

• Second layer with 10 features have the ability of tackling the harder negative-windows that survived first layer, etc.

• More effective detection rate is achieved via cascades of increasingly more difficult classifiers. Estimation regarding strong classifier which is created through learning procedure could be achieved in quick way. It is not adequately fast for running in real-time. Therefore, strong classifier will be organized in cascades for the purpose of complexity. All the successive classifiers are trained just on the chosen samples that are passing through previous classifiers. If at any of the stages

in cascades, the classifier reject sub-window that is studied, no additional processing will be achieved and keep searching succeeding sub-window. Thus, the cascade has the form of degenerate tree. With regard to faces, 1st cascade classifier will be referred to as the attention operator that applies just two features for the purpose of achieving false negative rate (FNR). The impact of such classifier is reduced the amount of time taken to half [43].

2.5 Feature Extraction

It's a process of the extraction of relevant data from face image. Where such data has to be useful for subsequent to recognizing the item with adequate error rate. In this thesis, these will be applicable technique is the Linear Discriminant Analysis, which is a standard approach for data assorting and dimensionality decreasing [44].

2.5.1 Linear Discriminates Analysis (LDA)

An approach is applied in machine learning, pattern recognition, and statistics for the purpose of obtaining a linear set of features which separate or define at least two object's classes or events. The set has resulted that could be applied as linear classifier or, generally, the for dimensionality reduction prior to more classification [45]. Such other approaches are used applications in which it is not possible to suppose that independent variables are typically distributed. The basic supposition is related to LDA approach. The LDA is related to factor analysis and the Principle Component Analysis (PCA) in a manner of the two search variable's linear combinations effectively describe the data [46].

Linear Discriminates Analysis operates in the case when measurements implemented on the independent variables for each one of the observations are considered to be continuous quantities. In the case of handling categorical

independent variables, equivalent method is the discriminant correspondence analysis.

Considering observations set referred to as measurements, variables, and attributes. For each one of the samples of events or objects with identified N class. Such sample's set is referred to training set. After that, classification problem will search for optimum predictor for N class that is related to any of the samples of distribution (not essentially from training set) specifying just the observation [47].

This is the first stage which calculates average vectors li, (i=1, 2, 3) classes:

$$\text{Ii}=\frac{1}{\text{ni}}\sum_{\text{i}=1}^{\text{n}} \boldsymbol{x}\ \text{i} \qquad (2.9)$$

where I is mean, $\boldsymbol{x}$ is a sample (i.e. row), and $\boldsymbol{n}$ is the total number of samples with a given class.

The next stage is to compute the scatter matrix including within–class scatter matrix (S_w), class-covariance matrices ($\sum$ j), and between-class scatter matrix (S_b).

$$S_w = \sum_{\text{j}=1}^{\text{c}} \sum_{\text{i}=1}^{\text{Nj}} (I_i^j - \bar{\text{I}}_j) \qquad (2.10)$$

where

I_i^j : can be defined as i^{Th} sample of class j,

$\bar{\text{I}}_j$: can be defined as the mean of class j,

c: can be defined as the number of classes, and

N_j: can be defined as the number of the samples in class j.

$$\sum \text{j}= \frac{1}{\text{N}} \sum_{\text{i}=1}^{\text{n}}(\text{x}_i - \bar{\text{I}}_j)(\text{x}_i - \bar{\text{I}}_j)^{\text{T}} \qquad (2.11)$$

$$S_b= \sum_{\text{j}=1}^{\text{c}} (\bar{\text{I}}_j - \bar{\text{I}}) (\bar{\text{I}}_j - \bar{\text{I}})^{T} \qquad (2.12)$$

where

$\bar{\text{I}}$: representing the mean of all classes, and

$\bar{\text{I}}_{\boldsymbol{j}}$: representing the mean of class j.

The goal is to maximize between classes measure S_b while minimize within classes measure S_w.

Popularize eigenvalue for array ${S_w}^{-1} S_b$, then checking that the eigenvector and eigenvalue calculation.

$$A\ V = \lambda\ V \tag{2.13}$$

where

$$A = {S_w}^{-1} S_b$$

V is denote the eigenvector, and λ is denote the eigenvalue.

The eigenvectors sorted through reducing eigenvalues, and choosing k eigenvectors with largest eigenvalues thus the W matrix is construct. The last stage has used the matrix W, which just calculated for transforming the samples to the new sub-space using the following formula [48].

$$\arg(w) = \frac{w^T * S_{b*}\, w}{w^T * S_{w*}\, w} \tag{2.14}$$

where w^T is transform W matrix.

The linear discriminant analysis is called LDA - Fisher Faces Feature Algorithm and is illustrated in Algorithm (2.2).

Algorithm (2.2): LDA - Fisher Faces Feature Algorithm [49]
Input: face image (cropping and resize face image)
Output: fisher face feature vectors
Begin Step1: Read transform face images. Step2: Store transform face image in matrix. Step3: Compute the d-dimensional mean vectors by using Equation (2.9) Step4: Compute the scatter matrices. Compute within –class scatter matrix (S_w) by using Equation (2.10)

Compute the class-covariance by using Equation (2.11)

Compute between –class scatter matrix (S_b) by using Equation (2.12)

Step5: Solving the generalized eigen value problem for matrix by Equation (2.13)

Step6: Choosing linear discriminates for the new property sub-space.

a. Sorting the eigenvectors via decreasing eigenvalues.
b. Selecting *k* Eigen vectors with the biggest eigenvalues.

Step7: Transforming the samples to the new sub-space by using Equation (2.14)

Step8: Return feature vectors.

End

2.6 Feature Selection Using Firefly Algorithm

This algorithm is categorizing as swarm intelligence approach, meta-heuristic and nature-inspired. It had been advanced by Yang in the year of 2008 by the stimulation of distinct fireflies' behaviors. In reality, the firefly population shows characteristic luminary flashing works as pulling partners, communications, and predator risk warnings [50]. Inspired by these actions, Yang had devised this way according to assuming that the fireflies were all unisexual, meaning all of them could attract one another and this attraction is directly proportional to how bright the individual is [51]. So, the brightest fireflies attracts the other fireflies to moves toward them. Aside to that in case where none of the fireflies is brighter than a specific firefly, it moves in a random manner.

This algorithm is implemented with three parameters, those parameters can be summarized as: randomization, absorption, and attractiveness. The latter has been based on the intensity of light between two fireflies and can be characterized using exponential functions. In the case where this parameter given the value of 0,

it occurs to random gait which corresponds to the parameter of randomizing, specified with the principle of Gauss distribution as the generation of a number in the range of [0,1] [52]. However, parameters of absorption have an impact on attractiveness parameters value like the change from 0 to ∞. And, in case of the convergence to ∞, fireflies movement appears as an arbitrary walk.

The flash of the firefly's primary objective is to jobs as a signal system for attracting other fireflies. This firefly algorithm was formulated by Xin-She Yang with the following presumption:

1. Unisexual are all the fireflies, that is why every firefly is attracted to others.
2. Brightness has immediate proportionality with attractiveness, and for pair of two fireflies, the one with less brightness will be attracted to the greater brightness and with the raise in the distance, there is a reduction in the distance.
3. If no brighter firefly is found than the given one, a motion will be random.

The objective function has correlated with brighter [53].

The major updating formula for the pair of fireflies ui and uj is

$$u_i^{T+1} = u_i^T + \beta \exp\left(-\gamma D_{ij}^2\right)\left(u_j^T - u_i^T\right) + \alpha_T \epsilon_T \tag{2.15}$$

Where ui is fireflies, β is the attractiveness of u_j , γ light absorption coefficient, D is distance, T is counter, α is parameter for the step length of the random movement and ϵ is a random vector from uniform distribution with values between 0 and 1[54].

The firefly algorithm was implementation by applying algorithm (2.3).

Algorithm (2.3): Feature Selection Using Firefly Algorithm [54]
Input: features
Output: optimized features

Begin
Step1: Objective function f(u), u= $(x_1, x_2, \ldots, x_d)$.
Step2: Initialize the firefly population ui (i=1,2,…,n).
Step3: Define light absorption coefficient γ
Step4: While T< MaxGeneration
Step5: for i=1: n (all n fireflies)
Step6: for j=1:i
light intensity li at ui is determined by f(ui)
if (Ii < Ij)
Move firefly j towards i v in all direction via Equation (2.15)
else
Move firefly i randomly
end if
Step7: Attractiveness changes with distance r via exp[-γ D^2]
Step8: Determine new solutions and revise light intensity
Step9: end for j
Step10: end for i
Step11: Rank the fireflies and find the best solution
Step12: end While
End

2.7 Feature Selection Using Bat Algorithm

Machine learning is majorly applied for learning knowledge from the experimental data. The aim of machine learning is creating a model which have the

ability of being used in weather prediction, data classification, and so on. Presently, increasing number of real-world data-sets with large numbers of features, that referred to as high-dimensional data-sets, are created. The learning tasks in the high-dimensional data-sets is very complex due to the difficulty of learning and learned model's lower popularization performance. Actually, many features in the high-dimensional data-sets are redundant and unrelated. Obviously, such features are not suitable for learning model, which could have bad impact on the performance of the model [55].

Bats are magical animals and their capacity for searching prey through echolocation. It has been the main focus for a lot of studies. Echo-location acts in a similar way as the sonar: bats, majorly micro-bats, emit short and loud pulses of sound, then it waits until it hit into the object, following fraction of time, echo will return to the ear of the bat. Therefore, the bat can estimate the distance between its location and the object. Also, echo-location might enables bats to differentiate hurdles from one prey in total darkness [56].

According to the bat's behavior, a novel meta-heuristic algorithm, referred to as the Bat Algorithm had been suggested via Yang in 2011 [57]. This approach is modeled through imitating the bat's behavior when trying to search for prey. For the purpose of modelling the algorithm, Yang had idealized certain rules in the following way:

(1) The bat applies its echo-location ability for the purpose of sensing the distance, also they can "know" the difference between hurdles and one prey.
(2) Bat b_i fly in a random way with velocity v_i at the position x_i with fixed frequency f min, varying wave-length λ and loud to search for prey. They can automatically fix the wave-length (or frequency) related to their emitted

pulses and alter the rate related to the pulse emission r [0, 1], according to their target proximity;

(3) Even though the loudness could be different in several ways it is supposed that the loudness changes from a positive large value A_0 to a minimum constant value A_{min} [58].

Velocity of i^{th} solution is updated by using equation (2.16).

$$V_i^{k+1} = V_i^k + (S_i^k - S_j^c) * F_i \tag{2.16}$$

Where V_i^{k+1} represents updated velocity of i^{th} solution, V_i^k denotes current velocity of i^{th} solution, S_i^k denotes the current best solution, and S_j^c denotes the current i^{th} solution.

The frequency of each solution is updated by the following equation

$$F_i = F_{min} + (F_{max} - F_{min}) * \beta \tag{2.17}$$

Bat algorithm can be defined as a major bio-inspired algorithm that works according to echolocation characteristics of bats and they are generate a sound pulse in the range of 20Hz to 150 Hz. This algorithm is enhanced by the integration of Chi-Square feature selection for selecting the best features in a random manner [59]. Bat algorithm is illustrated in (2.4).

Algorithm (2.4): Feature Selection Using Bat Algorithm [60]
Input: features
Output: optimized features
Begin Step1: Objective function f(x), x= (x_1,…,x_i) Step2: Initialize the bat population x_i (i=1,2,…,n) and velocity V_i

Step3: Define pulse frequency F_i at x_i

Step4: Initialize pulse rates r_i and the loudness A_i

Step5: for k=1: Max number of iterations

Generate new solutions by adjusting frequency, and updating velocities locations via the equations (2.16).

if (rand > r_i)

Select a solution among the best solutions

Generate a local solution around the selected best solution

end if

Step6: Generate a new solution by flying randomly

if (rand < A_i & F (x_i) < F (x_{best}))

Accept the new solutions

Increase r_i and reduce A_i

end if

Step7: Rank the bats and find the current best x_{best}

Step8: end for

End

2.8 J48 Classification Algorithm

C4.5 defined as a decision tree learner whose is an upgrade of ID-3 and is implemented in WEKA as J48 using Java.

Decision trees categorizing the approach that starts at the top and moves down to the point of reaching a leaf node is the decision tree. That leaf node's value results in the projected output for the sample. At each one of the nodes, a testing is performed on an attribute and on offshoots from the node which

corresponds to values which the attribute can be equal to. When the sample arrives at one of the nodes, the branch is taken according to its value for attribute which is under testing at that node. A decision tree is shown in Figure (2.5) [61].

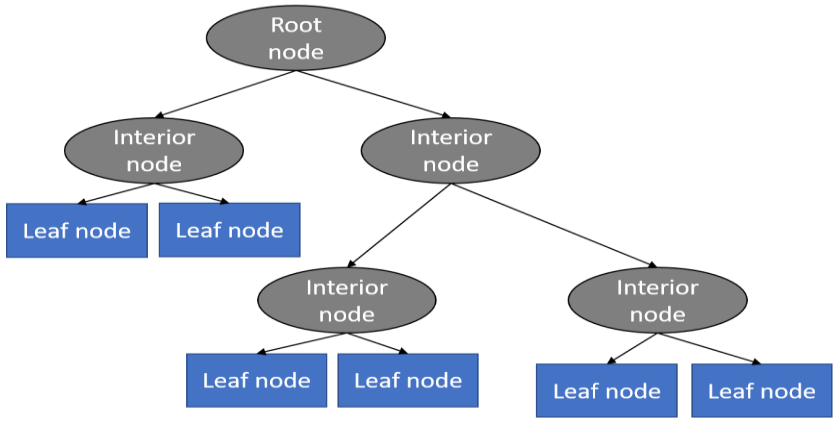


Figure (2.5): A Decision Tree Structure

ID-3 constructs a decision tree according to the training samples collection which is given to that tree. It follows a greedy top-down method to build this tree, beginning by creating a root node. At each one of the nodes, the attribute which can optimally classify each training instance which has arrived at that node is chosen as testing feature. At the node, only the attributes that haven't been utilized for classifying at other nodes above it in that tree are the ones that are considered [62]. For the sake of selecting the optimal feature at the node, the gains of information for every one of the attributes are computed and the feature that has the maximum information gain is chosen. Information gain for attributes which are characterized as entropy reduction which results from dividing the samples which are based on the values that have been obtained by the feature. The gain of information for a feature A at one of the nodes is computed with the use of the following equation:

$$InformationGain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \left(\frac{|S_v|}{|S|} Entropy(S) \right) \quad (2.18)$$

where S represents the group of samples at S and |S| represents the cardinality of nodes, S_v represents a sub-set of S, for which A is equal to a value v, and set S entropy is computes with the use of the following equation:

$$Entropy(S) = \sum_{i=1}^{numclasses} - p_i \log_2 p_i \tag{2.19}$$

where p_i represents the Probability of samples in S which have the value of the i^{-th} class as attributes of the outputs [63].

Expansions have been inserted in the main algorithm of ID-3 for (1) the transaction with the attributes of the continuous values, (2) the transaction with samples having attribute values which are missing and (3) preventing over-fitting of data [64]. J48 classifier algorithm is illustrated in (2.5).

Algorithm (2.5): J48 Classifier Algorithm [65]
Input: data set.
Output: Decision tree (Decision rule)
Begin Step1: Read features. Step2: Checking if algorithm satisfies the criteria of termination. Step3: Compute information-theoretical criteria for each attribute. Step4: Select the optimal attribute according to the information-theoretical criteria. Step5: Creating a decision node according to the optimal attribute in Step 4. Step6: Inducing the data-set according to the newly produced decision node in Step 5. Step7: For every sub-dataset in Step 6, call J48 algorithm to obtain a sub-tree.

Step8: Attaching the tree produced from Step 7 to the decision node in Step 5.

Step 9: Return tree.

End

2.9 Decision Table Classification Algorithm

This is a precise approach for the numerical predictions from the decision trees and it's an ordered group of rules of "If-Then" type, which can potentially be of a higher level of compactness and as a result more comprehensible compared to decision trees. Selecting for exploring decision tables due to the fact that they are less complicated and less compute-intensive compared to decision tree-based approaches [66]. Decision tables are a compact tabular notation for describing a sets of decision rules. Decision tables consist of four quadrants, which respectively describe the relevant conditions, entries of conditions, the relevant actions and the entries of actions. The decision rules are encoded in the columns of the decision table covering both condition and action entries. Figure (2.6) shows an overview of this general decision table structure [67].

	Rule1 Rule2….
Condition stub	Condition entries
Action stub	Action entries

Figure (2.6): The General Decision Table Structure

Decision tables represent two views, which are, the decision tables of limited entries, and the decision tables of extended entries. Every condition of decision

tables of limited entries include binary values that can either be false (F) or true (T), while each condition of decision tables with extended entries includes multiple values [68]. Decision table classifier algorithm is illustrated in (2.6).

Algorithm (2.6): Decision Table Classifier Algorithm [69]
Input: Decision Table, Condition Attributes , Decision Attributes.
Output: Decision rule
Begin Step1: Identify Conditions and Values Find the data attribute each condition tests and all of the attribute's values. Step2: Compute Max Number of Rules Multiply the number of values for each condition data attribute by each other. Step3: Identify Possible Actions Determine each independent action to be taken for the decision or policy. Step4: Enter All Possible Rules Fill in the values of the condition data attributes in each numbered rule column. Step5: Define Actions for each Rule For each rule, mark the appropriate actions with an X decision table. Step6: Verify the Policy Review completed decision table with end users. Step7: Simplify the Table Eliminate and/or consolidate rules to reduce the number of columns. End

2.10 FG-NET Aging Database

The FG-NET was generated as part of the Project FG-NET (Face and Gesture Recognition Network) [70]. The FG-NET contains 1002 images from 82 different subjects with ages ranging between 0 to 69 years old subjects. However, ages between 0 to 40 years are the most populated in the database. With the exception of images showing individuals at more recent ages, for which digital images were available, in most cases FG-NET images were collected by scanning photographs of subjects found in personal collections [71]. Finally, a certain of images will be displayed within a database FG-NET in Figure (2.7).

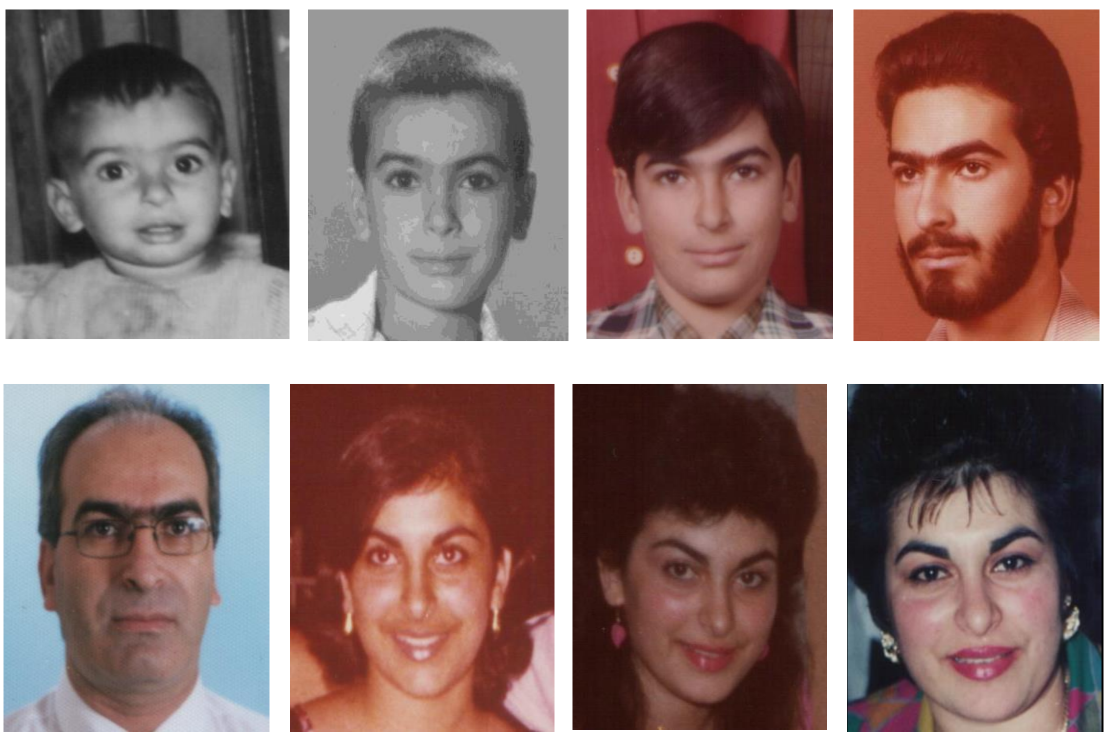


Figure (2.7): FG-NET Images

2.11 Evaluation of Measurements of Classification

Classification algorithms evaluation can be considered as a major issue in data mining methods. The majorly applied tools to investigate the results related to

the classification algorithms are learning curves, receiver operating curves (ROC) and the confusion matrix. Confusion matrix show number of incorrect and correct predictions carried out via model in comparison to true classes in test data.

True Positive rate (TP) is a portion of positive cases that classified as positive and are similar to the Recall.

True negative rate (TN) is a portion of negative cases that correctly classified as negative.

False-positive rate (FP) is a portion of negative cases which predicted as positive.

False-negative rate (FN) is a portion of positive cases which incorrectly labeled as negative.

F-measure is a measurement which merges sensitivity and precision is the harmonic average of the two parameters. It's computed utilizing equations (2.20 to 2.25) [72]:

Many measurements are used to evaluate the used algorithm which can be calculated as shown in the following equations:

True positive rate = diagonal element/ summation of relevant row (2.20)

True negative rate = non-diagonal element/ summation of relevant row (2.21)

Precision = TP/ (TP + FP) (2.22)

Recall = TP / (TP + FN) (2.23)

F-measure = 2*(Precision*recall / Precision+ recall) (2.24)

$$MAE=1/N \sum_{i=1}^{N} |\hat{}yi - yi| \quad (2.25)$$

Where a vector ˆy of N predictions and the vector y of N actual observed target values.

Chapter Three

Chapter Three

Intelligent Human Age Estimation System Based on Hybrid Techniques

3.1 Introduction

Age estimated system is a new challenge for researchers. The system aims to guess the age of the human via image by extracting features and classify it to estimate the age of the human-based on these features.

In this chapter, section (3.2) introduces the dataset and section (3.3) introduces the details of design and the used techniques for the proposed system.

3.2 Database

The proposed system uses the FG-NET database which is summarized on the table (3.1).

Table (3.1) The FG-NET Database Summarization

Element	**DETAILS**
No. of faces images	720
No. of persons	82 different individuals
Age scope	3 years to 50 years

The data which is used in this thesis is grouped into seven classes. Each class has its very own age scope which is appeared in Table (3.2). The estimation

of age will allude to the yield of every order procedure done which will identify to which class this information has belonged.

Table (3.2): Classes Summarizations

class number	**Class scope**
Class 1	3 to 7 years
Class 2	8 to 13 years
Class 3	14 to 19 years
Class 4	20 to 25 years
Class 5	26 to 30 years
Class 6	31 to 40 years
Class 7	41 to 50 years

The extracted feature for each class consolidates since it has a similar number of features which prompted bunch the classes into three classes which are shown in Table (3.3).

Table (3.3) Final Classes Summarization

Class number	**Age scope**
Class 1	3-7 and 26-30 years
Class 2	8-13, 14-19 and 20-25 years.
Class 3	31- 40 and 41-50 years

3.3 Proposed System

The entry of the suggested system is the face image of the human that the system need to guess their age based on the face features. Figure (3.1) illustrates the human age estimation system architecture which is suggested in this thesis.

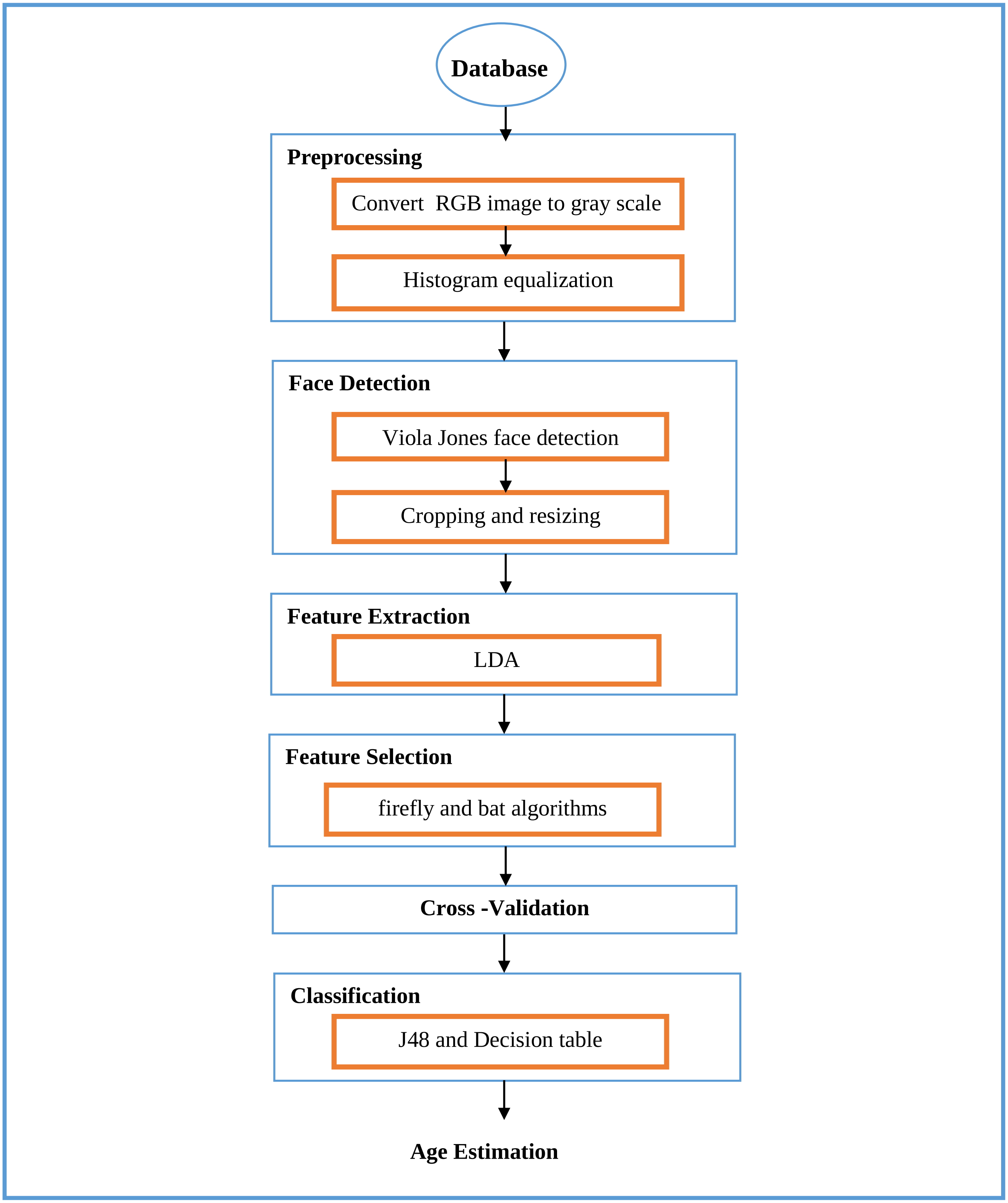


Figure (3.1): Age estimation System Architecture

The main step in such systems is face detection step where the human face image is detected cropped and any other details may discard, this detected face from the view will be used for feature extraction, the age estimation system will work on the data that result from face detection.

Before starting the face detection algorithm which will crop the human face in the image proper pre-processing may be applied to the group of data, the resulted data will enter the face detection algorithm after reducing the amount of information needed, which will lead to extract only the strong data from it to use it later.

- The detected face will be entered to the feature extraction algorithm which will result in a unique feature from each face that is detected, the number of the features in this step is huge and need to be reduced without reducing the quality of the way that features describe the image (face image) some feature selection algorithm is used to improve the results of the proposed system as will discussed later in chapter four.
- **Preprocessing Phase:** Image processed usually contains huge information. Each image which is colored image, contain a lot of information to process. Preprocessing is used for the reduction of the data amount that needs processing and to remove the noise that may affect the extraction of features later. In this thesis simple RGB to gray scale convertor is used and histogram equalization.
- **Face Detection Phase**: Face detection aim is to extract the region of interest which is person face in the proposed system. This detection is used for further processing by extract features and classifies them. The algorithm that is used in this thesis for face detection is viola

Jones algorithm of facial detection. The image is cropped and resized to unique the whole data which is used in the system.

- **Feature Extraction Phase:** The amount of data that need to be processed are still huge and the elimination of the weak data is done via extract the correlated data using feature extraction algorithms, these feature will be used later for classifications, in this proposed system the applied algorithm is LDA feature extraction.
- **Feature Selection Phase:** The number of the obtained features from the previous step is huge and it needs to be reduced which will be done by using a firefly and bat swarm techniques algorithm.
- **Classifications Phase**: The age period is classified to different classes. These classes are represent the age periods and times. Each feature belong to certain age will be classified in the system via Machine learning classification algorithms and the used techniques are J48 classifier and decision table classifier. Finally, Age estimation which is the prediction of human age is done according to the classification step which will verify the age group of each person.

3.3.1 Preprocessing Phase

In this phase, a preprocessing step is applied to the entered image. This phase of the proposed system consists of the following two internal steps:

- **RGB to GrayScale**

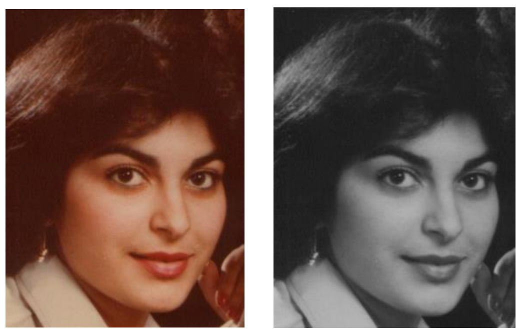
The RGB image converts to grayscale by using proper equation, the main equations used in the proposed system to convert the images is:

Gray RGB color representation has equal values of each of red, green and blue:

R = G = B

For every one of the image pixels with values of (R, G, and B), implements equation (2.1).

The result is the gray level value. Figure (3.2) shows the original image and the gray image.



(a) (b)

Figure (3.2): a: Original Face Image, b: Gray Scaled Image

- **Histogram Equalization**

This technique has been utilized in computer vision to improve contrast in the images, this approach typically maximizes global image contrast in the case where its usable data can be characterized by values of close contrast. This permits the lower local contrast areas for gaining more contrast. Figure (3.3) illustrates the image after applying equalization. Histogram equalization algorithm is implemented as illustrated in algorithm (2.1).

Figure (3.3): Histogram Equalization

3.3.2 Face Detection Phase

This phase in the proposed system consists of two steps as explained in algorithm (3.1).

3.3.2.1 Face Detection Using Viola Jones

Viola Jones is an approach which is based on machine learning technique, in which a cascade function is trained from a massive number of negative and positive images. After that, it is utilized for detecting objects in other images.

There are four internal steps of viola Jones algorithm applied to face images, these steps are:

- **Haar Feature Selection:** since all faces of humans have some similar characteristics, those regularities could be matched with the use of the Haar Features.
- **Integral Image:** that will evaluates rectangular features in constant time.
- **Adaboost Learning:** which choose the optimal characteristics and perform training of the classifiers using them.
- **Cascade Classifier:** all features are grouped to a number of stages, every one of which has a particular number of features.

The results of applying viola Jones algorithm is shown in Figure (3.4)

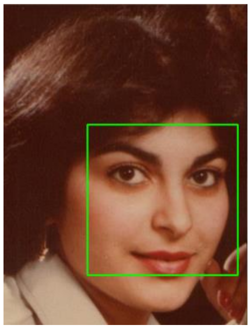


Figure (3.4): The Image after applying Viola-Jones Algorithm

3.3.2.2 Cropping and Resizing

In order to adopt a uniform size for images and to keep only the part of the face which will be used for further processing and for feature extraction which will be applied to this part of face image, the image will be cropped and resized. The cropped image is with size 200*200 and is shown in Figure (3.5).

Figure (3.5): Cropped and Resized Image

The reconstruction of objects is an important step in face detection algorithm to guarantee that the face image is saved and keep all the facial details of the person. Face detection algorithm is illustrated in algorithm (3.1).

Algorithm (3.1): Face Detection Phase Algorithm
Input: equalized gray image Output: face detected
Begin Step1: apply viola jones to detect faces **1. Haar Feature Selection** Value = summation of (pixels in the black area) - summation of (pixels in the white area) Calculate f by using equation (2.4) **2. Creating an Integral Image** Calculate the integral image *ii (x, y)* by using equation (2.5) **3. Adaboost Training** Calculate the weak classifier by using equation (2.7) Calculate the strong classifier by using equation (2.8) 4. **Cascading Classifiers** discard unpromising regions of the image and the result is Z where z is the region of interest (face) Step2: Cropping and resizing Cropping: the image is cropped to Z which is the face image only. Resizing: to unique the size of images the size turns to 200x200. End

3.3.3 Feature Extraction Using Linear Discriminates Analysis (LDA)

Using LDA feature extraction in proposed system will led to reduce the dimension of data and keep the information which will result to speed up the

training and classification process. LDA algorithm is implemented as explained in algorithm (2.2).

3.3.4 Feature Selection Phase

In this phase, two algorithms are used to optimize the features that extracted from image by using FIREFLY and BAT algorithms. Figure (3.6) shows the feature selection phase.

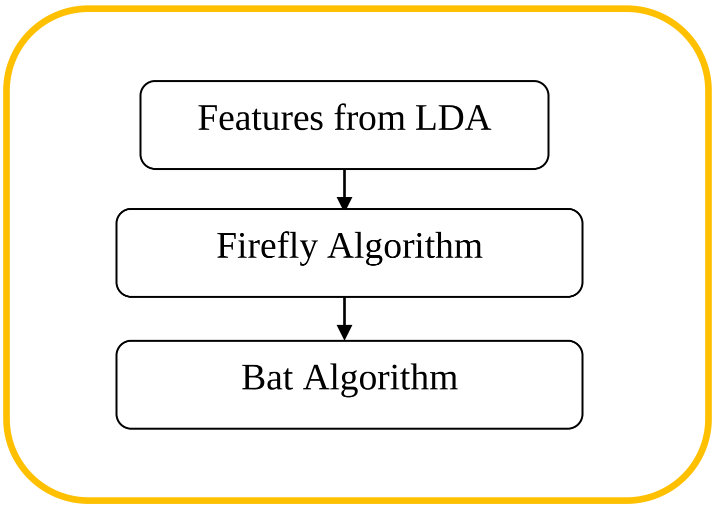


Figure (3.6): Feature Selection phase

The choice of feature optimization algorithm which is done by the system in this thesis is not mandatory in other systems. So, if no feature selection algorithm is chosen then the system will ignore this phase and classify the whole features obtained from the face detection phase. In our system two of these algorithms is chosen, then the number of obtained features will be minimized and only the strong features will be kept. This will reduce the computation time of the system which needs to classify these features and to improve the result of the classification phase since only the strong features are kept and the weak features are ignored. The features obtained from dataset is organized as groups of data that built up

hierarchy to construct the data which are only relevant to same class and other data moved to another classes.

3.3.4.1 Feature Selection Using Firefly Algorithm

Firefly algorithm is one of the feature selection algorithms which uses to solve the degree of optimality (number of features to be selected or ignored).

This algorithm works on finding the attractiveness features which are prepared by the previous phase looking for brightness spots which are the facial details in our database. The other details are ignored by calling the other fireflies to the same features and its neighbor. However, the intensity decreases as their mutual distance increase. Firefly algorithm is implemented as illustrated in algorithm (2.3).

3.3.4.2 Feature Selection Using Bat Algorithm

Using the behavior of the bat this feature selection algorithm will choose the strong features that is the proximate and nearest of data that is belong to the similar class which will drive the discriminative features for each level to defined age range.

The poor features (redundant and irrelevant face information) will be moved, features are selected from the dataset based on the types of attributes and the irrelevant features are eliminated by considering them as obstacles. Bat algorithm is implemented as in algorithm (2.4). The hybrid (Firefly and bat) algorithm is illustrated in algorithm (3.2).

Algorithm (3.2): The Hybrid (Firefly – Bat) Algorithm
Input: features Output: optimized features
Begin Step1: Generate N initial solutions randomly Step2: Firefly procedure: Set the number of iterations to process the firefly algorithm. The update will do by the equation (2.15). Step3: Every iteration of firefly algorithm will generate N number of fitness values. Step4: Bat procedure: Set a number of iterations to process the BAT algorithm. Velocity of i^{th} solution is updated by using the equation (2.16). Frequency of each solution is updated by the equation (2.17). Step5: Every iteration of the Bat algorithm will generate N number of fitness values. Step6: Using step 4 to 6, can choose top fitness values. Step7: The algorithm stops only if maximum number of iterations is attained and the solution which contains the best fitness value that is chosen by Firefly-BAT and it is given as the best rule for classification. End

3.3.5 Cross-validation

The data set images need to split up into two main groups (test and train), the data take it role simply by choosing the number k=10 to split the images to ten groups where the first group is test and the other is train. Then the operation moves

forward and group two will be test and the other is train and so on. This guarantees that all data will be used as train or test as shown in algorithm (3.3).

Algorithm (3.3): Cross – Validation Algorithm
Input: features Output: train and test
Begin Step1: apply cross validation Step2: set the k to 10 Step3: for i=1 to k Step4: Split the data to k folds where one fold is for testing and the remains folds are for training. Step5: fitting the model using k-1 Step6:end for End

3.3.6 Classifications Phase

This phase is done by applying the J48 tree and DT classification algorithms to classify the selected features obtained from the previous phase. The classification is applied to many classes of ages as describes in Table (3.3) which contains three classes.

3.3.6.1 J48 Classification Algorithm

The J48 classifier (which is the next version of Iterative Dichotomiser (ID3) algorithm) has many extra features which will account the missing details in the optimized features, reduce the pruning of errors and derivative rules, and so on.

This algorithm works in steps with the following features:

1. All the data will be trained before choosing the root which will help the j48 to choose better trained features as root.
2. The obtained features from the feature optimization phase will portioned to build new attributes.
3. The information gain and entropy are calculated for all attributes, the column with the biggest information gain (IG) will be selected as decision node, and the branches features with no entropy (zero entropy) marked as leaf of the decision tree.
4. This step run in recursive manner for only non-leaf nodes (this leaf nodes do not require further calculating since it has zero entropy).
5. All samples then belong to same class (which is originally one of the three classes of the FG-NET database). It is stop when no other column needs further partitioning.

   Decision tree algorithm is implemented as describes in algorithm (2.5).

3.3.6.2 Decision Table

A decision table is a tabular form that presents a set of conditions and their corresponding actions. Condition stubs describe the conditions or factors that will affect the decision or policy. They are listed in the upper section of the decision table. Action stubs describe in the form of statements, the possible policy actions or decisions. They are listed in the lower section of the decision table. Rules describe which actions are to be taken under a specific combination of conditions. They are specified by first inserting different combinations of condition attribute values and then putting X's in the appropriate columns of the action section of the table.

Decision table algorithm is implemented as describes in algorithm (2.6).

Chapter Four

Chapter Four

Experimental Results

4.1 Introduction

Results that are obtained thereafter execution of the proposed system are discussed in this chapter, the effect of each proposed algorithm are presented. Three scenarios are given in this system which will discuss in detail.

4.2 Proposed System Environments

The proposed system environment consists of two parts as following:

4.2.1 Software Details

The proposed system is implemented by using three programming languages which are: python v. 3.6.5(SwarmPackagePy), java (net beans IDE 8.0.2) and C++, they have characteristics and important tools in the field, In addition to high speed compared to other languages. Then the proposed system is applied under windows ten (Win10) operating system. The windows ten (Win10) operating system provides the user-friendly environment to implement the human interface and menus.

4.2.2 Hardware Details

The proposed system implemented on a computer with the following specifications: Intel core i3 due with 4GB RAM Pentium IV, 1.8E GHz with 256 cache memory (half cache), 360 GB Hard Disk and with video card Geforce2.

4.3 System Experimental Results Using J48 Classification Algorithm

Many measurements are used to evaluate the used algorithm which can be calculated as shown in Table (4.1, 4.2, and 4.3)

Three scenarios of the proposed system can be implemented if it is applied on images in the data set using J48 classification:

1. Where no feature selection phase run and only Original (LDA – J48) algorithms runs, LDA for feature extraction and J48 for classification in classification phase the outcome of this scenario is shown in Table (4.1)

Table (4.1) Original (LDA – J48)

	Class (0-7)(26-30)	Class (31-40)(41-50)	Class (8-13)(14-19)(20-25)
True Positive Rate	93.48837209302326	73.95833333333334	95.35452322738386
True Negative Rate	98.90599819768772	98.94363120579781	99.67342764118143
Precision	93.9388692969274	78.47387566137564	95.52111488069121
Recall	93.48837209302326	73.95833333333334	95.35452322738386
F-Measure	93.13668792323632	74.72551226687256	95.34999065695482
Mean Absolute Error	1.383801122694	2.487847222222	0.596695792916
Root Mean Squared Error	11.42397010	15.15286689	7.180863889
Total Number of Instances	215.0	96.0	409.0
Correct	201.0	71.0	390.0
Incorrect	14.0	25.0	19.0

Table (4.1) illustrates the values of True positive rate, True negative rate, Precision, Recall, F-measure, Mean absolute error, Root mean squared error, Total number of instances, correct and incorrect for three classes and by using only LDA with J48 classifier.

2. After selection the features in the features selection phase, the LDA algorithm is applied for feature extraction, and the firefly algorithm is used

for feature selection in the feature optimization phase. While for classification, the J48 algorithm is used in the classification phase the outcome of this scenario is shown in Table (4.2)

Table (4.2) (LDA - firefly– J48)

	Class (0-7)(26-30)	Class (31-40)(41-50)	Class (8-13)(14-19)(20-25)
True Positive Rate	93.95348837209302	77.08333333333334	95.59902200488997
True Negative Rate	99.22125941561	98.82766063956	99.73759825365
Precision	94.04551263589892	76.6216856060606	95.67203205650637
Recall	93.95348837209302	77.08333333333334	95.59902200488997
F-Measure	93.96119064763812	76.45652958152958	95.5985747302097
Mean Absolute Error	1.48698887913	2.25446428571	0.48899755501
Root Mean Squared Error	10.9419621	14.4611417	6.99283601
Total Number of Instances	215.0	96.0	409.0
Correct	202.0	74.0	391.0
Incorrect	13.0	22.0	18.0

Table (4.2) illustrates the values of True positive rate, True negative rate, Precision, Recall, F-measure, Mean absolute error, Root mean squared error, Total number of instances, correct and incorrect for three classes and by using LDA and Firefly algorithm with J48 classifier.

3. After selection the features in the features selection phase, the LDA algorithm is applied for feature extraction, and hybrid technique (firefly and bat) algorithms are used for feature selection in the feature optimization phase. While for classification, the J48 algorithm is used in the classification phase the outcome of this scenario is shown in Table (4.3)

Table (4.3) (LDA-hybrid- J48)

	Class (0-7)(26-30)	Class (31-40)(41-50)	Class (8-13)(14-19)(20-25)
True Positive Rate	95.81395348837209	78.125	95.59902200488997
True Negative Rate	99.54655454282	99.0758423079068	99.74755109842
Precision	95.86257947867452	79.67881944444444	95.80134474327629
Recall	95.81395348837209	78.125	95.59902200488997
F-Measure	95.80796240493702	78.13283025454078	95.60039981760937
Mean Absolute Error	0.83720930230232	2.09722222222	0.48899755501
Root Mean Squared Error	9.14991421	13.8744369	6.99283601
Total Number of Instances	215.0	96.0	409.0
Correct	206.0	75.0	391.0
Incorrect	9.0	21.0	18.0

Table (4.3) illustrates the values of True positive rate, True negative rate, Precision, Recall, F-measure, Mean absolute error, Root mean squared error, Total number of instances, correct and incorrect for three classes and by using LDA and Hybrid technique with J48 classifier.

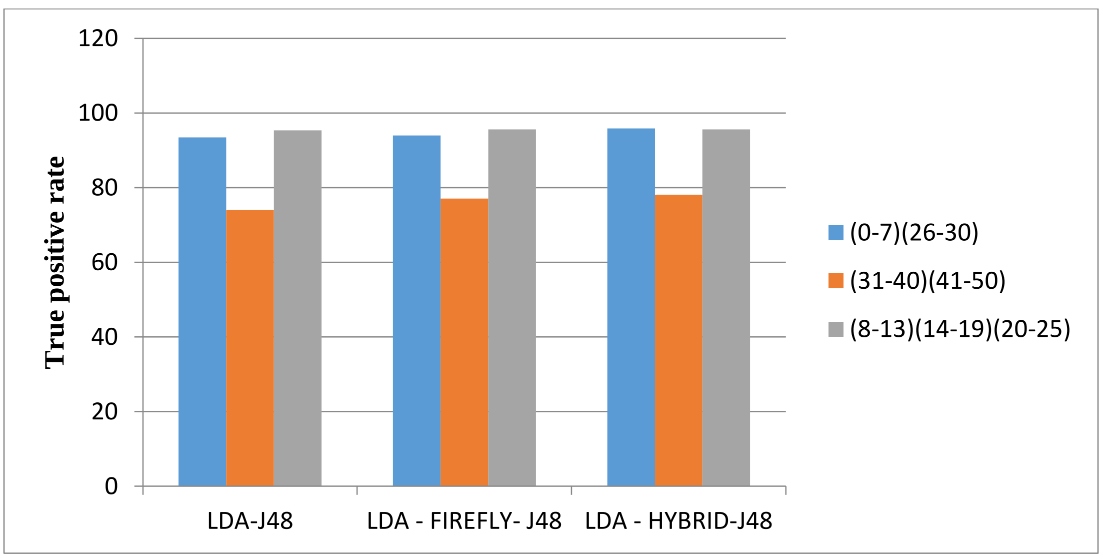


Figure (4.1) True Positive Rate of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the True Positive Rate characteristics of using (LDA-J48) is

93.48837209302326, 73.95833333333334 and 95.35452322738386 for the three classes respectively in addition to 93.95348837209302, 77.08333333333334 and 95.59902200488997 for (LDA- FIREFLY-J48) and 95.81395348837209, 78.125 and 95.59902200488997 for (LDA- HYBRID- J48) in the three classes respectively and according to Figure (4.1) the (LDA- HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

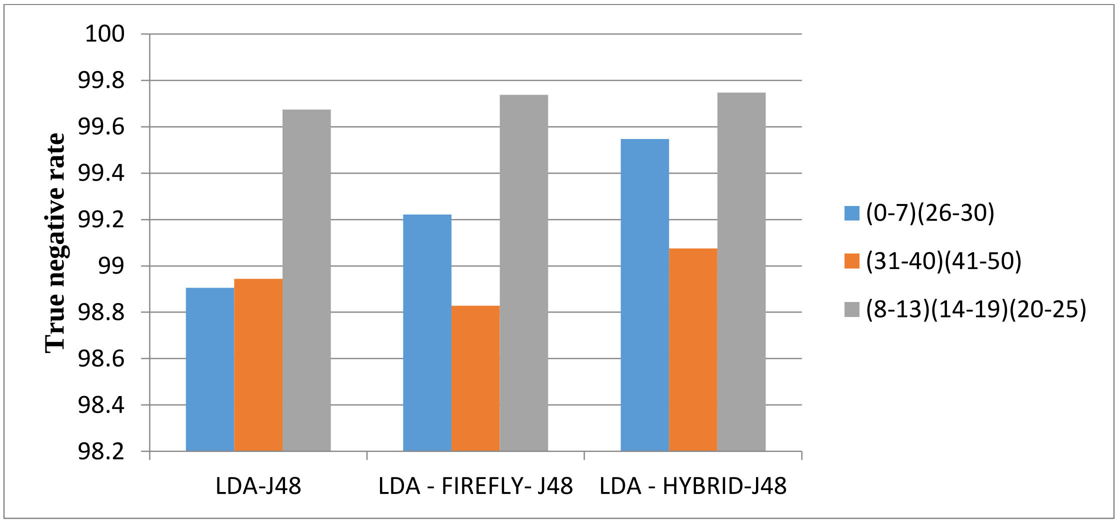


Figure (4.2) True Negative Rate of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the True Negative Rate characteristics of using (LDA-J48) is 98.90599819768772, 98.94363120579781 and 99.67342764118143 for the three classes respectively in addition to 99.22125941561, 98.82766063956 and 99.73759825365 for (LDA-FIREFLY-J48) and 99.54655454282, 99.0758423079068 and 99.74755109842for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.2) the (LDA- HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

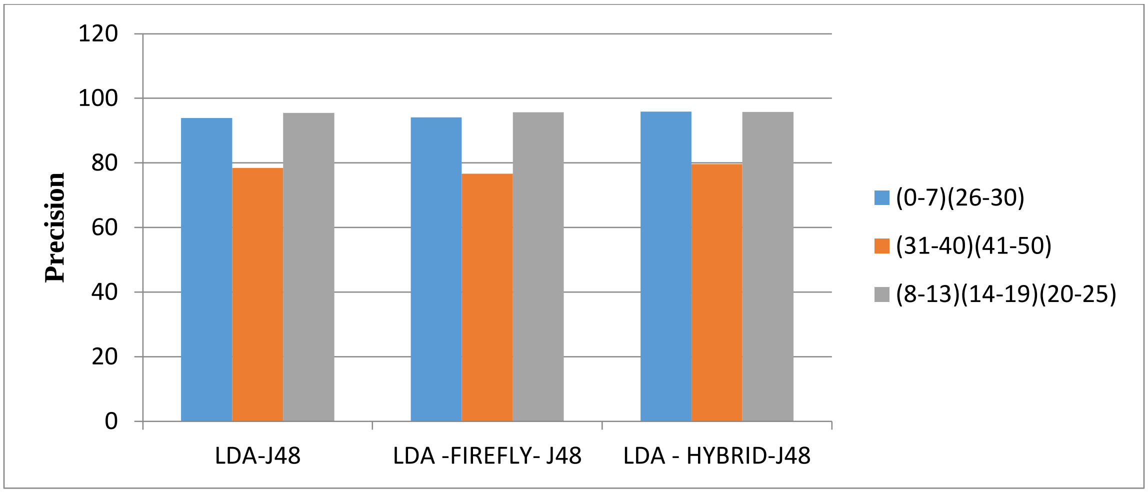


Figure (4.3) Precision of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the Precision characteristics of using (LDA-J48) is 93.9388692969274 ,78.47387566137564 and 95.52111488069121 for the three classes respectively in addition to 94.04551263589892, 76.6216856060606 and 95.67203205650637 for (LDA-FIREFLY-J48) and 95.86257947867452, 79.67881944444444 and 95.80134474327629 for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.3) the (LDA-HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

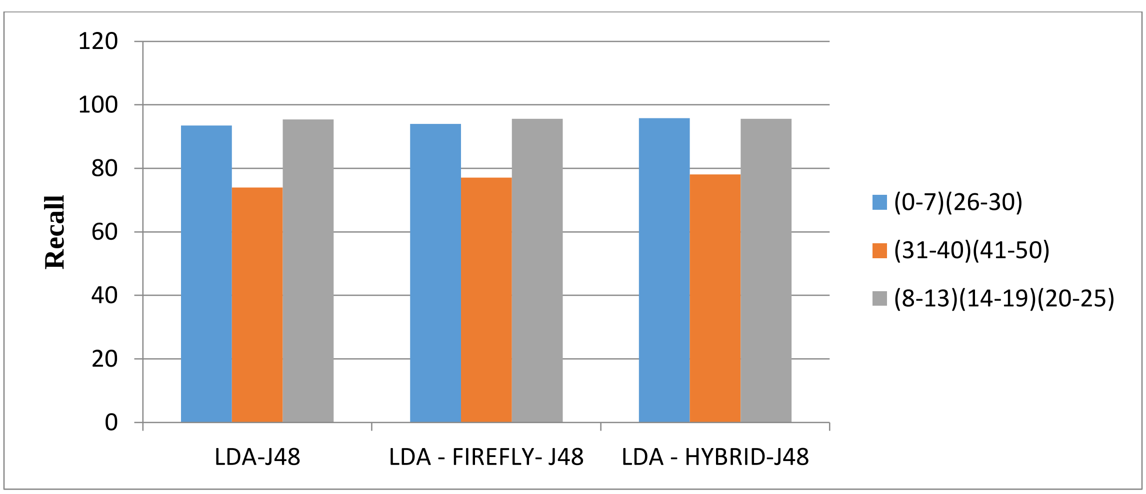


Figure (4.4) Recall of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the Recall characteristics of using (LDA-J48) is 93.48837209302326 ,73.95833333333334 and 95.35452322738386 for the three classes respectively in addition 93.95348837209302, 77.08333333333334 and 95.59902200488997 for (LDA-FIREFLY-J48) and 95.81395348837209, 78 .125 and 95.59902200488997 for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.4) the (LDA-HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

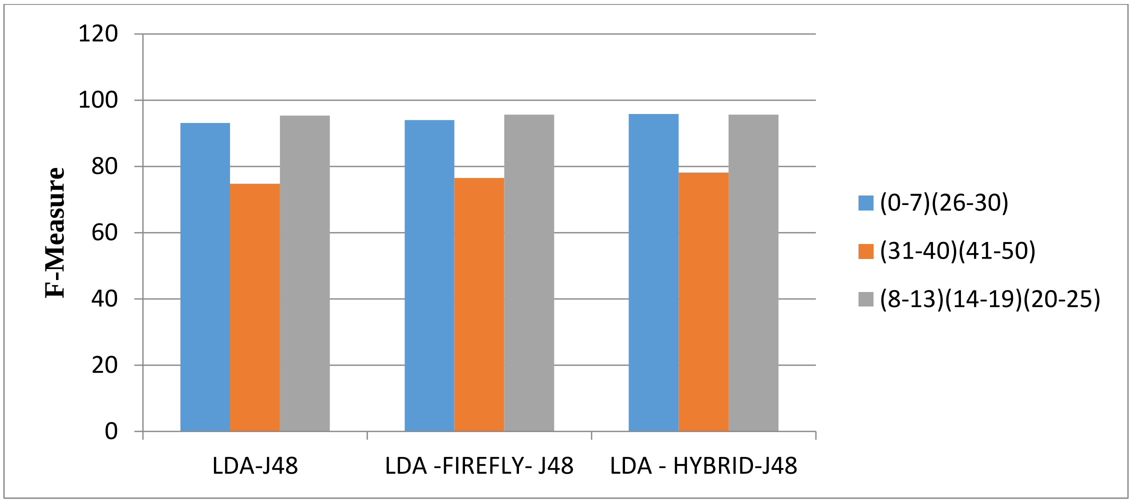


Figure (4.5) F-Measure of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the F-Measure characteristics of using (LDA-J48) is 93.13668792323632, 74.72551226687256 and 95.34999065695482 for the three classes respectively in addition 93.96119064763812, 76.45652958152958 and 95.5985747302097 for (LDA-FIREFLY-J48) and 95.80796240493702, 78.13283025454078 and 95.60039981760937 for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.5) the (LDA-HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

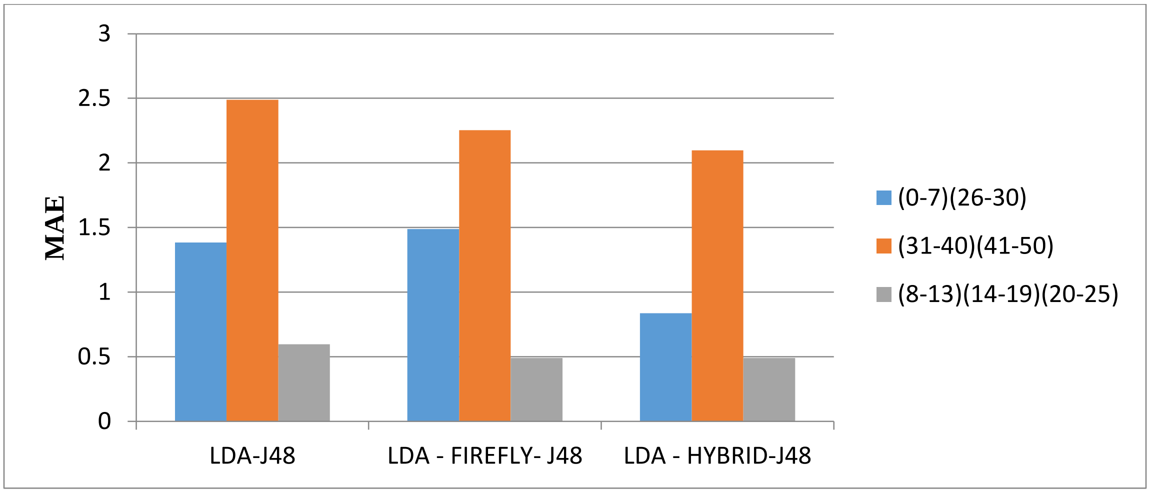


Figure (4.6) Mean Absolute Error of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the Mean Absolute Error characteristics of using (LDA-J48) is 1.383801122694, 2.487847222222 and 0.596695792916 for the three classes respectively in addition 1.48698887913, 2.25446428571and 0.48899755501for (LDA-FIREFLY-J48) and 0.83720930230232, 2.09722222222 and 0.48899755501 for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.6) the (LDA-HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

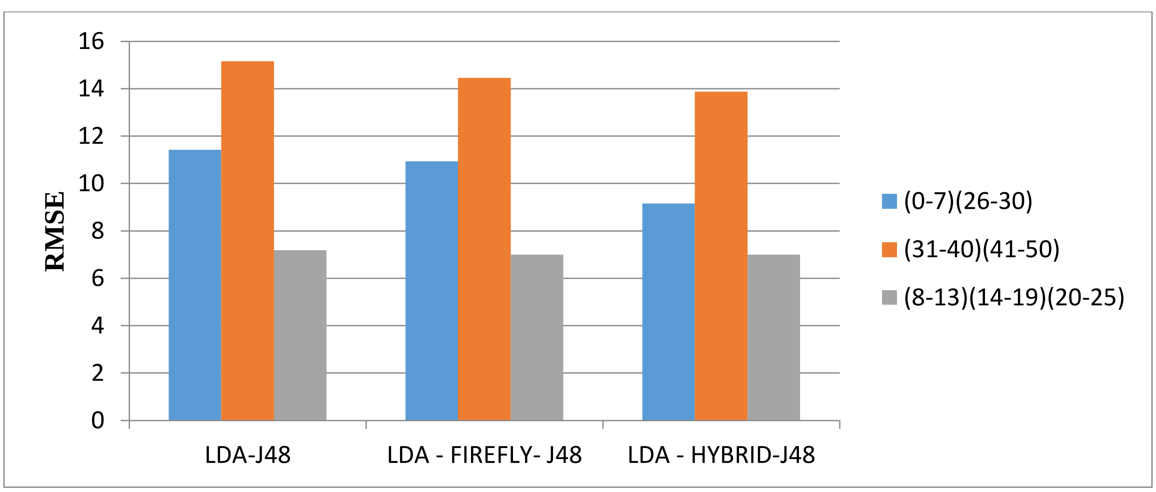


Figure (4.7) Root Mean Squared Error of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the Root Mean Squared Error characteristics of using (LDA-J48) is 11.42397010, 15.15286689 and 7.180863889for the three classes respectively in addition 10.9419621, 14.4611417 and 6.99283601for (LDA-FIREFLY-J48) and 9.14991421 , 13.8744369 and 6.99283601 for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.7) the (LDA-Hybrid(firefly-bat)) gives better results than other scenario algorithms used.

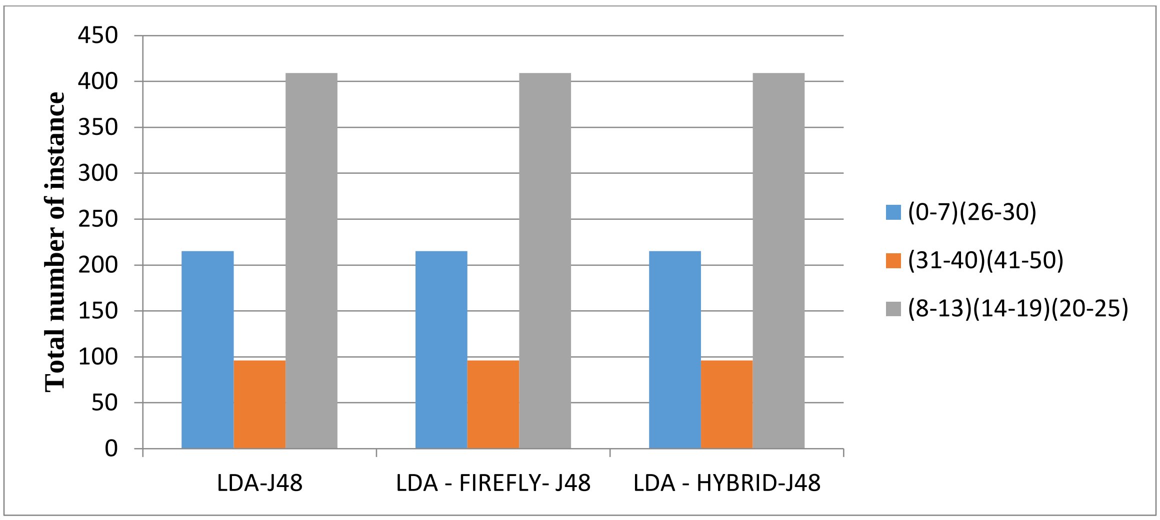


Figure (4.8) Total Number of Instances of the three scenarios

According to Table (4.1, 4.2, and 4.3) and results obtained for applying three scenarios the Total Number of Instances characteristics of using (LDA-J48) is 215.0 , 96.0 and 409.0 for the three classes respectively in addition 215.0 , 96.0 and 409.0 for (LDA-FIREFLY-J48) and 215.0 , 96.0 and 409.0 for (LDA-HYBRID-J48) in the three classes respectively and according to Figure (4.8) the algorithms shows same number of the instances in the dataset and no data is ignored before the feature selection phase.

4.4 System Experimental Results Using Decision Table Classification Algorithm

Many measurements are used to evaluate by using algorithm which can be calculated as shown in Table (4.4, 4.5, and 4.6):

Three scenarios of the proposed system can be implemented if it is applied on images in the dataset by using decision table classification:

1. Where no feature selection phase run and only Original (LDA – decision table) algorithms runs, LDA for feature extraction, and decision table for classification in classification phase the outcome of this scenario is shown in Table (4.4)

Table (4.4) Original (LDA – decision table)

	Class (0-7)(26-30)	Class (31-40)(41-50)	Class (8-13)(14-19)(20-25)
True Positive Rate	89.30232558139535	72.91666666666666	88.50855745721272
True Negative Rate	98.40419012066116	97.72061988057516	99.03087595192117
Precision	90.31859613254962	68.0955762987013	91.75955565600323
Recall	89.30232558139535	72.91666666666666	88.50855745721272
F-Measure	88. 48444618516871	68.04083180621805	89.19431497686537
Mean Absolute Error	7.822143506362	8.27272698802423	6.716573735099
Root Mean Squared Error	15.50302812	19.17962071	15.40093805
Total Number of Instances	215.0	96.0	409.0
Correct	192.0	70.0	362.0
Incorrect	23.0	26.0	47.0

Table (4.4) illustrates the values of True positive rate, True negative rate, Precision, Recall, F-measure, Mean absolute error, Root mean squared error, Total number of instances, correct and Incorrect for three classes and by using only LDA with Decision Table classifier.

2. After selection the features in the features selection phase, the LDA algorithm is applied for feature extraction, and the firefly algorithm is used for feature selection in the feature optimization phase. While for classification, the decision table algorithm is used in the classification phase the outcome of this scenario is shown in Table (4.5)

Table (4.5) (LDA - firefly– decision table)

	Class (0-7)(26-30)	Class (31-40)(41-50)	Class (8-13)(14-19)(20-25)
True Positive Rate	82.325581395348	85.416666666666	97.799511002444
True Negative Rate	96.98282952745	98.79859991922	99.84483867368
Precision	84.00916273129585	78.53050595238096	97.84268314769537
Recall	82.32558139534883	85.41666666666666	97.79951100244499
F-Measure	82.45090196387105	81.13089213418161	97.80254251516345
Mean Absolute Error	9.49488673995	8.09830443943	5.03799852674
Root Mean Squared Error	18.7414212	18.7203978	11.4897594
Total Number of Instances	215.0	96.0	409.0
Correct	177.0	82.0	400.0
Incorrect	38.0	14.0	9.0

Table (4.5) illustrates the values of True positive rate, True negative rate, Precision, Recall, F-measure, Mean absolute error, Root mean squared error, Total number of instances, correct and Incorrect for three classes and by using LDA and Firefly algorithm with Decision Table classifier.

3. After selection the features in the features selection phase, the LDA algorithm is applied for feature extraction, and hybrid technique (firefly and bat) algorithms are used for feature selection in the feature optimization phase. While for classification, the decision table algorithm is used in the classification phase the outcome of this scenario is shown in Table (4.6)

Table (4.6) (LDA - hybrid– decision table)

	Class (0-7)(26-30)	Class (31-40)(41-50)	Class (8-13)(14-19)(20-25)
True Positive Rate	84.651162790697	81.25	97.799511002444
True Negative Rate	97.41425189033	98.34292208229	99.84483867368
Precision	87.07597328323605	76.23766447368422	97.84268314769537
Recall	84.65116279069768	81.25	97.7995110024449
F-Measure	84.3935278086204	77.4022289550874	97.80254251516345
Mean Absolute Error	9.52849963505	8.14445052950	5.03799852674
Root Mean Squared Error	18.1928970	18.8364221	11.4897594
Total Number of Instances	215.0	96.0	409.0
Correct	182.0	78.0	400.0
Incorrect	33.0	18.0	9.0

Table (4.6) illustrates the values of True positive rate, True negative rate, Precision, Recall, F-measure, Mean absolute error, Root mean squared error, Total number of instances, correct and Incorrect for three classes and by using LDA and Hybrid technique with Decision Table classifier.

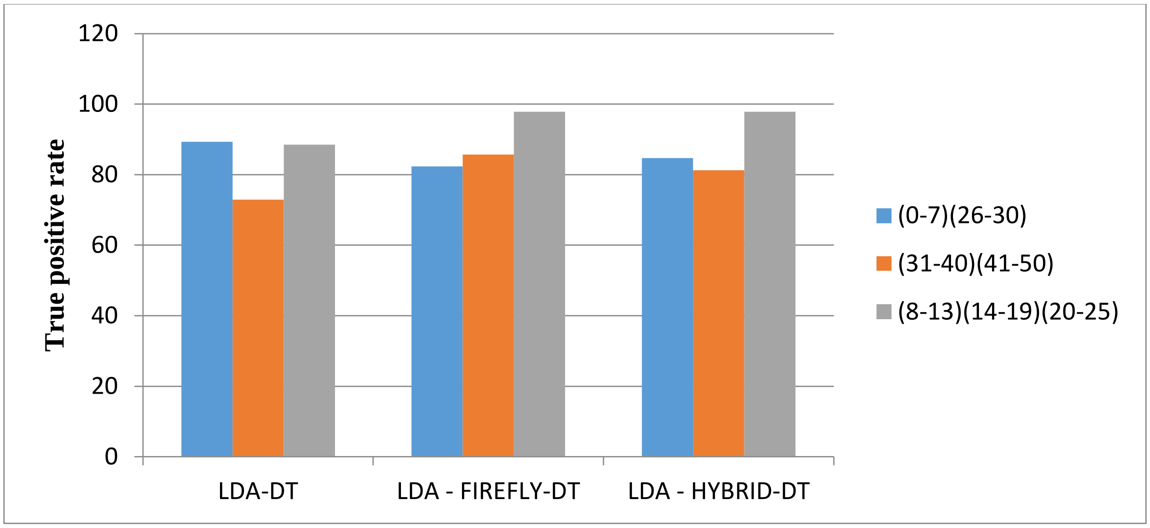


Figure (4.9) True Positive Rate of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the True Positive Rate characteristics of using (LDA- decision table) is

89.30232558139535, 72.91666666666666 , 88.50855745721272 for the three classes respectively in addition 82.325581395348 , 85.416666666666, 97.799511002444 for (LDA-FIREFLY-decision table) and 84.651162790697 ,81.25, 97.799511002444 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.9) the (LDA-FIREFLY) gives better results than other scenario algorithms used.

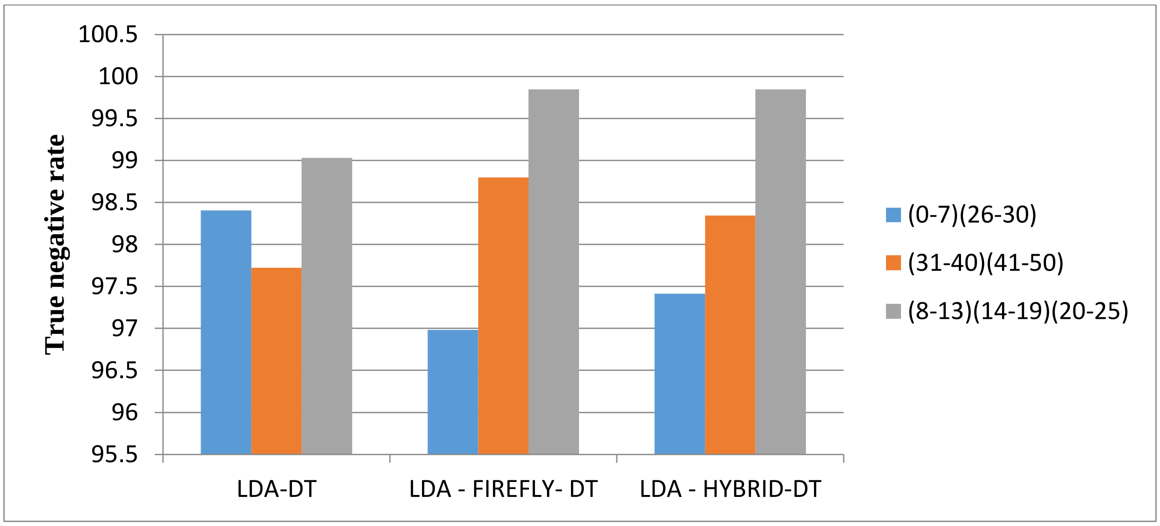


Figure (4.10) True Negative Rate of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the True Negative Rate characteristics of using (LDA-decision table) is 98.40419012066116, 97.72061988057516 , 99.03087595192117 for the three classes respectively in addition 96.98282952745, 98.79859991922 , 99.84483867368 for (LDA-FIREFLY-decision table) and 97.41425189033, 98.34292208229, 99.84483867368 for (LDA-HYBRID-decision table) in the three classes respectively .

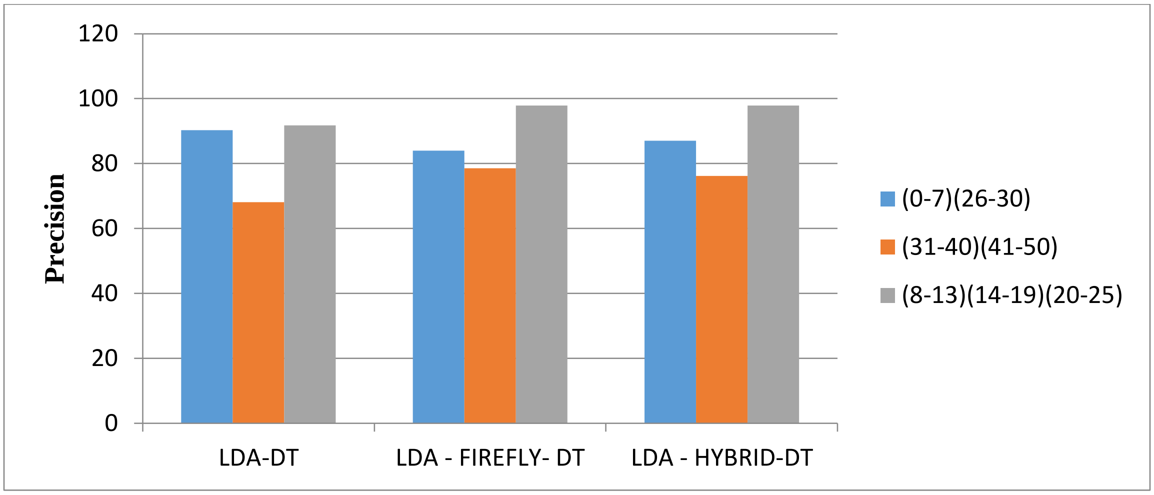


Figure (4.11) Precision of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the Precision characteristics of using (LDA-DECISION TABLE) is 90.31859613254962, 68.0955762987013, 91.75955565600323 for the three classes respectively in addition 84.00916273129585, 78.53050595238096, 97.84268314769537 for (LDA-FIREFLY-decision table) and 87.07597328323605, 76.23766447368422, 97.84268314769537 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.11) the (LDA-HYBRID(firefly-bat)) gives better results than other scenario algorithms used.

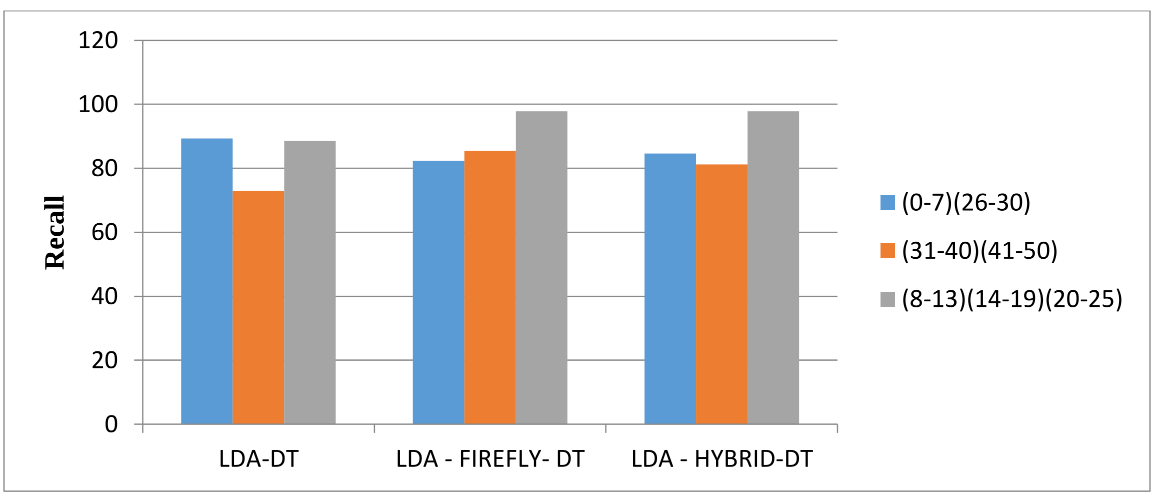


Figure (4.12) Recall of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the Recall characteristics of using (LDA-decision table) is 89.30232558139535, 72.91666666666666 , 88.50855745721272 for the three classes respectively in addition 82.32558139534883, 85.41666666666666, 97.79951100244499 for (LDA-FIREFLY-decision table) and 84.65116279069768, 81.25, 97.7995110024449 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.12) the (LDA-FIREFLY) gives better results than other scenario algorithms used.

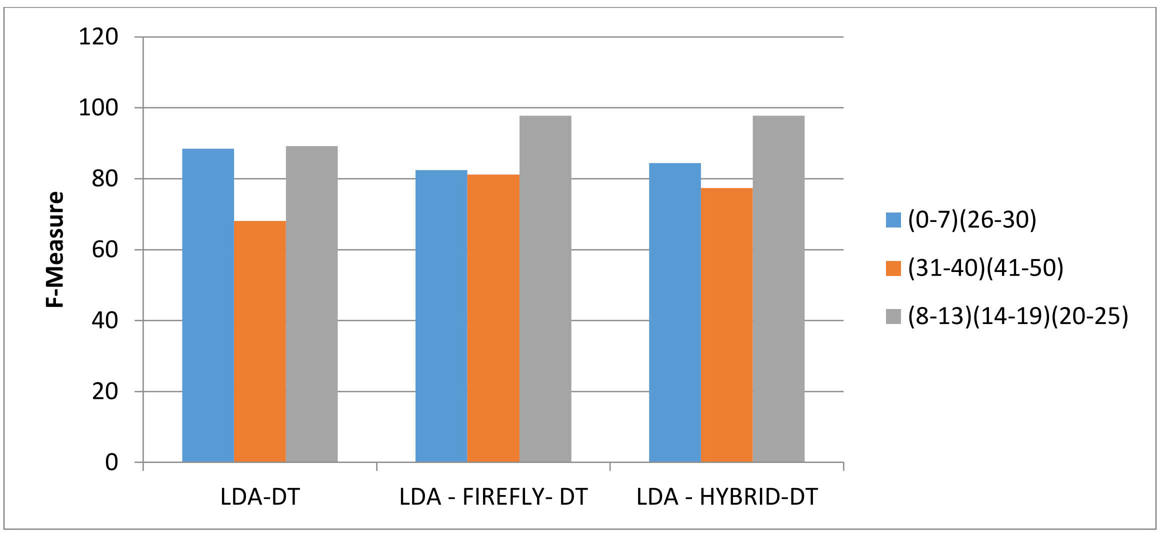


Figure (4.13) F-Measure of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the F-Measure characteristics of using (LDA-decision table) is

88. 48444618516871, 68.04083180621805, 89.19431497686537 for the three classes respectively in addition 82.45090196387105, 81.13089213418161, 97.80254251516345 for (LDA-FIREFLY-decision table) and 84.3935278086204, 77.4022289550874, 97.80254251516345 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.13) the (LDA-FIREFLY) gives better results than other scenario algorithms used.

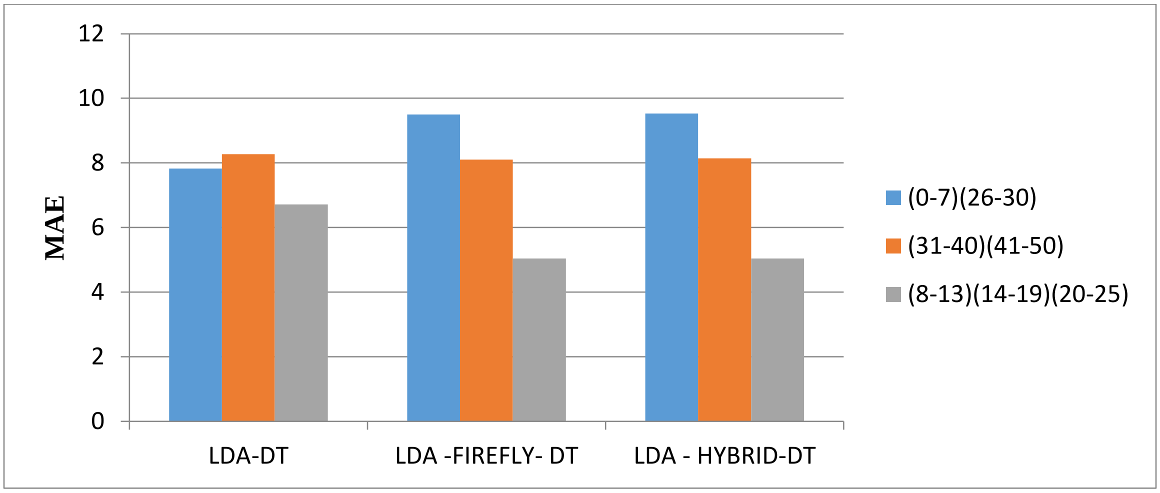


Figure (4.14) Mean Absolute Error of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the Mean Absolute Error characteristics of using (LDA-decision table) is 7.822143506362, 8.27272698802423, 6.716573735099 for the three classes respectively in addition 9.49488673995, 8.09830443943 , 5.03799852674 for (LDA-FIREFLY-decision table) and 9.52849963505, 8.14445052950, 5.03799852674 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.14) the (LDA- FIREFLY) gives better results than other scenario algorithms used.

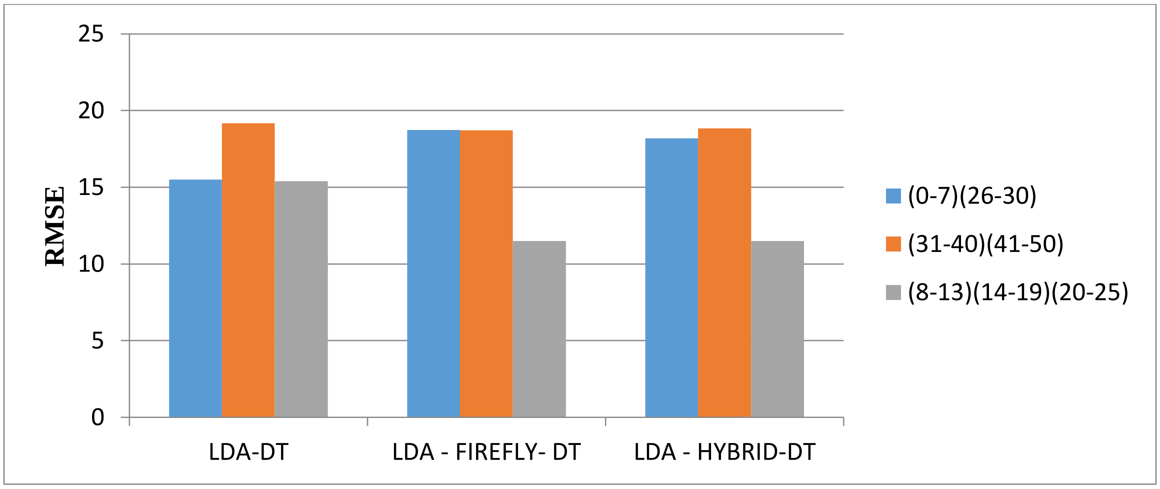


Figure (4.15) Root Mean Squared Error of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the Root Mean Squared Error characteristics of using (LDA-decision table) is 15.50302812, 19.17962071, 15.40093805 for the three classes respectively in addition 18.7414212, 18.7203978 ,11.4897594 for (LDA-FIREFLY-decision table) and 18.1928970, 18.8364221, 11.4897594 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.15) the (LDA-HYBRID) gives better results than other scenario algorithms used.

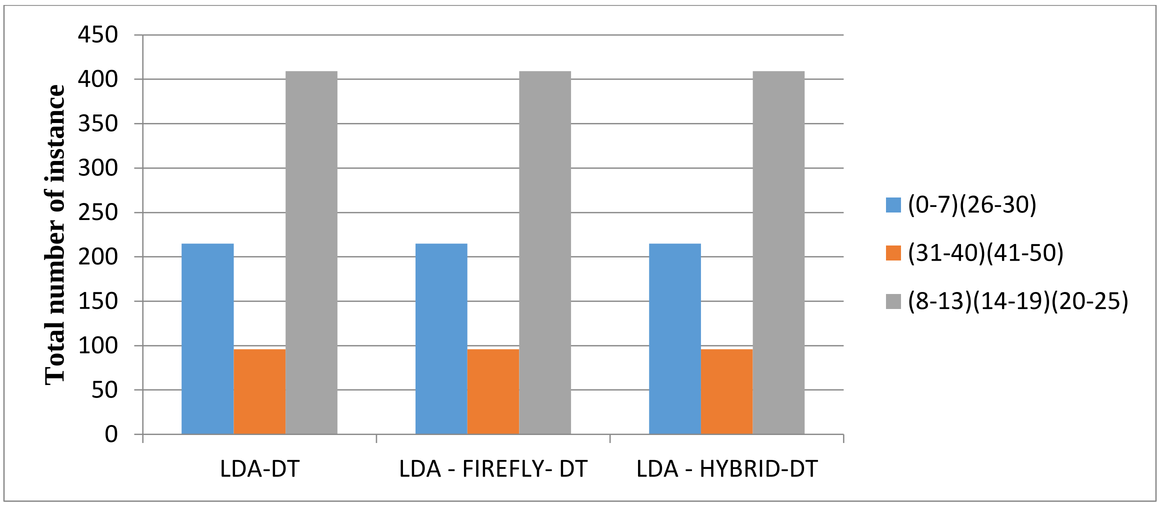


Figure (4.16) Total Number of Instances of the three scenarios

According to Table (4.4, 4.5, and 4.6) and results obtained for applying three scenarios the Total Number of Instances characteristics of using (LDA-decision table) is 215.0, 96.0, 409.0 for the three classes respectively in addition 215.0, 96.0, 409.0 for (LDA-FIREFLY-decision table) and 215.0, 96.0, 409.0 for (LDA-HYBRID-decision table) in the three classes respectively and according to Figure (4.16) the algorithms shows same number of the instances in the dataset and no data is ignored before the feature selection phase.

4.5 Comparison of The Proposed Method To The Related Works

The comparison of MAE for the proposed system with MAE for related work and these values evidence proposed system effectiveness and performance.

Table (4.7) Comparison MAE of the proposed method with MAE for related works on FG-NET database

Year	**Authors**	**Method**			**MAE**
		Feature extraction	**Feature Selection**	**Classification Algorithm**	
2015	J. K. Pontes, A. S. Britto, and et al	AAM, LBP, GW, LPQ	-	SVM, and SVR	4.50
2015	Pablo Pardo García, and Sergio Escalera Guerrero	BIF	-	SVM	7.99
2016	Wei Zhao and Han Wang	SDM-LDL	-	SVM, and KKN	5.07
2017	Xiaolong Wang, Robert Li, et al	CSC, CNN and STD	-	SVR	4.01
2018	Soumaya Zaghbani, et al	ANN	-	Autoencoder	3.75
2019	Shahram Taheri and Önsen Toygar	BIF, MRELBP, HOG,and KFA	-	KKN	4.06
The proposed method		LDA	-	J48	1.48
				DT	7.60
			hybrid firefly and bat algorithm	J48	1.14
				DT	7.53

Chapter Five

Chapter Five

Conclusions and Suggestions for Future Works

5.1 Conclusions

This section summarizes the results of the thesis in short, the main objective of which is to estimate the age of the human using machine learning techniques and the conclusion shows the following points:

1- The experimental results indicate that feature selection by using Firefly and Bat algorithms enhanced the performance of the proposed model in comparison with same proposed model without feature selection.

2- The proposed system when using the feature extraction approach based on LDA with J48 classifier is give a result of precision 89.30% while when use LDA with Decision table the obtained precision was 83.39%.

3- The system proved its effectiveness to evaluate human age where the precision reached by using Firefly algorithm with J48 classifier 88.77% ,and reached by using Firefly algorithm with a decision table the precision was 86.79%.

4- The proposed system produce high precision when using the Hybrid algorithm (firefly and bat) with J48 classifier which is 90.44%, and the precision when using the Hybrid algorithm with a decision table is 87.05%.

5- Comparing the proposed system with other related works shows that the proposed system is more effectiveness over other related works.

5.2 Suggestions for Future Works

Future works can summarize as follows:

1- Implementing system by using other databases or gathering data (images) to people in our region.
2- Optimizing extraction features by other selection optimization algorithms instead of Firefly and Bat algorithms like (flower pollination, fish school, etc).
3- Implementing a human age estimate system using other types of machine learning techniques like Sequential Minimal Optimization (SMO).
4- Implementing system with other Media like video.
5- Using deep learning technique like (Reinforcement Learning to Neural Networks) instead of this method.

References

[1] R. Ranjan, S. Zhou, and et al, **“Unconstrained Age Estimation with Deep Convolutional Neural Networks”**, *IEEE International Conference on Computer Vision Workshop (ICCVW)*, 2015.

[2] A. A. Shejul, K. S. Kinage, and B. E. Reddy, **“Comprehensive Review on Facial based Human Age Estimation”**, *International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 2017.

[3] I. Huerta, C. Fernández, and et al, **“A deep analysis on age estimation”**, *Pattern Recognition Letters*, vol. 68, pp. 239 - 249, 2015.

[4] C. Belver, I. Arganda-Carreras, and F. Dornaika, **“Evaluating age estimation using deep convolutional neural nets”**, *IS&T International Symposium on Electronic Imaging Computational Imaging Science and Technology*, pp. 100–105, 2017.

[5] K. H. Liu, S. Yan, and C. C. J. Kuo, **“Age Estimation via Grouping and Decision Fusion”**, *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2408–2423, 2015.

[6] C. Belver, I. Arganda-Carreras, and F. Dornaika, **“Comparative Study of Human Age Estimation Based on Hand-Crafted and Deep Face Features”**, *Springer International Publishing AG*, pp. 98-112, 2017.

[7] S. Shigh, and S.V. A.V. Prasad, **“ Techniques and Challenges of Face Recognition: A Critical Review”**, *Procedia Computer Science*, 143, pp. 536-543, 2018.

[8] J. I. Olszewska, **“Automated Face Recognition: Challenges and Solutions”**, *Pattern Recognition - Analysis and Applications*, 2016.

[9] S. Yang, P. Luo, and et al, **“Faceness-Net: Face Detection through Deep Facial Part Responses”**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 8, pp. 1845–1859, 2018.

[10] N. Jagtap, P. P. Shevatekar, and N. Mustary, **“A Comparative Study of Classification Techniques in Data Mining Algorithms”**, *International Journal of Modern Trends in Engineering and Research (IJMTER)*, Vol. 04, 2017.

[11] F. Fabris, J. P. d. Magalha˜es, and A. A. Freitas, **“A review of supervised machine learning applied to ageing research**”, *Biogerontology*, vol. 18(2), pp. 171–188, 2017.

[12] A. O. abass, G. Naghdy, and C. Todd, **“Unsupervised Classification of Images: A Review”**, *International Journal of Image Processing (IJIP)*, Vol. 8, 2014.

[13] J. K. Pontes, A. S. Britto, and et al, **“A flexible hierarchical approach for facial age estimation based on multiple features”**, *Pattern Recognition*, vol. 54, pp. 34–51, 2015.

[14] P. P. García, **“Automatic Real and Apparent Age Estimation in Still Images”**, Thesis for the degree of Master of Science in Artificial Intelligence (UPC-UB-URV), at University of Barcelona (UB), on 20th April 2015.

[15] W. Zhao, and H. Wang, **“Strategic decision-making learning from label distributions: An approach for facial age estimation”**, *Sensors*, vol. 16, no. 7, 2016.

[16] X. Wang, R. Li, Y. Zhou, and C. Kambhamettu, **“A Study of Convolutional Sparse Feature Learning for Human Age Estimate”**, *12th IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, pp. 566–572, 2017.

[17] S. Zaghbani, N. Boujneh, and M. S. Bouhlel, **“Age estimation using deep learning”**, *Computers and Electrical Engineering*, vol. 68, no. April, pp. 337–347, 2018.

[18] S. Taheri, and Ö. Toygar, **“Integrating Feature Extractors for the Estimation of Human Facial Age”**, *Applied Artificial Intelligence*, vol. 33, no. 5, pp. 379–398, 2019.

[19] R. Angulu, J. R. Tapamo, and A. O. Adewumi, **“Age estimation via face images: a survey”**, *EURASIP Journal on Image and Video Processing*, 2018.

[20] O. F. Osman, and M. H. Yap, **“Computational Intelligence in Automatic Face Age Estimation: A Survey”**, *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2018.

[21] C. - C. NG, **“Face Age Estimation Using Wrinkle Patterns”**, A thesis submitted in partial fulfilment of the requirements for the degree of doctor of philosophy school of computing, mathematics and digital technology manchester metropolitan university, 2015.

[22] R. B. G. Luta, R. G. Baldovino, and N. T. Bugtai, **“Image Preprocessing using Quick Color Averaging Approach for Color Machine Vision (CMV) Systems”**, *IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control,*

Environment and Management (HNICEM), 2017.

[23] F. Ma, and Y. Wan, **“A New Color-to-Gray Conversion Framework Based on Normalized Cross Correlation”**, *3rd IEEE International Conference on Computer and Communications (ICCC)*, 2017.

[24] Y. Wan, and Q. Xie, **“A Novel Framework for Optimal RGB to Grayscale Image Conversion”**, *8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2016.

[25] W. A. Mustafa, and M. M. M. Abdul Kader**, “A Review of Histogram Equalization Techniques in Image Enhancement Application”**, *Journal of Physics: Conference Series 1019 012026*, 2018.

[26] M. Abdul, H. Radhi, and et al, **“Enhancement of The Captured Images Under Different Lighting Conditions Using Histogram Equalization Method”**, *International Journal of Latest Research in Science and Technology*, vol. 3, no. 3, pp. 25–28, 2014.

[27] Histogram Equalization, https://digitalpadm.com/image-contrast-enhancement-cumulative-histogram-equalization/.

[28] S. Raja, S. Rajb, and S. Kumar, **“An Improved Histogram Equalization Technique forImage Contrast Enhancement”**, *International Conference on Communication and Computing (ICC)*, 2015.

[29] M. Chauhan, and M. Sakle, **“Study & Analysis of Different Face Detection Techniques”**, *(IJCSIT) International Journal of Computer Science and Information Technologies*, vol. 5 (2), pp. 1615-1618, 2014.

[30] H. Hatem, Z. Beiji, and R. Majeed, **“A Survey of Feature Base Methods for**

Human Face Detection", *International Journal of Control and Automation*, vol. 8, No.5, pp.61-78, 2015.

[31] M. Sharif, F. Naz, and et al, **"Face Recognition: A Survey"**, *Journal Of Engineering Science and Technology Review*, vol. 10, pp. 166–177, 2017.

[32] A. Sharifara, M. S. M. Rahim, and Y. Anisi, **"A General Review of Human Face Detection Including a Study of Neural Networks and Haar Feature-based Cascade Classifier in Face Detection"**, *International Symposium on Biometrics and Security Technologies (ISBAST)*, 2014.

[33] S. Sarkar, V. M. Patel, and R. Chellappa, **"Deep feature-based face detection on mobile devices"**, *ISBA - IEEE International Conference on Identity, Security and Behavior Analysis*, 2016.

[34] N. Wang, X. Gao, and et al, **"Facial feature point detection: A comprehensive survey"**, *Neurocomputing*, vol. 275, pp. 50–65, 2017.

[35] X. Zhang, Y. Sugano, and et al, **"Appearance-based gaze estimation in the wild"**, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 07-12-June, pp. 4511–4520, 2015.

[36] D. Meena, and R. Sharan, **"An Approach to Face Detection and Recognition"**, *IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, 2016.

[37] T. Hassner et al, **"Pooling Faces: Template Based Face Recognition with Pooled Face Images"**, *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 127–135, 2016.

[38] N. Hazim, S. Sameer, and et al, **"Face Detection and Recognition Using**

Viola-Jones with PCA-LDA and Square Euclidean Distance", *International Journal of Advanced Computer Science and Applications (IJACSA),* vol. 7, no. 5, pp. 371–377, 2016.

[39] Z. Hao, Q. Feng, and L. Kaidong, **"An Optimized Face Detection Based on Adaboost Algorithm"**, *International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pp. 375 – 378, 2018.

[40] N. T. Deshpande and S. Ravishankar, **"Face Detection and Recognition using Viola - Jones algorithm and fusion of LDA and ANN"**, *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 18, no. 6, pp. 1– 6, 2016.

[41] R. R. Damanik, D. Sitanggang, and et al, **"An application of viola jones method for face recognition for absence process efficiency"**, *IOP Conference Series: Journal of Physics: Conference Series*, 2018.

[42] E. Olmedo, J. de la Calleja, and et al, **"A parallel approach for the training stage of the Viola-Jones face detection algorithm"**, *Intelligent Data Analysis*, 21(5), pp. 1097–1115, 2017.

[43] J. Kaur, and A. Sharma, **"Performance Analysis of Face Detection by using Viola-Jones algorithm"**, *International Journal of Computational Intelligence Research*, vol. 13, pp. 70 -717, 2017.

[44] P. Lalitha and R. Nedunchelianr, **"A Case Study on Analysis of Face Recognition Techniques"**, *International Journal of Scientific and Research Publications,* vol. 4, no. 7, pp. 1–4, 2014.

[45] F. Juefei-Xu and M. Savvides, **"Multi-class Fukunaga Koontz discriminant analysis for enhanced face recognition"**, *Pattern Recognition*, vol. 52, pp. 186–205, 2016.

[46] Q. Ye, L. Fu, and et al, **“Lp- and Ls-Norm Distance Based Robust Linear Discriminant Analysis”**, *Neural Networks*, vol. 105, pp. 393–404, 2018.

[47] L. Wu, Chunhua Shen, and A. van den Hengel, **“Deep linear discriminant analysis on fisher networks: A hybrid architecture for person re-identification”**, *Pattern Recognition*, vol. 65, pp. 238–250, 2017.

[48] S. E. Safo, and J. Ahn, **“General sparse multi-class linear discriminant analysis”**, *Computational Statistics and Data Analysis*, vol. 99, pp. 81–90, 2016.

[49] A. Tharwat, T. Gaber, and et al, **“Linear discriminant analysis : A detailed tutorial”**, *AI Communications*, vol. 30, pp. 169–190, 2017.

[50] L. Zhang, K. Mistry, and et al, **“Feature Selection Using Firefly Optimization for classification and regression models”**, *Decision Support Systems*, 2017.

[51] Y. Zhang, X. Song, and D. Gong, **“A return-cost-based binary firefly algorithm for feature selection”**, *Information Sciences*, vol. 419, pp. 561–574, 2017.

[52] A. Liu, K. Chen, and et al, **“Feature Selection for Motor Imagery EEG Classification Based on Firefly Algorithm and Learning Automata”**, *sensors*, 2017.

[53] I. Fister, X. Yang, and et al, **“Firefly Algorithm: A Brief Review of the Expanding Literature”**, *Springer International Publishing Switzerland*, vol. 516, pp. 1–26, 2014.

[54] I. Fister, M. Perc, and et al, **“A review of chaos-based firefly algorithms:**

Perspectives and research challenges", *Applied Mathematics and Computation*, vol. 252, pp. 155–165, 2015.

[55] B. Yang, Y. Lu, and et al, **"Feature Selection Based on Modified Bat Algorithm"**, *EICE Transactions on Information and Systems*, no. 8, pp. 1860–1869, 2017.

[56] H. Patel and M. Shah, **"Rough Set Feature Selection Using Bat Algorithm"**, *International Journal of Computer Science and Mobile Computing*, vol. 55, no. 5, pp. 312–321, May, 2016.

[57] L. Y. Chuang, S. W. Tsai, and C. H. Yang, **"Improved binary particle swarm optimization using catfish effect for feature selection"**, *Expert Systems with Applications*, vol. 38, no. 10, pp. 12699–12707, 2011.

[58] M. Tuba, M. Jordanski, and A. Arsic, **"Improved weighted thresholded histogram equalization algorithm for digital image contrast enhancement using the bat algorithm"**, *Bio-Inspired Computation and Applications in Image Processing*, pp. 61- 86, 2016.

[59] P. Shrivastava, A. Shukla, and et al, **"A survey of nature-inspired algorithms for feature selection to identify Parkinson's disease"**, *Computer Methods and Programs in Biomedicine*, vol. 139, pp. 171–179, 2016.

[60] S. Yılmaz, and E. U. Kuc, **"A New Modification Approach on Bat Algorithm for Solving Optimization Problems"**, *Applied Soft Computing Journal*, vol. 28, pp. 259–275, 2015.

[61] P. Pandey and R. Prabhakar, **"An Analysis of Machine Learning Techniques (J48 & AdaBoost) -for Classification"**, *1st India*

International Conference on Information Processing (IICIP). IEEE, pp. 0–5, 2016.

[62] N. Saravanan and V. Gayathri, **“Classification of dengue dataset using J48 algorithm and ant colony based AJ48 algorithm”**, *Proceedings of the International Conference on Inventive Computing and Informatics, ICICI*, no. Icici, pp. 1062–1067, 2017.

[63] U. Bashir, and M. Chachoo, **“Performance Evaluation of J48 and Bayes Algorithms For Intrusion Detection System”**, *International Journal of Network Security & Its Applications (IJNSA),* Vol.9, No.4, 2017.

[64] N. V. Kalyankar, **“Effect of Training Set Size in Decision Tree Construction by using GATree and J48 Algorithm”**, *Proceedings of the World Congress on Engineering*, vol. I, pp. 4–7, 2018.

[65] S. Aljawarneh, M. Bani, and M. Aljundi, **“An enhanced J48 classification algorithm for the anomaly intrusion detection systems”**, *Cluster Computer*, 2017.

[66] S. Kalmegh, **“Comparative Analysis of the WEKA Classifiers Rules Conjunctive rule & Decision table on Indian News Dataset by Using Different Test Mode”**, *International Journal of Engineering Science Invention (IJESI)*, vol. 7, no. 2, pp. 1–9, 2018.

[67] P. Verkade, **“Model-Based Unit Testing using Decision Tables”**, Thesis for degree of Master in software engineering, at Universiteit van Amsterdam, on July, pp.1-60, 2017.

[68] A. Jamali, and A. Abdul Rahman, **“Evaluation of Advanced Data Mining Algorithms in Land Use/Land Cover Mapping”**, *International Archives of*

the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-4/W16, pp. 283–289, 2019.

[69] V. Veeralakshmi, and D. Ramyachitra, **"Ripple Down Rule learner (RIDOR) Classifier for IRIS Dataset"**, *International Journal of Computer Science Engineering (IJCSE)*, vol.4, no.03, May 2015.

[70] FG-NET (Face and Gesture Recognition Network). http://yanweifu. github. io /FG_NET_data/FGNET.zip, 2014.

[71] G. Panis, and A. Lanitis, **"An Overview of Research Activities in Facial Age Estimation Using the FG-NET Aging Database"**, *Springer International Publishing Switzerland*, 2015.

[72] C. Oprea, **"Performance Evaluation of the Data Mining Classification methods"**, *Information society and sustainable development*, pp. 249–253, 2014.

الملخص

في العقود القليلة الماضية ، يعد تقدير العمر للبشر من البيانات المرئية مثل بيانات الفيديو والصور الثابتة أحد الأبحاث المهمة نظرًا لوجود العديد من التطبيقات المحتملة مثل تطوير واجهات ذكية بين الإنسان والآلة وتحسين السلامة والحماية في أقسام مختلفة مثل النقل والأمن والدواء. من ناحية أخرى ، لا يزال تقدير العمر أمرًا صعبًا للغاية نظرًا لأن الشيخوخة تتأثر بعدة عوامل مثل الحوادث وشعر الوجه والتجاعيد وحالة الجلد. في الآونة الأخيرة ، يتم استخدام العديد من تقنيات التعلم الآلي لتقدير الأعمار.

في هذه الأطروحة ، تم اقتراح نظام ذكي جديد لتقدير عمر الإنسان من صور الوجه باستخدام تقنيات التعلم الآلي وخوارزميتان مبنيتان على ذكاء السرب. يتكون الهيكل النموذجي للنظام المقترح من عدة مراحل: مرحلة ما قبل المعالجة ، مرحلة الكشف عن الوجه ، مرحلة استخراج المعالم ، مرحلة اختيار الميزات للتخلص من الميزات الزائدة ، ومرحلة التصنيف. جرب النظام المقترح استخدام قاعدة بيانات الشيخوخة FG-NET. تنقسم قاعدة البيانات إلى سبعة فصول تبين أن بعض الفئات لها نفس عدد الميزات. لذلك ، تم دمج هذه الفئات السبعة في ثلاث فئات بناءً على عدد الميزات التي تحتوي عليها لتقليل الخطأ المطلق المتوسط وزيادة الدقة. يتم تنفيذ مرحلة الكشف عن الوجه باستخدام خوارزمية فيولا جونز. يتم استخدام استخراج الميزة باستخدام تحليل خطي للتمييز LDA ، ويتم تنفيذ اختيار الميزة باستخدام اثنين من خوارزميات تعتمد على الذكاء ، خوارزميات اليراع والخفافيش. أخيرًا ، يتم استخدام خوارزميات مبنية على التعلم الآلي ، مصنف J48 وجدول القرار لتقدير العمر.

أظهرت النتائج التي تم الحصول عليها أن النظام المقترح قد أعطى دقة 88.77 ٪ عند استخدام Firefly مع مصنف J48 ، بينما حقق دقة 86.79 ٪ عند استخدام Firefly مع مصنف جدول القرار. علاوة على ذلك ، تم تطبيق نهج هجين من خوارزميات Firefly و Bat مع مصنف J48 وأعطى دقة 90.45٪ ، بينما حقق النهج الهجيني نفسه دقة 87.05٪ عند استخدام مصنف جدول القرار. أخيرًا ، تم الحصول على أفضل نتيجة لمتوسط الخطأ المطلق عند 1.14 ، من خلال النهج الهجين لخوارزمية Firefly و Bat مع مصنف J48.

جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة ديالى

كلية العلوم

نظام ذكي لتقديرعمر الانسان القائم على تقنيات هجينة

رسالة

مقدمة الى قسم علوم الحاسوب / كلية العلوم/ جامعة ديالى وهي جزء من متطلبات نيل درجة الماجستير في علوم الحاسوب

من قبل

دلال عدنان عباس

بإشراف

أ.م. د. جمال مصطفى عباس

2020م

1441 هـ