Tribhuvan University

Institute of Science and Technology

Kathford International College of Engineering and Management

A Facial Expression Recognition System

A Project Report

Submitted To

Department of Science and Information Technology,

Kathford International College of Engineering and Management,

Balkumari, Lalitpur, Nepal

Submitted By

Nalina Matang (2203/069)

Shreejana Sunuwar (2218/069)

Sunny Shrestha (2228/069)

Sushmita Parajuli (2229/069)

Under the Supervision of

Mr. Ashok Kumar Pant

(Sr. Software Engineer, Innovisto Pvt. Ltd.)

Date: September 2016

Tribhuvan University

Institute Of Science and Technology

STUDENT'S DECLARATION

We, the undersigned solemnly declare that the report of the project work entitled **"FACIAL EXPRESSION RECOGNITION SYSTEM",** is based on our work carried out during the course of study under the supervision of **Mr. Ashok Kumar Pant.**

We assert that the statements made and conclusions drawn are an outcome of the project work. We further declare that, to the best of our knowledge and belief that the project report does not contain any part of any work which has been submitted for the award of any other degree/diploma/certificate in this University.

………………

Ms. Nalina Matang

(2203/069)

…………………

Ms. Shreejana Sunuwar

(2218/069)

…………………

Ms. Sunny Shrestha

(2228/069)

……………………

Ms. Sushmita Parajuli

(2229/069)

Tribhuvan University

Institute Of Science and Technology

Supervisor's Recommendation

I hereby recommend that this project work report is satisfactory in the partial fulfillment for the requirement of Bachelor of Science in Computer Science and Information Technology and be processed for the evaluation.

…………………...........................
Mr. Ashok Kumar Pant
Sr. Software engineer
Innovisto Pvt. Ltd.
(Supervisor)

Date:

Tribhuvan University

Institute Of Science and Technology

LETTER OF APPROVAL

This is to certify that the project prepared by **Ms. Nalina Matang (2203/069), Ms. Shreejana Sunuwar (2218/069), Ms. Sunny Shrestha (2228/069)** and **Ms. Sushmita Parajuli (2229/069)** entitled **"FACIAL EXPRESSION RECOGNITION SYSTEM"** in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Information Technology has been well studied. In our opinion it is satisfactory in the scope and quality as a project for the required degree.

…………………..........................
Department of Computer Science
and Information Technology
Kathford International College of
Engineering and Management

…………………..........................
Mr. Ashok Kumar Pant
Sr. Software engineer
Innovisto Pvt. Ltd.
(Supervisor)

…………………..........................
(External Examiner)

…………………..........................
(Internal Examiner)

iv

ACKNOWLEDGEMENT

It is a great pleasure to have the opportunity to extend our heartfelt gratitude to everyone who helped us throughout the course of this project. We are profoundly grateful to our supervisor **Mr. Ashok Kumar Pant**, Sr. Software Engineer of Innovisto Pvt. Ltd., for his expert guidance, continuous encouragement and ever willingness to spare time from his otherwise busy schedule for the project's progress reviews. His continuous inspiration has made us complete this project and achieve its target.

We would also like to express our deepest appreciation to **Mr. Sushant Poudel** Head of Department, Kathford International College of Engineering and Management, Department of Computer Science and Information Technology, for his constant motivation, support and for providing us with a suitable working environment. We would also like to extend our sincere regards to **Ms. Deni Shahi** and all the faculty members for their support and encouragement.

At last our special thanks go to all staff members of BSc CSIT department who directly and indirectly extended their hands in making this project works a success.

ABSTRACT

Facial Expression conveys non-verbal cues, which plays an important roles in interpersonal relations. The Facial Expression Recognition system is the process of identifying the emotional state of a person. In this system captured image is compared with the trained dataset available in database and then emotional state of the image will be displayed.

This system is based on image processing and machine learning. For designing a robust facial feature descriptor, we apply the Local Binary Pattern. Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. The histogram will be formed by using the operator label of LBP.

The recognition performance of the proposed method will be evaluated by using the trained database with the help of Support Vector Machine. Experimental results with prototypic expressions show the superiority of the LBP descriptor against some well-known appearance-based feature representation methods.

We evaluate our proposed method on the JAFFE and COHN-KANADE dataset. The Precision, Recall and Fscore from the COHN-KANADE dataset were 83.6142%, 95.0822% and 88.9955% respectively and that of JAFFE dataset were 91.8986%, 98.3649%, 95.0218% respectively. Experimental results demonstrate the competitive classification accuracy of our proposed method.

Keywords:

Facial expression recognition (FER), Local Binary pattern (LBP), Support Vector Machine (SVM)

List of Figures

List of Tables:

Table of Contents

CHAPTER 1

1. INTRODUCTION

A Facial expression is the visible manifestation of the affective state, cognitive activity, intention, personality and psychopathology of a person and plays a communicative role in interpersonal relations. It have been studied for a long period of time and obtaining the progress recent decades. Though much progress has been made, recognizing facial expression with a high accuracy remains to be difficult due to the complexity and varieties of facial expressions [2].

Generally human beings can convey intentions and emotions through nonverbal ways such as gestures, facial expressions and involuntary languages. This system can be significantly useful, nonverbal way for people to communicate with each other. The important thing is how fluently the system detects or extracts the facial expression from image. The system is growing attention because this could be widely used in many fields like lie detection, medical assessment and human computer interface. The Facial Action Coding System (FACS), which was proposed in 1978 by Ekman and refined in 2002, is a very popular facial expression analysis tool [3].

On a day to day basics humans commonly recognize emotions by characteristic features, displayed as a part of a facial expression. For instance happiness is undeniably associated with a smile or an upward movement of the corners of the lips. Similarly other emotions are characterized by other deformations typical to a particular expression. Research into automatic recognition of facial expressions addresses the problems surrounding the representation and categorization of static or dynamic characteristics of these deformations of face pigmentation [8].

The system classifies facial expression of the same person into the basic emotions namely anger, disgust, fear, happiness, sadness and surprise. The main purpose of this system is efficient interaction between human beings and machines using eye gaze, facial expressions, cognitive modeling etc. Here, detection and classification of facial

expressions can be used as a natural way for the interaction between man and machine. And the system intensity vary from person to person and also varies along with age, gender, size and shape of face, and further, even the expressions of the same person do not remain constant with time.

However, the inherent variability of facial images caused by different factors like variations in illumination, pose, alignment, occlusions makes expression recognition a challenging task. Some surveys on facial feature representations for face recognition and expression analysis addressed these challenges and possible solutions in detail [5].

1.1. Motivation

In today's networked world the need to maintain security of information or physical property is becoming both increasingly important and increasingly difficult. In countries like Nepal the rate of crimes are increasing day by day. No automatic systems are there that can track person's activity. If we will be able to track Facial expressions of persons automatically then we can find the criminal easily since facial expressions changes doing different activities. So we decided to make a Facial Expression Recognition System.

We are interested in this project after we went through few papers in this area. The papers were published as per their system creation and way of creating the system for accurate and reliable facial expression recognition system.

As a result we are highly motivated to develop a system that recognizes facial expression and track one person's activity.

1.2. Problem Statement

Human emotions and intentions are expressed through facial expressions and deriving an efficient and effective feature is the fundamental component of facial expression system. Face recognition is important for the interpretation of facial expressions in applications such as intelligent, man-machine interface and communication, intelligent visual surveillance, teleconference and real-time animation from live motion images. The facial expressions are useful for efficient interaction Most research and system in facial expression recognition are limited to six basic expressions (joy, sad, anger, disgust, fear,

surprise). It is found that it is insufficient to describe all facial expressions and these expressions are categorized based on facial actions [7].

Detecting face and recognizing the facial expression is a very complicated task when it is a vital to pay attention to primary components like: face configuration, orientation, location where the face is set.

1.3. Objectives

1. To develop a facial expression recognition system.
2. To experiment machine learning algorithm in computer vision fields.
3. To detect emotion thus facilitating Intelligent Human-Computer Interaction.

1.4. Scope and Applications

The scope of this system is to tackle with the problems that can arise in day to day life. Some of the scopes are:

1. The system can be used to detect and track a user's state of mind.
2. The system can be used in mini-marts, shopping center to view the feedback of the customers to enhance the business,
3. The system can be installed at busy places like airport, railway station or bus station for detecting human faces and facial expressions of each person. If there are any faces that appeared suspicious like angry or fearful, the system might set an internal alarm.
4. The system can also be used for educational purpose such as one can get feedback on how the student is reacting during the class.
5. This system can be used for lie detection amongst criminal suspects during interrogation
6. This system can help people in emotion related -research to improve the processing of emotion data.
7. Clever marketing is feasible using emotional knowledge of a person which can be identified by this system.

CHAPTER 2

2. REQUIREMENT ANALYSIS

2.1.   Planning

In planning phase study of reliable and effective algorithms is done. On the other hand data were collected and were preprocessed for more fine and accurate results. Since huge amount of data were needed for better accuracy we have collected the data surfing the internet. Since, we are new to this project we have decided to use local binary pattern algorithm for feature extraction and support vector machine for training the dataset. We have decided to implement these algorithms by using OpenCv framework.

2.2.   Literature Reviews

Research in the fields of face detection and tracking has been very active and there is exhaustive literature available on the same. The major challenge that the researchers face is the non-availability of spontaneous expression data [1]. Capturing spontaneous expressions on images and video is one of the biggest challenges ahead [2]. Many attempts have been made to recognize facial expressions. Zhang et al investigated two types of features, the geometry-based features and Gabor wavelets based features, for facial expression recognition.

Appearance based methods, feature invariant methods, knowledge based methods, Template based methods are the face detection strategies whereas Local Binary Pattern phase correlation, Haar classifier, AdaBoost, Gabor Wavelet are the expression detection strategies in related field [3]. Face reader is the premier for automatic analysis of facial expression recognition and Emotient, Affectiva, Karios etc are some of the API's for expression recognition. Automatic facial expression recognition includes two vital aspects: facial feature representation and classifier problem [2].

Facial feature representation is to extract a set of appropriate features from original face images for describing faces. Histogram of Oriented Gradient (HOG), SIFT, Gabbor Fitters and Local Binary Pattern (LBP) are the algorithms used for facial feature representation [3,4]. LBP is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. The operator labels the pixels of an image by thresholding the 3X3 neighborhood of each pixel with the center value and considering the result as a binary number [3]. HOG was first proposed by Dalal and Triggs in 2005. HOG numerates the appearance of gradient orientation in a local path of an image.

For classifier problem we use algorithms like Machine learning, Neural Network, Support Vector Machine, Deep learning, Naive Bayes. The formation of histogram by using any of facial feature representation will use Support Vector Machine (SVM) for expression recognition. SVM builds a hyperplane to separate the high dimensional space. An ideal separation is achieved when the distance between the hyper plane and the training data of any class is the largest [4].

The size of the block for the LBP feature extraction is chosen for higher recognition accuracy. The testing results indicate that by using LBP features facial expressions recognition accuracy is more than 97%. The block LBP histogram features extract local as well as global features of face image resulting higher accuracy. LBP is compatible with various classifiers, filters etc. [3].

2.3.   Data collection

Some of the public databases to evaluate the facial expression recognition algorithms are:

2.3.1. COHN-KANADE AU Coded Facial Expression Database

Subjects in the released portion of the COHN-KANADE AU-Coded Facial Expression Database are 100 university students. They ranged in age from 18 to 30 years. Sixty-five percent were female, 15 percent were African-American, and three percent were Asian or Latino. Subjects were instructed by an experimenter to perform a series of 23 facial displays that included single action units and combinations of action units. Image

sequences from neutral to target display were digitized into 640 by 480 or 490 pixel arrays with 8-bit precision for grayscale values. Included with the image files are "sequence" files; these are short text files that describe the order in which images should be read. The seven expressions are angry, surprise, contempt, fear, and disgust [4]. Fig.1 shows the 8 expressions with each from a different subject.
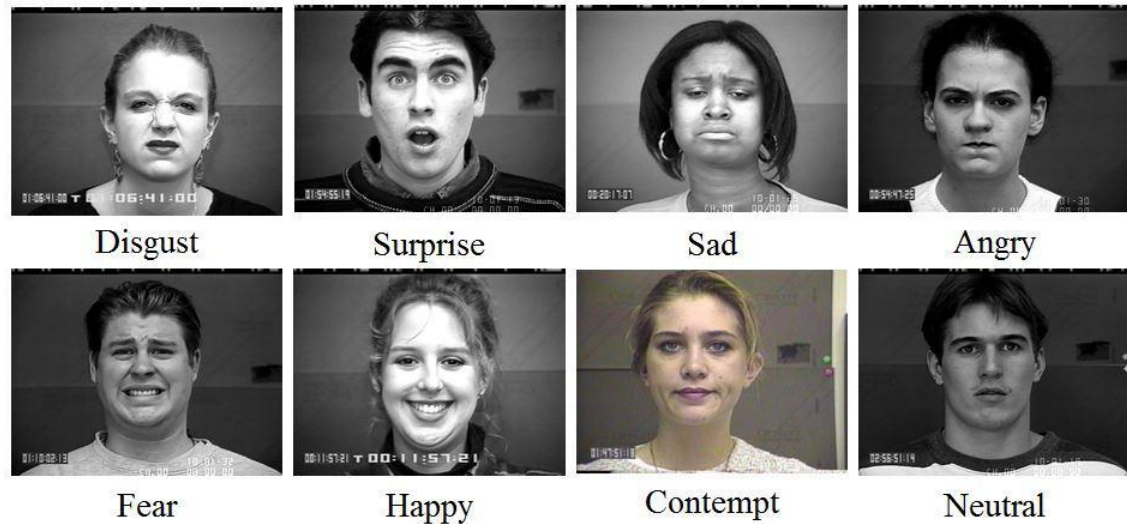

Figure 1: The eight expression from one subject

2.3.2. Japanese Female Facial Expression (JAFFE) Database

This database contains 213 images in total. There are 10 subjects and 7 facial expressions for each subject. Each subject has about twenty images and each expression includes two to three images. The seven expressions are angry, happy, disgust, sadness, surprise, fear and neutral respectively [4]. Fig.2 shows the seven expressions from one subject.

Figure 2: The seven expression from one subject

6

Table 1: Data Collections

Database	Sample Details	Available Descriptions
COHN-KANADE Database (also known as CMU-Pittsburg data-base) [1].	• 585 image sequences from 97 subjects • Age: 18 to 30 years • Gender: 65% female • Ethnicity: 15% African - Americans and 3% of Asians and Latinos	"Annotation of FACS Action Units and emotion-specified expressions"
The Japanese Female Facial Expression (JAFFEE) Data-base [1].	• 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) • 10 Japanese female models.	"Each image has been rated on 6 emotion adjectives by 92 Japanese subjects"

2.4. Dataset Preparation

The proposed system was trained and tested using two datasets namely COHN-KANADE and JAFFE. The COHN-KANADE dataset consists of 500 images sequences from 100 subjects whereas the JAFFE dataset consists of 213 images. For our experiment we have used 6481 images for training and 1619 images for testing fromm different subjects of Cohn-Kanade dataset. Similarly, 107 images were used for training and 106 images were used for testing from JAFFE dataset.

We normalized the faces to 72 pixels. Based on the structure of face facial images of 256*256 pixels were cropped from original images. To identify the facial image automatic face detection was performed by using the face detector of our own system based on Haar classifier. From the results of face detection including face location, face width and face height were automatically created. Finally images were cropped in accordance to the result given by the face detector and further cropped images were used for training and testing.

7

Figures below show the original image and cropped image:
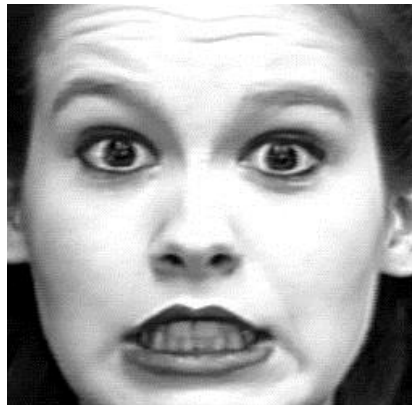

Figure 3: Original Image

Figure 4: Cropped Image

2.5.   Software Requirement Specification:

Requirement analysis is mainly categorized into two types:

2.5.1. Functional requirements:

The functional requirements for a system describe what the system should do. Those requirements depend on the type of software being developed, the expected users of the software. These are statement of services the system should provide, how the system should react to particular inputs and how the system should behave in particular situation.

2.5.2. Non-Functional requirements:

Nonfunctional requirements are requirements that are not directly concerned with the specified function delivered by the system. They may relate to emergent system properties such as reliability, response time and store occupancy. Some of the nonfunctional requirements related with this system are hereby below:

a) Reliability:

Reliability based on this system defines the evaluation result of the system, correct identification of the facial expressions and maximum evaluation rate of the facial expression recognition of any input images.

b) Ease of Use:

The system is simple, user friendly, graphics user interface implemented so any can use this system without any difficulties.

2.6.   Feasibility Study

Before starting the project, feasibility study is carried out to measure the viable of the system. Feasibility study is necessary to determine if creating a new or improved system is friendly with the cost, benefits, operation, technology and time. Following feasibility study is given as below:

2.6.1. Technical Feasibility

Technical feasibility is one of the first studies that must be conducted after the project has been identified. Technical feasibility study includes the hardware and software devices. The required technologies (C++ language and CLion IDE) existed.

2.6.2. Operational Feasibility

Operational Feasibility is a measure of how well a proposed system solves the problem and takes advantage of the opportunities identified during scope definition. The following points were considered for the project's technical feasibility:

- The system will detect and capture the image of face.

- The captured image is then (identified which category)

2.6.3. Economic Feasibility

The purpose of economic feasibility is to determine the positive economic benefits that include quantification and identification. The system is economically feasible due to availability of all requirements such as collection of data from

- JAFFE
- COHN-KANADE

2.6.4. Schedule Feasibility

Schedule feasibility is a measure of how reasonable the project timetable is. The system is found schedule feasible because the system is designed in such a way that it will finish prescribed time.

2.7. Software and Hardware Requirement

2.7.1. Software Requirement

Following are the software requirement necessary of the project:

a) C++ programming language
b) CLion IDE (selective)
c) OpenCV framework
d) Linux platform Ubuntu OS

2.7.2. Hardware Requirement

Following are the hardware requirement that is most important for the project:

a) Fluently working Laptops
b) RAM minimum 4Gb
c) Web Camera

CHAPTER 3

3. PROJECT METHODOLOGY

3.1. System Design

System design shows the overall design of system. In this section we discuss in detail the design aspects of the system:
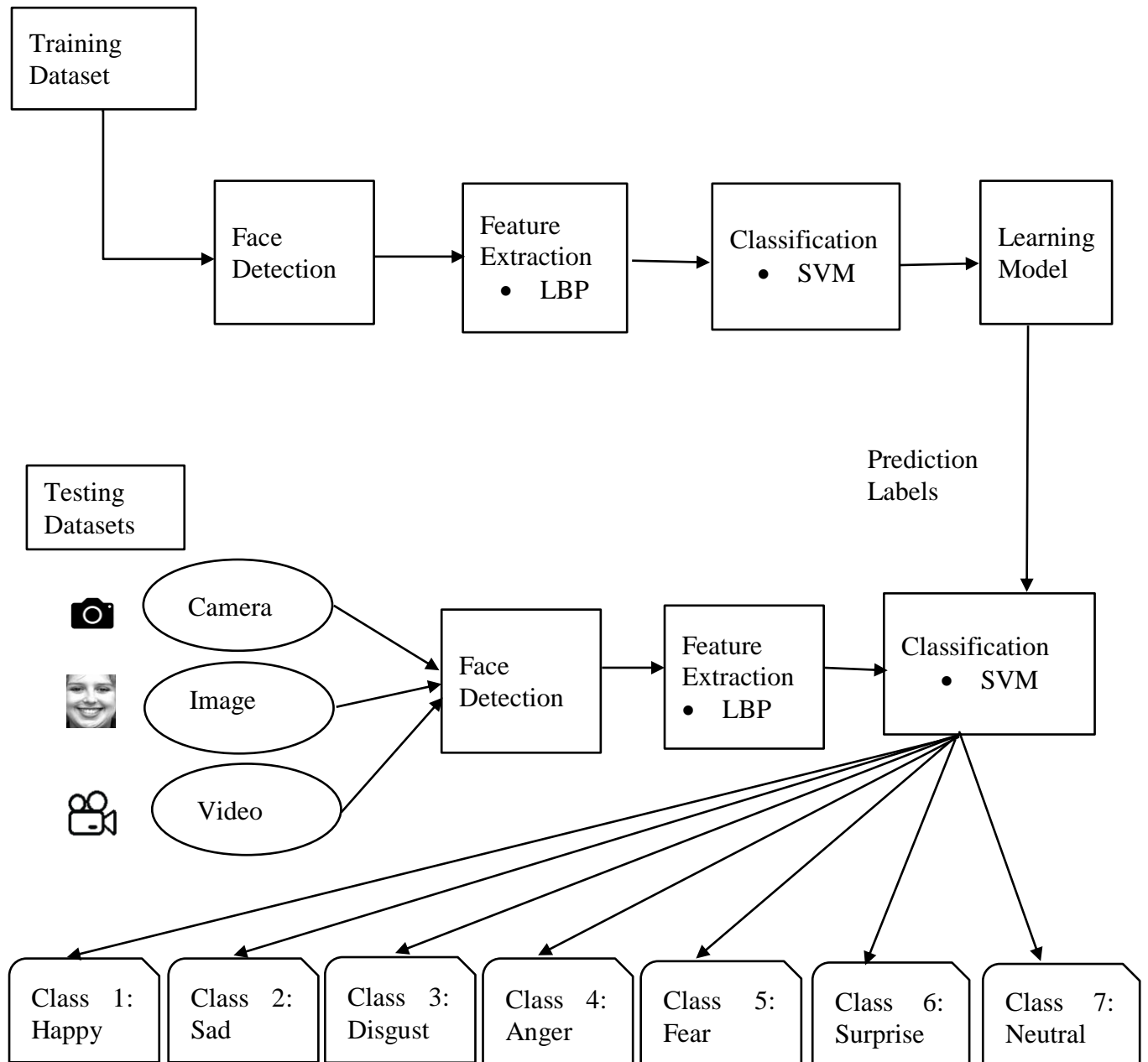
3.1.1. System Diagram

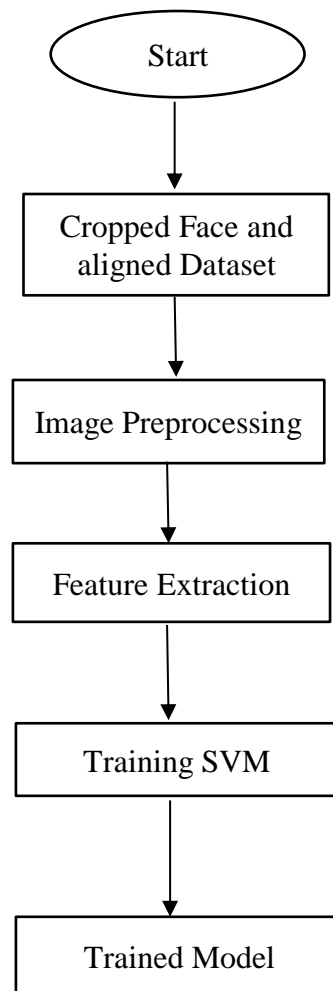

Figure 5: System Diagram
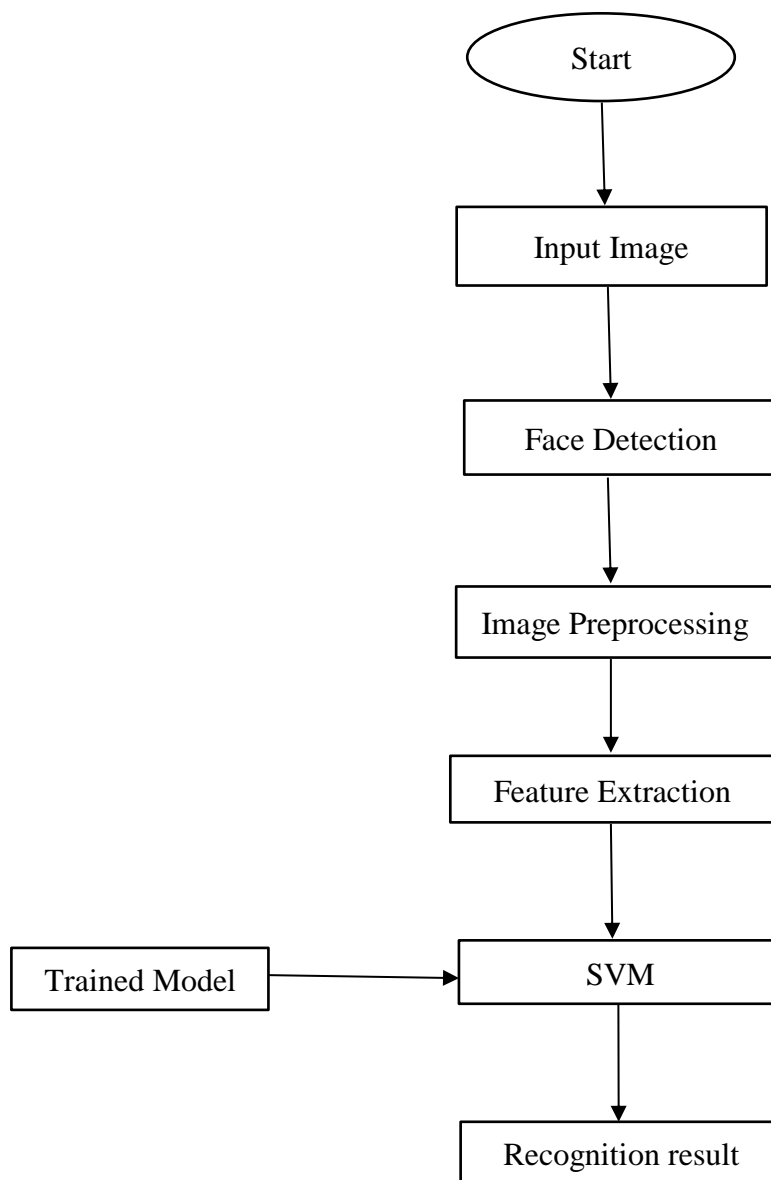
3.1.2. System Flowchart

Figure 6: Flowchart of Training

```
                    ┌─────────────┐
                    │    Start    │
                    └──────┬──────┘
                           │
                           ▼
                    ┌─────────────┐
                    │ Input Image │
                    └──────┬──────┘
                           │
                           ▼
                    ┌──────────────┐
                    │ Face Detection│
                    └──────┬───────┘
                           │
                           ▼
                  ┌──────────────────┐
                  │ Image Preprocessing│
                  └────────┬─────────┘
                           │
                           ▼
                  ┌──────────────────┐
                  │ Feature Extraction│
                  └────────┬─────────┘
                           │
┌──────────────┐          ▼
│ Trained Model│────────►┌──────┐
└──────────────┘         │  SVM │
                         └───┬──┘
                             │
                             ▼
                    ┌──────────────────┐
                    │ Recognition result│
                    └──────────────────┘
```

Figure 7: Flowchart of Testing/Prediction

3.1.3. Class Diagram

FersCore

TRAIN_FILE: String
TEST_FILE:String
LABEL_FILE:String
MEAN_FILE: String
NUM_FEATURES_PER_GRIDS:Int
NUM_GRIDS:Size
Image_Size:Size

init()
setLabels()
train()
evaluate()
predict()
predictLabel()
calculateMean()

1..* ——— 1

Fers

commands: string
modelPath: string
params: string
windowName: string
image: string
file: string
video: string
camera: int
display: bool
shuff: bool
detectModelFile: string
pause:int

displayDetections( )
fileMode( )
videoMode( )
cameraMode( )
imageMode( )

Evaluation

_avgAccuracy:double
_errRate:double
_precisionMicro:double
_recallMicro:double
_fScoreMicro:double

Evaluation ()

1   1   1   1..*

1

1..*

1

FeatureExtraction

Img:mat
Hist:mat
spatialHist:int
numPatterns:int
grid:size
images:<string,int>

convert()
computeLbpFeatures()
LBP()
Histogram()
spatialHistogram()

Confusion

_classes: int
_samples: int
_c: double
_per: double
_ind: string
_cm: int

convertToBooleanMatrix( )
confusion( )

1

Detector

Cascade_:string
imgNewW:int
imgNewH_: int
minWH_: Size
scale_: double
minNeighbours_: int
padW_: int
padH_: int
cascadeFile_: string

detect( )

Figure 8: Class Diagram

14

3.1.4. Sequence Diagram
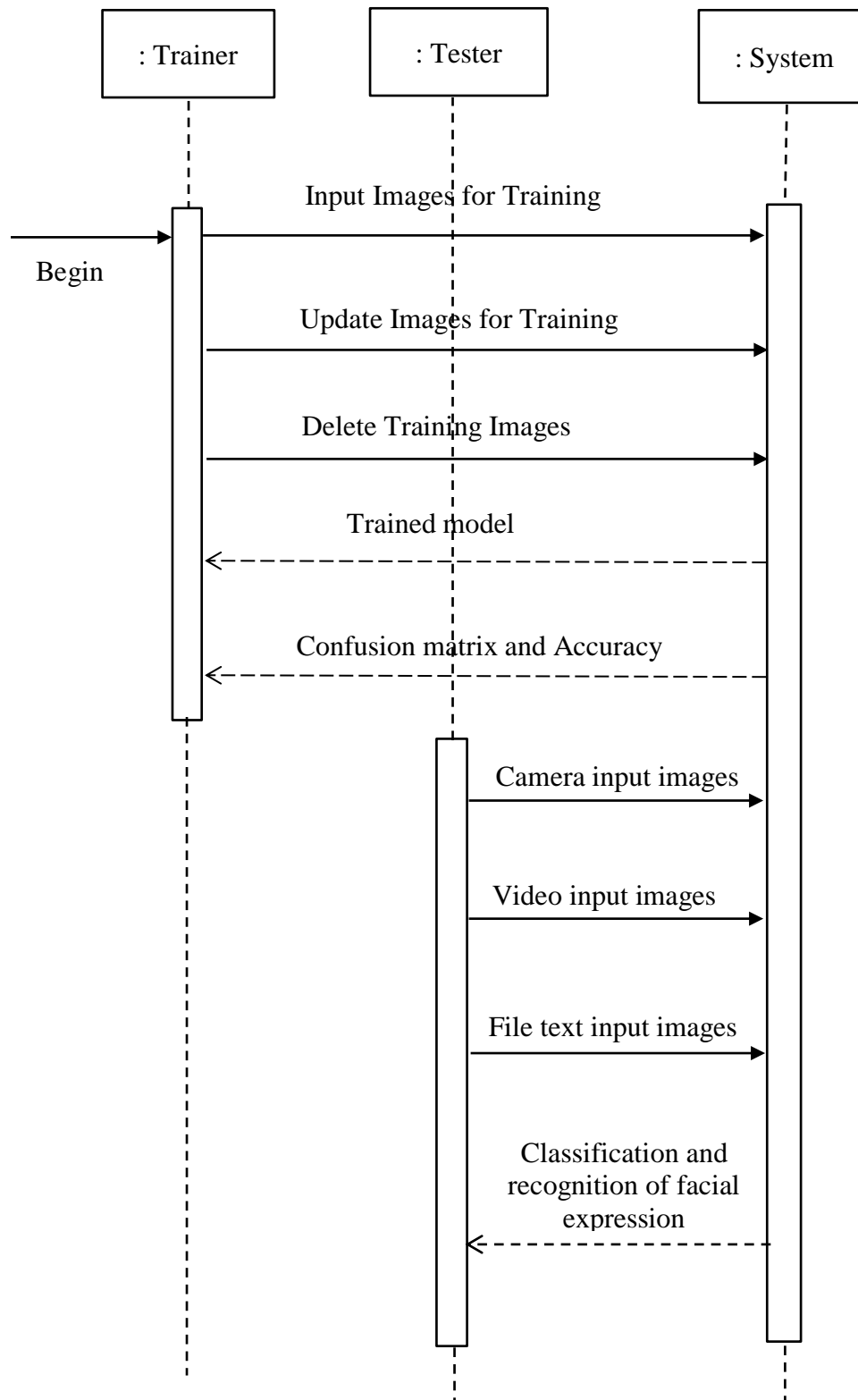
Figure 9: Sequence Diagram

3.2. Phases in Facial Expression Recognition

The facial expression recognition system is trained using supervised learning approach in which it takes images of different facial expressions. The system includes the training and testing phase followed by image acquisition, face detection, image preprocessing, feature extraction and classification. Face detection and feature extraction are carried out from face images and then classified into six classes belonging to six basic expressions which are outlined below:

3.2.1. Image Acquisition

Images used for facial expression recognition are static images or image sequences. Images of face can be captured using camera.

3.2.2. Face detection

Face Detection is useful in detection of facial image. Face Detection is carried out in training dataset using Haar classifier called Voila-Jones face detector and implemented through Opencv. Haar like features encodes the difference in average intensity in different parts of the image and consists of black and white connected rectangles in which the value of the feature is the difference of sum of pixel values in black and white regions [6].

3.2.3. Image Pre-processing

Image pre-processing includes the removal of noise and normalization against the variation of pixel position or brightness.

  a) Color Normalization
  b) Histogram Normalization

3.2.4. Feature Extraction

Selection of the feature vector is the most important part in a pattern classification problem. The image of face after pre-processing is then used for extracting the important features. The inherent problems related to image classification include the scale, pose,

translation and variations in illumination level [6]. The important features are extracted using LBP algorithm which is described below:

3.2.4.1. Local Binary Pattern

LBP is the feature extraction technique. The original LBP operator points the pixels of an image with decimal numbers, which are called LBPs or LBP codes that encode the local structure around each pixel. Each pixel is compared with its eight neighbors in a 3 X 3 neighborhood by subtracting the center pixel value. In the result, negative values are encoded with 0 and the others with 1. For each given pixel, a binary number is obtained by merging all these binary values in a clockwise direction, which starts from the one of its top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel. The derived binary numbers are referred to be the LBPs or LBP codes [6].
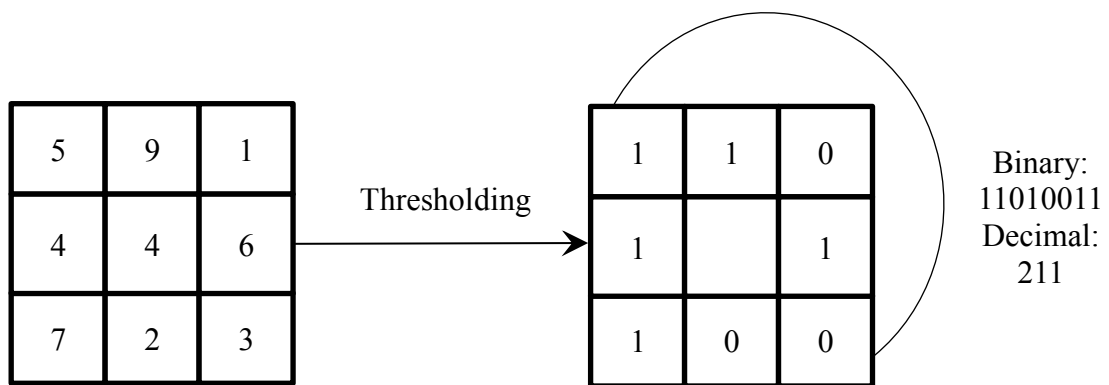
Figure 10: The Basic LBP Operator

(P=8,R=1.0)   (P=12,R=1.5)
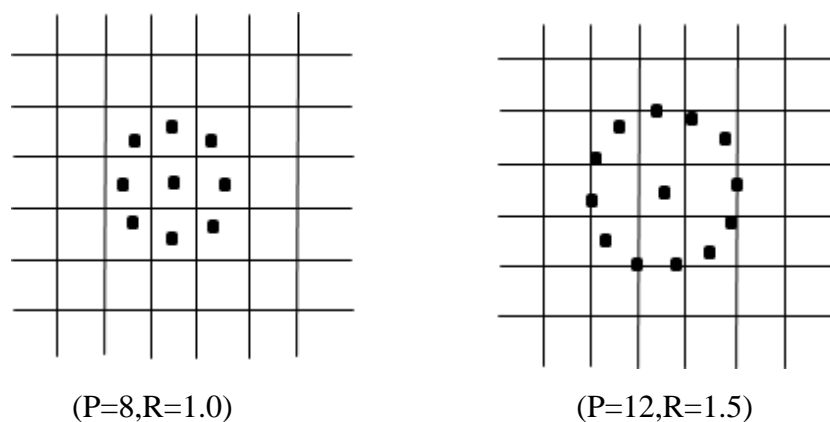
Figure 11: Two examples of extended LBP

17

The limitation of the basic LBP operator is that its small 3×3 neighborhood cannot capture the dominant features with large scale structures. As a result, to deal with the texture at different scales, the operator was later extended to use neighborhoods of different sizes [7]. Using circular neighborhoods and bilinearly interpolating the pixel values allow any radius and number of pixel in the neighborhood. Examples of the extended LBP are shown above (Figure 5.3.2), where (P, R) denotes sampling points on a circle of radius of R.

Further extension of LBP is to user uniform patterns. A LBP is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. E.g.00000000, 001110000 and 11100001 are uniform patterns. A histogram of a labelled image f1(x, y) can be defined as

$$Hi = \sum_{x,y} I\left(f_l(x, y) = i\right), \quad i = 0, \dots, n - 1 \qquad (1)$$

Where n is the number of different labels produced by the LBP operator and

$$I(A) = \begin{cases} 1 \text{ A is true} \\ 0 \text{ A is false} \end{cases} \qquad (2)$$

This histogram contains information about the distribution of the local micro-patterns, such as edges, spots and flat areas, over the whole image. For efficient face representation, feature extracted should retain also spatial information. Hence, face image is divided into m small regions R0, R1,…,Rm and a spatially enhanced histogram is defined as [2]

$$Hi = \sum_{x,y} I\left(f_l(x, y) = i\right) I\left((x, y) \epsilon R_j\right) \qquad (3)$$

3.2.5. Classification

The dimensionality of data obtained from the feature extraction method is very high so it is reduced using classification. Features should take different values for object belonging to different class so classification will be done using Support Vector Machine algorithm.

3.2.5.1. Support Vector Machines

SVM is widely used in various pattern recognition tasks. SVM is a state-of-the-art machine learning approach based on the modern statistical learning theory. SVM can achieve a near optimum separation among classes. SVMs is trained to perform facial expression classification using the features proposed. In general, SVM are the maximal hyperplane classification method that relies on results from statistical learning theory to guarantee high generalization performance.

Kernel functions are employed to efficiently map input data which may not be linearly separable to a high dimensional feature space where linear methods can then be applied. SVMs exhibit good classification accuracy even when only a modest amount of training data is available, making them particularly suitable to a dynamic, interactive approach to expression recognition [10].

An ideal separation is achieved when the hyper plane and the training data of any class is the largest. This separating hyper plane works as the decision surface. SVM has been successfully employed for a number of classification tasks such as text categorization, genetic analysis and face detection [11].

Given a training set of labeled samples:

$$D = \{(x_i, y_i | x_i \in R^n, y_i \in \{-1,1\}\}_{i=1}^{p} \qquad (1)$$

A SVM tries to find a hyperplane to distinguish the samples with the smallest errors.

$$w.x - b = 0 \qquad (2)$$

For an input vector xi, the classification is achieved by computing the distance from the input vector to the hyperplane. The original SVM is a binary classifier [4].

3.2.6. System Evaluation

Evaluation of the system can be done using following methods:

a) Precision

Precision estimates the predictive value of a label, either positive or negative, depending on the class for which it is calculated; in other words, it assesses the predictive power of the algorithm. Precision is the percentage of correctly assigned expressions in relation to the total number of aspects.

$$precision = \frac{tp}{tp + fp} \qquad (1)$$

b) Recall

Recall is a function of its correctly classified examples (true positives) and its misclassified examples (false negatives). Recall is the percentage of correctly assigned expressions in relation to the total number of expressions.

$$recall = \frac{tp}{tp + fn} \qquad (2)$$

c) F-score

F-score is a composite measure which benefits algorithms with higher sensitivity and challenges algorithms with higher specificity. The F-score is evenly balanced when $\beta = 1$. It favours precision when $\beta > 1$, and recall otherwise.

$$F - measure = \frac{(\beta^2 + 1) * precision * recall}{\beta^2 * precision * recall} \qquad (3)$$

All three measures distinguish the correct classification of labels within different classes. They concentrate on one class (positive examples). Hence, precision and recall do measure different properties and we therefore need a combined quality measure in order to determine the best matching aspect to expression category mappings. The so called F-measure fm computes the harmonic mean of precision and recall and allows taking into account both properties at the same time [9]. Note that the overall recall recallov is also known as accuracy.

CHAPTER 4

4. DEVELOPMENT AND TESTING

4.1. Implementation Tools

4.1.1. Programming Language and Coding Tools

a) C++

C++ is a general-purpose programming language. It has imperative, object-oriented and generic programming features, while also providing facilities for low-level memory manipulation. It was designed with a bias toward system programming and embedded, resource-constrained and large systems, with performance, efficiency and flexibility of use as its design highlights. C++ has also been found useful in many other contexts, with key strengths being software infrastructure and resource-constrained applications, including desktop applications, servers (e.g. e-commerce, web search or SQL servers), and performance-critical applications (e.g. telephone switches or space probes). C++ is a compiled language, with implementations of it available on many platforms and provided by various organizations, including the Free Software Foundation (FSF's GCC), LLVM, Microsoft, Intel and IBM.

b) IDE CLion for C++

CLion is a cross-platform C/C++ IDE which is more than just an editor as it offers intelligent Cmake support. CLion helps in knowing codes through and through and can boost the productivity with smart and relevant code completion, instant navigation and reliable refactoring.

4.1.2. Framework

a) OpenCV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the

commercial products. Being a BSD-licensed product, OpenCV makes it easy for businesses to utilize and modify the code.

The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc. It has C++, C, Python, Java and MATLAB interfaces and supports Windows, Linux, Android and Mac OS.

4.2. System Testing

System testing was done by giving different training and testing datasets. This test was done to evaluate whether the system was predicting accurate result or not. During the phase of the development of the system our system was tested time and again.
The series of testing conducted are as follows:

4.2.1. Unit Testing

In unit testing, we designed the whole system in modularized pattern and each module was tested. Till we get the accurate output from the individual module we worked on the same module.

4.2.2. Integration Testing

After constructing individual modules all the modules were merged and a complete system was made. Then the system was tested whether the prediction given by training dataset to testing set was correct or not. We tried to meet the accuracy as higher as much as we can get. After spending a couple of days in integration testing the average accuracy of our system was 91%.

4.2.2.1. Alpha testing

Alpha testing is the first stage of software engineering which is considered as a simulated or actual operational testing done by the individual member of the project. Alpha testing is conducted by the project developers, in context of our project.

4.2.2.2. Beta Testing

Beta testing comes continuously after alpha testing which is considered as a form of external user acceptance testing. The beta version of the program is developed to and provided to limited audience. This is the final test process in the case of this project. In this system the beta-testing is done by our colleagues and the project supervisor.

CHAPTER 5

5. EXPERIMENTATION AND RESULTS

The aim of this project work is to develop a complete facial expression recognition system. Two datasets, COHN_KANADE and JAFFE were used for the experimentations. First of all, system was trained using different random samples in each dataset by supervised learning. In each datasets the data were partitioned into two parts for training and testing. Every dataset have completely different samples which are selected randomly in uniform manner from the pool of given dataset. The COHN_KANADE datasets included 585 directories of both subject and session where there were 97 subject directories and 8795 image files in total and partitioned was made in the ratio of 8:2 i.e. 6481 (80%) for train and 1619 (20%) for test. Similarly, JAFFE dataset included 213 images which was partitioned in the ratio of 7.5:2.5 i.e. 160 (75%) for train and 53 (25%) for test.

The confusion and accuracy evaluation results of COHN-KANADE and JAFFE datasets are as below:

Table 2: Confusion matrix of COHN-KANADE

Labels	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	259	0	0	0	0	1	0
Disgust	1	182	0	0	0	0	0
Fear	2	1	219	0	0	0	1
Happy	25	40	173	98	1	19	0
Neutral	1	1	12	0	111	0	0
Sad	1	1	1	1	0	228	0
Surprise	12	15	141	1	0	11	60

In the above table, row shows the actual classes and column shows the predicted classes. The classifier made a total of 1619 predictions where the classifier predicted angry for

300 times , disgust for 239 times, fear for 545 times, happy for 99 times, neutral for 112 times, sad for 259 times and surprise for 61 times. Whereas in reality 260 cases was angry, 183 was disgust, 223 was fear, 356 was happy, 125 was neutral, 228 was sad and 240 was surprise.

Table 3: Accuracy of COHN-KANADE

Evaluation Types	Results Percentages
Precision	83.6412
Recall	95.0822
F-score	88.9955

The above table shows that 83.6412% of the expressions were predicted 95.0822% of the expressions were correctly assigned. The harmonic mean of precision and recall was 88.9955%.

Table 4: Confusion matrix of JAFFE

Labels	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	4	1	0	0	0	1	0
Disgust	0	6	0	0	0	0	0
Fear	0	0	10	0	0	0	0
Happy	0	0	0	10	2	0	0
Neutral	0	0	0	0	6	0	0
Sad	0	0	0	0	0	10	0
Surprise	0	0	0	0	1	0	2

In the above table, row shows the actual classes and column shows the predicted classes. The classifier made a total of 53 predictions where the classifier predicted angry for 4 times, disgust for 7 times, fear for 10 times, happy for 10 times, neutral for 9 times, sad for 11 times and surprise for 2 times. Whereas in reality 6 cases were angry, 6 was disgust, 10 was fear, 12 was happy, 6 was neutral, 10 was sad and 3 was surprise.

Table 5: Accuracy of JAFFE

Evaluation Types	Percentages
Precision	91.8986
Recall	98.3649
F-score	95.0218

The above table shows that 91.8986% of the expressions were predicted, 98.3649% of the expressions were correctly assigned. The harmonic mean of precision and recall was 95.0218%.

CHAPTER 6

6. CONCLUSION AND RECOMMENDATION

6.1. Conclusion

This project proposes an approach for recognizing the category of facial expressions. Face Detection and Extraction of expressions from facial images is useful in many applications, such as robotics vision, video surveillance, digital cameras, security and human-computer interaction. This project's objective was to develop a facial expression recognition system implementing the computer visions and enhancing the advanced feature extraction and classification in face expression recognition.

In this project, seven different facial expressions of different persons' images from different datasets have been analyzed. This project involves facial expression preprocessing of captured facial images followed by feature extraction using feature extraction using Local Binary Patterns and classification of facial expressions based on training of datasets of facial images based on Support Vector Machines. This project recognizes more facial expressions based on JAFFE, COHN-KANADE face database. To measure the performance of proposed algorithm and methods and check the results accuracy, the system has been evaluated using Precision, Recall and Fscore. The same datasets were used for both training and testing by dividing the datasets into training samples and testing samples in the ratio of 8:2 of COHN-KANADE and 7.5:2.5 of JAFFE. The Precision, Recall and Fscore from the COHN-KANADE dataset were 83.6142%, 95.0822% and 88.9955% respectively and JAFFE dataset were 91.8986%, 98.3649% and 95.0218% respectively.

Experiment results on two databases, JAFFE and the COHN-KANADE dataset, show that our proposed method can achieve a good performance. Facial expression recognition is a very challenging problem. More efforts should be made to improve the classification performance for important applications. Our future work will focus on improving the performance of the system and deriving more appropriate classifications which may be useful in many real world applications.

6.2.   Future Scope

Face expression recognition systems have improved a lot over the past decade. The focus has definitely shifted from posed expression recognition to spontaneous expression recognition. Promising results can be obtained under face registration errors, fast processing time, and high correct recognition rate (CRR) and significant performance improvements can be obtained in our system. System is fully automatic and has the capability to work with images feed. It is able to recognize spontaneous expressions. Our system can be used in Digital Cameras wherein the image can be captured only when the person smiles. In security systems which can identify a person, in any form of expression he presents himself. Rooms in homes can set the lights, television to a person's taste when they enter the room. Doctors can use the system to understand the intensity of pain or illness of a deaf patient. Our system can be used to detect and track a user's state of mind, and in mini-marts, shopping center to view the feedback of the customers to enhance the business etc.

References
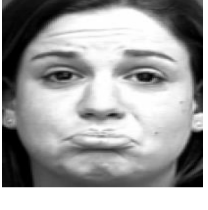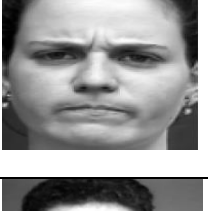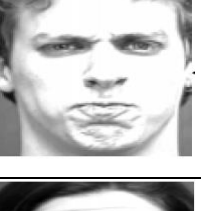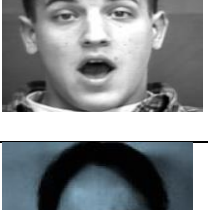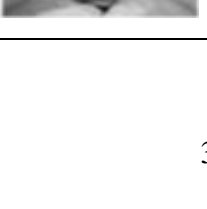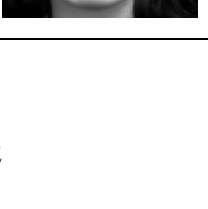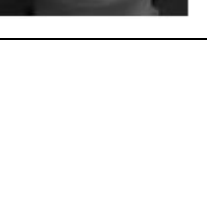
[1]   Bettadapura, V. (2012). Face expression recognition and analysis: the state of the art. *arXiv preprint arXiv:1203.6722*.

[2]   Shan, C., Gong, S., & McOwan, P. W. (2005, September). Robust facial expression recognition using local binary patterns. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on* (Vol. 2, pp. II-370). IEEE.

[3]   Bhatt, M., Drashti, H., Rathod, M., Kirit, R., Agravat, M., & Shardul, J. (2014). A Studyof Local Binary Pattern Method for Facial Expression Detection. *arXiv preprint arXiv:1405.6130*.

[4]   Chen, J., Chen, Z., Chi, Z., & Fu, H. (2014, August). Facial expression recognition based on facial components detection and hog features. In *International Workshops on Electrical and Computer Engineering Subfields* (pp. 884-888).

[5]   Ahmed, F., Bari, H., & Hossain, E. (2014). Person-independent facial expression recognition based on compound local binary pattern (CLBP). *Int. Arab J. Inf. Technol.*, *11*(2), 195-203.

[6]   Happy, S. L., George, A., & Routray, A. (2012, December). A real time facial expression classification system using Local Binary Patterns. In *Intelligent Human Computer Interaction (IHCI), 2012 4th International Conference on* (pp. 1-5). IEEE.

[7]   Zhang, S., Zhao, X., & Lei, B. (2012). Facial expression recognition based on local binary patterns and local fisher discriminant analysis. *WSEAS Trans. Signal Process*, *8*(1), 21-31.

[8]   Chibelushi, C. C., & Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. CVonline: On-Line Compendium of Computer Vision, 9.

[9]   Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006, December). Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In Australasian Joint Conference on Artificial Intelligence (pp. 1015-1021). Springer Berlin Heidelberg.

[10]   Michel, P., & El Kaliouby, R. (2005). Facial expression recognition using support vector machines. In The 10th International Conference on Human-Computer Interaction, Crete, Greece.

[11]   Michel, P., & El Kaliouby, R. (2003, November). Real time facial expression recognition in video using support vector machines. In Proceedings of the 5th international conference on Multimodal interfaces (pp. 258-264). ACM.

Appendix

Datasets Collection

Table 6: Dataset images of facial recognition

Happy			
Sad			
Angry			
Surprise			
Disgust			
Fear			
Neutral			

Experimental Demonstration

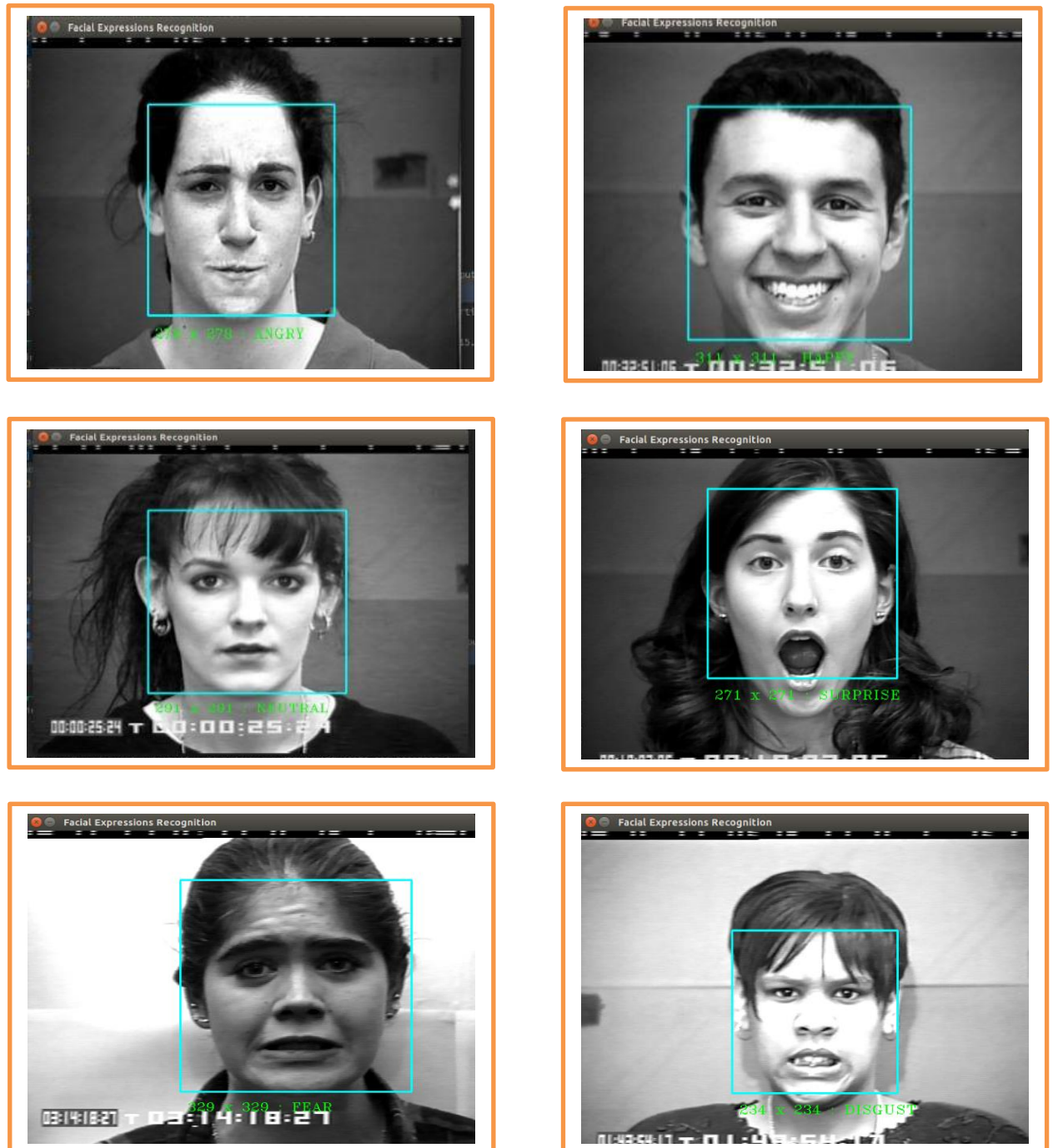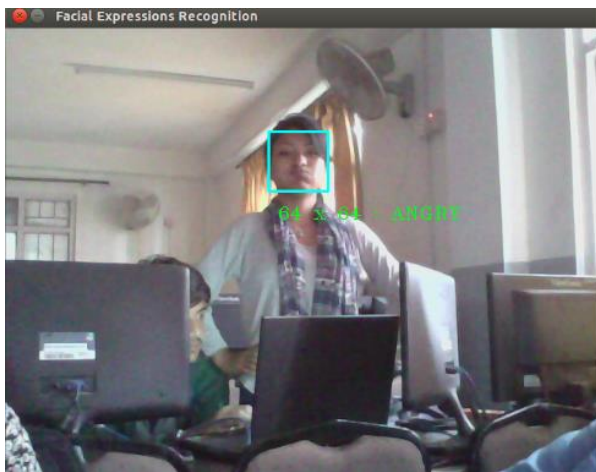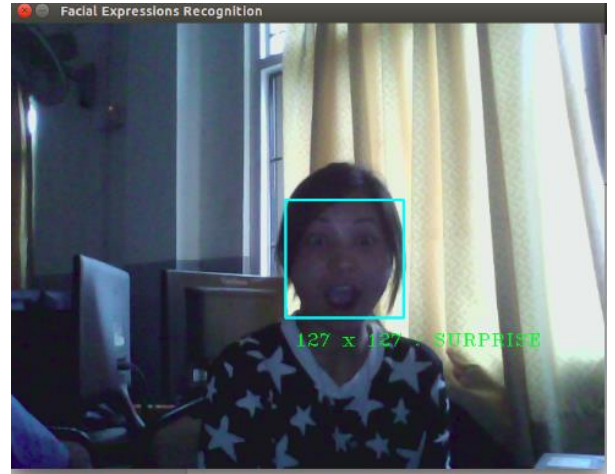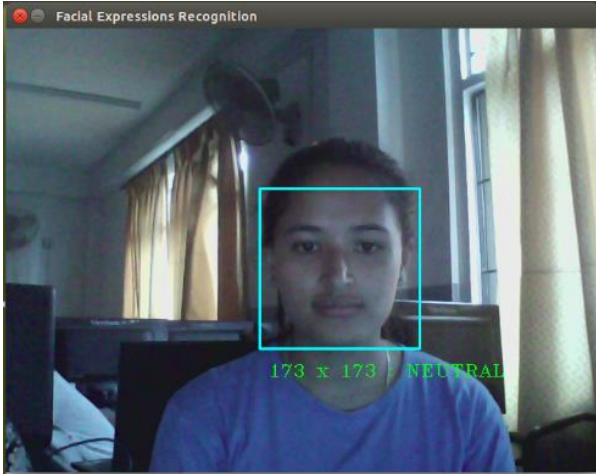## Experimental Demonstration from Image File

Figure 12: Experimental Demonstration from Image File

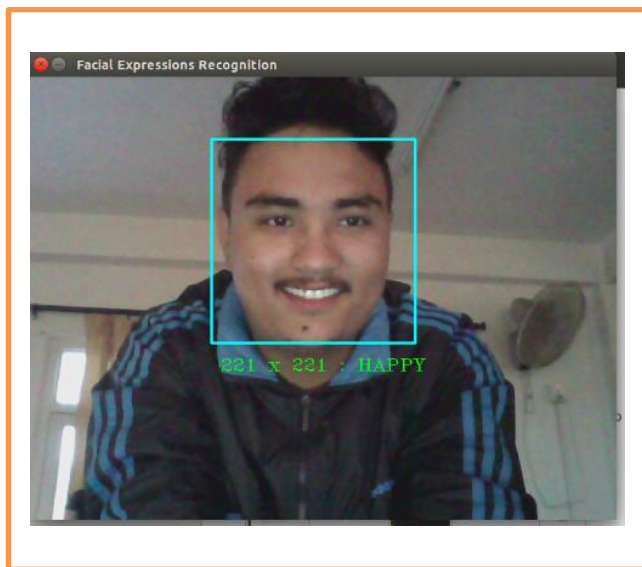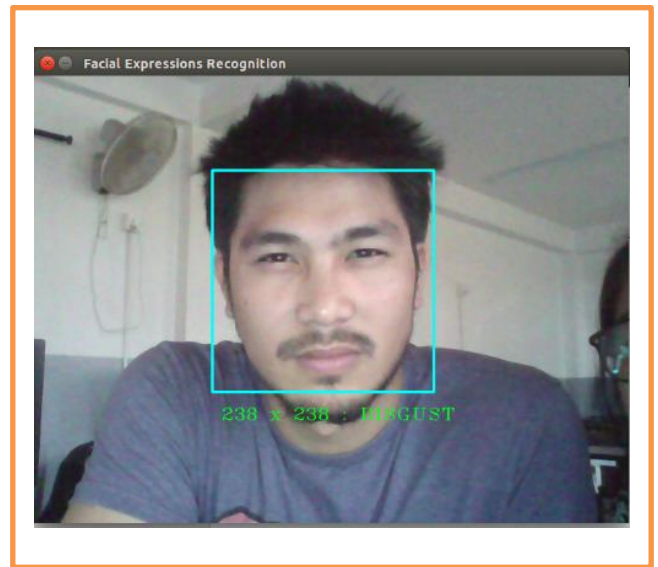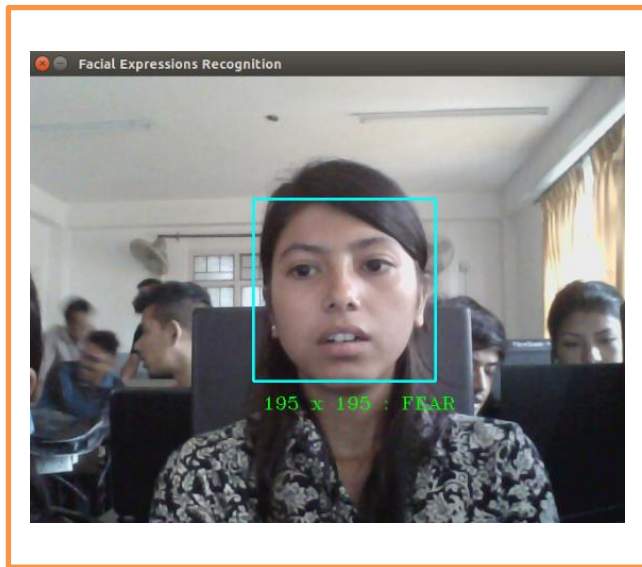## Experimental Demonstration from Camera

Figure 13: Experimental Demonstration from Camera