**Report on**

**Summer Internship Project-1**

“A Facial Emotion Recognition System Using Machine Learning”

In partial fulfillment of requirements for the degree of

**Bachelor of Technology (B. Tech.)**

in

**Computer Science and Engineering**

**Submitted by**

Ms. Ritika Rathi

200381

*Department of Computer Science and Engineering*

**SCHOOL OF ENGINEERING AND TECHNOLOGY**

**Mody University of Science and Technology Lakshmangarh, Distt. Sikar-332311**

December, 2022

**Declaration**

I hereby declare that the Report on Summer Internship Project-1 entitled " A Facial Emotion Recognition System using Machine Learning" is an authentic record of my own work as requirements of Industrial Training during the period of June 2022 to July 2022 for the award of degree of B.Tech. (Computer Science and Engineering), School of Engineering and Technology, Mody University of Technology and Science, Lakshmangarh, Sikar.

**Date**: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Ritika Rathi

(200381)

**Acknowledgement**

As it is rightly said that success taste better when it comes after immense hard work, dedication and failures to done a task.

I would also like to express my deepest appreciation to Mody University of Science and Technology for their constant support, guidance and encouragement.

Apart from the efforts put in by myself, I have a great pleasure to have the opportunity to extend my heartfelt gratitude to everyone who helped me throughout the course of this project.

**Table of Contents**

**Chapter 1 6**

1 Introduction 6

*1.1 History of Machine Learning 7*

*1.2 Working of Machine Learning 7*

*1.3 Need for Machine Learning 8*

*1.4 Types of Machine Learning*

**Chapter 2**

2 Tools and Technology Used 9

*2.1 Tools* 9

*2.2 Data Collection* 9

*2.3 Dataset Preparation* 11

**Chapter 3 13**

3 Literature Survey 13

*3.1 Literature Survey* 13

*3.2 Future Scope* 14

**Chapter 4 16**

4 Technology Implemented

*4.1 C++*  16

*4.2 IDE CLion For C++* 17

*4.3 Open CV Framework* 17

**Chapter 5 18**

5 Diagram and Working

*5.1 System design and diagram* 18

*5.2 Phase in facial expression recognition* 23

**Chapter 6 27**

6 Experimentation and Results 27

**Chapter 7 34**

7 Conclusion and Recommendation 34

**References 35**

**Outcome of Course 37**

**Chapter 1**

**1 Introduction**

Human emotion detection is one of the most advanced feature evolved in recent times. It will enhance the security of the devices as along with face lock we can add the second layer of security as emotion detection. This will be useful to verify the person standing in front of camera is not just 2-dimensional representation.

Another important use of emotion detection is for business promotions. In any business customer responses and feedback matters the most. If we can design an artificial intelligent system through which we can capture the user image or video of the customer so we can take the decision whether they liked or disliked the product or offer.

As we know that now-a-days security is priority and it also plays a significant role in identifying a person. We have many ways through which we can identify a person like voice recognition, finger print matching, passwords, retina detection etc. Identifying the intension of a person can also be one reason to reduce threats. This type of technology is mostly useful in vulnerable places like airports, concerts, railway stations, parliament house, major public gatherings where the breaches in recent years have increased.

There are many human emotions but if we focus on some major emotions so they can be classified as fear, disgust, anger, happiness, surprise, sad, contempt and neutral. These emotions are very subtle and facial muscles help in detecting these emotions. The detection of these emotions is difficult and very challenging as even a small difference in facial muscle contortions can result in change of emotion. Also, the expressions of different or the same people may differ for the same emotion, as emotion are hugely context and situation dependent. In this we can focus on more prominent areas of face for detecting a particular emotion like around the mouth and eyes, but how to extract and categorize these emotions is still important. So, for this we will be using neural networks and machine learning for these tasks and machine learning algorithms have proven to show good results in pattern recognition and classification. We can generalize detection of these emotions as following steps:

1. Dataset pre-processing
2. Face detection
3. Feature extraction
4. Classification based on the features
   1. **History of Machine Learning**

According to some, machine learning is a branch of artificial intelligence that focuses primarily on creating algorithms that enable a computer to independently learn from data and previous experiences. Arthur Samuel coined the phrase "machine learning" in 1959.Machine learning algorithms create a mathematical model with the aid of historical sample data, or "training data," that aids in making predictions or judgments without being explicitly programmed. Computer science and statistics are used with machine learning to create prediction models.

**1.2 Working of Machine Learning**

When a machine learning system learns from previous data, it creates prediction models and predicts the outcome for fresh data whenever it is received. The quantity of data used to create the model affects how precisely the outcome is predicted since a larger data set allows for the development of a better model.

Imagine that we have a difficult problem that requires certain predictions. Instead of creating code for it, we can just input the data to generic algorithms, and the machine would develop the logic according to the data and forecast the result. Our perspective on the issue has altered as a result of machine learning. The machine learning algorithm's operation is explained in the block diagram below:



**1.3 Need for Machine Learning**

Increased data production quickly solving difficult-for-a-human to solve complicated challenges decision making in different sectors, including finance discovering hidden patterns and removing information that is helpful from data.

**1.4 Types of Machine Learning**

Different machine learning algorithms have different approaches, input and output data types, and tasks or problems that they are meant to address. Mainly it can be classified as:

1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning
4. Semi-supervised Learning

**Chapter 2**

**2.1 Tools and Technology Used**

* Hardware Platform(Recommended)

1. Fluently working Laptops
2. RAM minimum 4Gb
3. Web Camera

* Software Platform(Recommended)

1. C++ programming language
2. CLion IDE (selective)
3. OpenCV framework
4. Linux platform Ubuntu OS

**2.2 Data Collection**

**2.2.1 Cohn-Kanade AU Coded Facial Expression Database**

100 college students make up the subjects in the COHN-KANADE AU-Coded Facial Expression Database that has been made public. They were between the ages of 18 and 30. 65 percent of the population was female, 15 percent African-American, and 3 percent Asian or Latino. An investigator gave the subjects instructions on how to make a sequence of 23 facial displays, which included both single action units and combinations of action units. Sequences were digitalized into 640 by 480 or 490 pixel arrays with 8-bit precision for grayscale values from neutral to the desired display. "Sequence" files, which are brief text files that specify the order in which photos should be viewed, are included with the image files. The seven expressions are angry, surprise, contempt, fear, and disgust. Fig.1 shows the 8 expressions with each from a different subject.



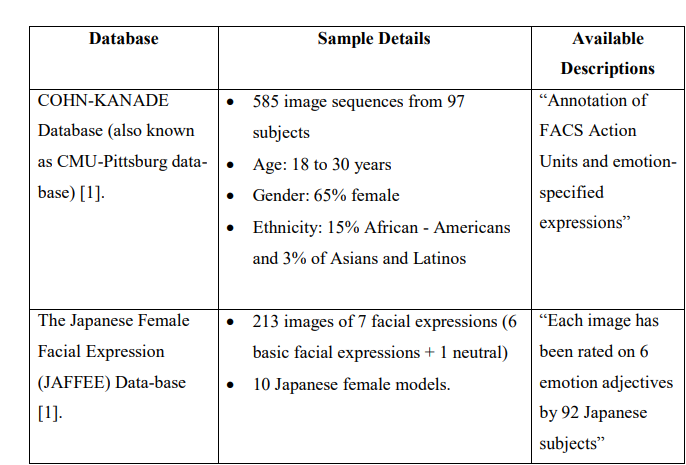
**Fig.1: The eight expression from one subject**

**2.2.2 Japanese Female Facial Expression(JAFFE) Database**

There are 213 total photos in this database. Each of the 10 individuals has seven different face expressions. About twenty photographs per topic and two to three images per expression are used in each subject. Anger, joy, disgust, sorrow, surprise, fear, and neutral are the seven expressions, in that order. Fig.2 shows the seven expressions from one subject.



**Fig. 2: The seven expressions from one subject**



**Table 1: Data Collections**

**2.3 Dataset Preparation**

Two datasets, COHN-KANADE and JAFFE, were used to train and test the suggested method. The JAFFE dataset only contains 213 photos, whereas the COHN-KANADE dataset contains 500 image sequences from 100 patients. We utilized 6481 photographs from the Cohn-Kanade dataset from various participants for training and 1619 images for testing in our experiment. Similar to this, 107 photos from the JAFFE dataset were utilized for training and 106 images for testing.

We reduced the faces' resolution to 72 pixels. The original photos of faces were cropped to 256\*256 pixels based on the anatomy of the face. The face detector of our own system, which is based on the Haar classifier, was used to automatically detect faces in order to identify facial images. Face location, face breadth, and face height were automatically constructed from the results of the face detection. After being cropped, photos were utilized for training and testing after being further cropped in line with the face detector's output.



**Fig. 3: Original Image**



**Fig. 4: Cropped Image**

**Chapter 3**

**3.1 Literature Survey**

There is a wealth of material on face tracking and detection that has been the subject of intense research. The lack of data on spontaneous expression is the main issue the researchers are dealing with. Unpredictable facial emotions are one of the hardest things to capture on film and in photos. There have been several attempts to identify facial expressions. For the purpose of recognizing facial expressions, Zhang et al. explored two categories of features: geometry-based features and Gabor wavelets-based features.

Face detection strategies include appearance-based approaches, feature-invariant approaches, knowledge-based approaches, and template-based approaches, whereas expression detection strategies in a related field include local binary pattern phase correlation, Haar classifier, AdaBoost, and Gabor Wavelet. Emotient, Affectiva, Karios, etc. are some of the APIs for expression recognition, while Face Reader is the industry leader for automatic analysis of facial expression detection. The encoding of facial features and the classifier problem are two essential components of automatic facial expression recognition.

In order to describe faces, a set of suitable characteristics must be extracted from the original face photos. The techniques utilized for representing face features include Histogram of Oriented Gradient (HOG), SIFT, Gabbor Fitters, and Local Binary Pattern (LBP) [3,4]. LBP is a straightforward yet very effective texture operator that identifies each pixel in an image by thresholding its immediate surroundings and treating the result as a binary integer. By thresholding the 3X3 neighborhood around each pixel with its center value and using the result as a binary integer, the operator labels the pixels in a picture. Dalal and Triggs initially suggested HOG in 2005. HOG assigns a numerical value to the occurrence of gradient orientation along a local picture path.

We employ methods like machine learning, neural networks, support vector machines, deep learning, and naive bayes for classifier problems. Support Vector Machine (SVM) will be used to create a histogram utilizing any representation of a face characteristic for the purpose of identifying facial expressions. In order to divide the high dimensional space, SVM constructs a hyperplane. When the distance between the hyper plane and the training data of any class is the greatest, an optimum separation is realized.

For greater identification accuracy, the block size for the LBP feature extraction is set. According to the test findings, facial expressions may be recognized with an accuracy of more than 97% when LBP characteristics are used. The local and global aspects of the face picture are extracted more accurately using the block LBP histogram features. LBP works with a variety of classifiers, filters, etc.

**3.2 Future Scope**

Over the past ten years, face expression recognition technologies have significantly advanced. Posted expression recognition is no longer the main emphasis; instead, spontaneous expression recognition has taken center stage. Under face registration mistakes, quick processing, high correct recognition rate (CRR), and considerable performance gains in our system, promising results may be obtained. The system is entirely autonomous and is equipped to handle an image stream. It can identify unprompted facial expressions. Digital cameras that employ our method can only take pictures of people who are grinning. In security systems that can recognize a person, he can present himself in any way. When someone enters a room in a house, the lighting and television may be adjusted to their preferences. The technique helps doctors comprehend a deaf patient's level of disease or agony. Our technology may be utilized in mini-marts and shopping centers to examine client feedback in order to improve the business, as well as to identify and track a user's mental condition.

**Chapter 4**

**Technology Implemented**

* 1. **C++**

An all-purpose programming language is C++. It provides features for imperative, object-oriented, and generic programming and offers tools for low-level memory management. Its design strengths were performance, efficiency, and flexibility of usage, with a slant toward system programming and embedded, resource-constrained, and massive systems. With its key strengths being software infrastructure and resource-constrained applications, such as desktop applications, servers (such as e-commerce, web search, or SQL servers), and performance-critical applications (such as telephone switches or space probes), C++ has also been found to be helpful in many other contexts. The Free Software Foundation's GCC, LLVM, Microsoft, Intel, and IBM all offer C++ implementations that may be used on a variety of platforms. C++ is a compiled language.

**Features**

* Object oriented programming language

Object-oriented programming is the primary improvement between C and C++. It adheres to the basic concepts of polymorphism, inheritance, encapsulation, and abstraction. This facilitates development and upkeep.

* Memory Management

Dynamic memory allocation is supported by C++. The RAM that has been allocated can always be freed. Additionally, C++ has dynamic memory management strategies.

* Platform Dependent

Platform dependent languages are those that can only be used on the operating system on which they were written and compile. On any other operating system, it cannot be run or executed.

* 1. **IDE CLion for C++**

The cross-platform C/C++ IDE CLion has clever Cmake support, making it more than just an editor. With intelligent and pertinent code completion, fast navigation, and trustworthy refactoring, CLion assists in thoroughly understanding programmes and can increase productivity.

* 1. **Open CV Framework**

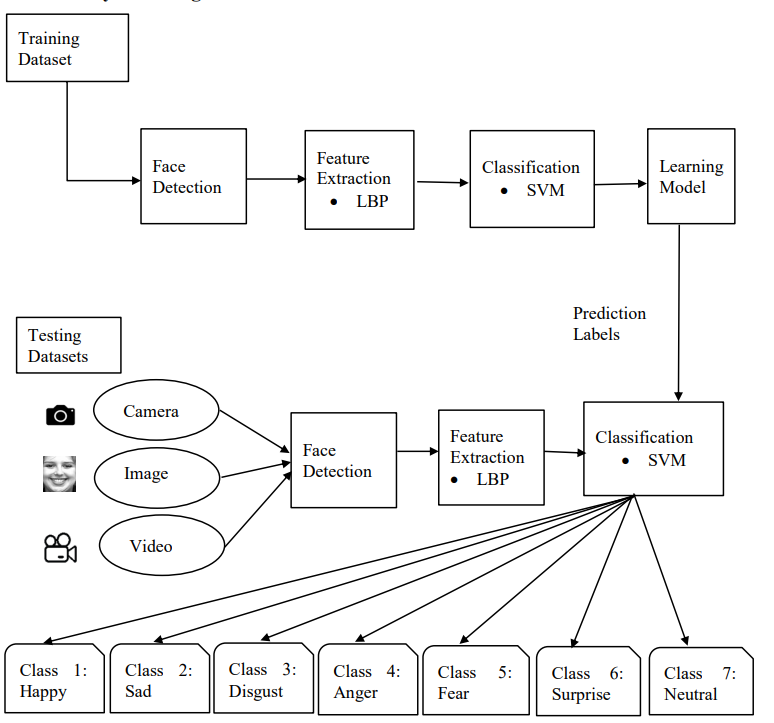
A computer vision and machine learning software library called OpenCV is available for free use. A standard infrastructure for computer vision applications was created with OpenCV in order to speed up the incorporation of artificial intelligence into products. OpenCV makes it simple for companies to consume and alter the code because it is a BSD-licensed product.

More than 2500 optimized algorithms are available in the collection, including a wide range of both traditional and cutting-edge computer vision and machine learning techniques.

These algorithms can be used to find similar images from an image database, remove red eyes from flash-taken photos, follow eye movements, 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, and establish markers to over time. It has C++, C, Python, Java and MATLAB interfaces and supports Windows, Linux, Android and Mac OS.

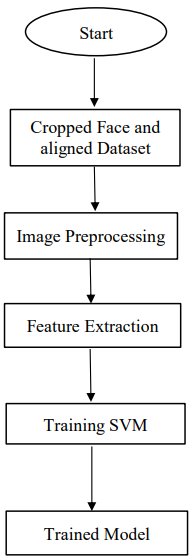
**Chapter 5**

5.1 System Design and Diagram

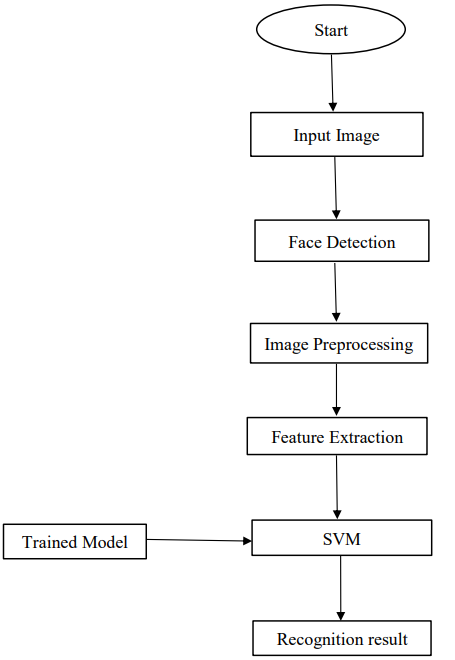


**Fig. 5: System Diagram**

5.1.1 System Flowchart

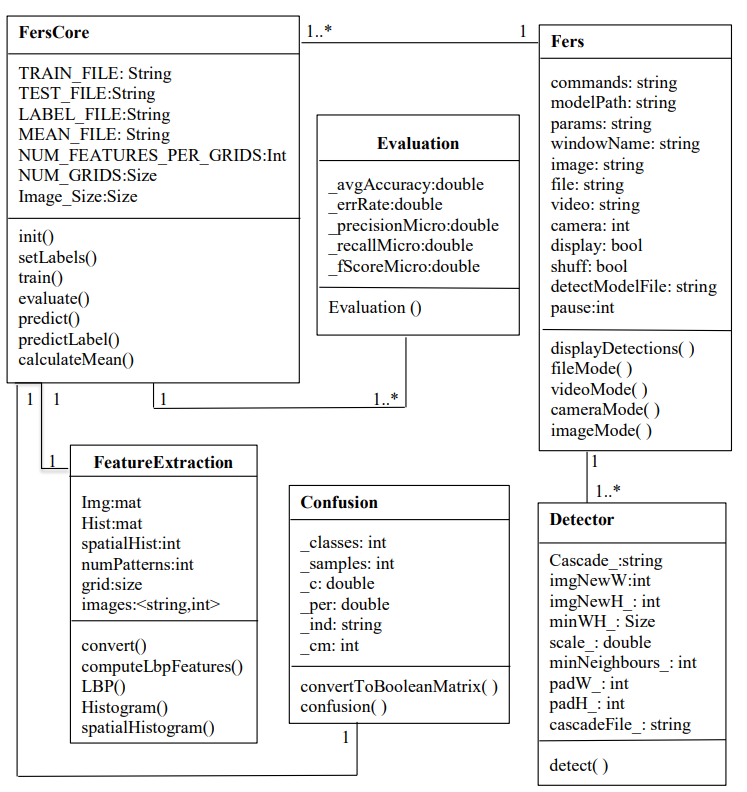


**Fig.6: Flowchart of training**



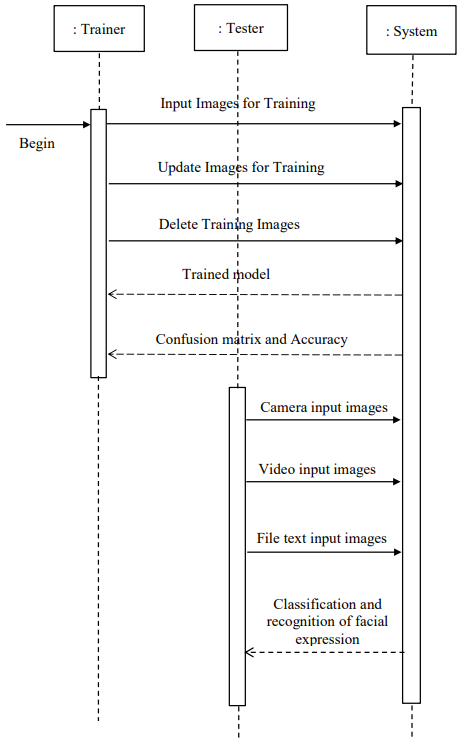
**Fig.7: Flowchart of testing/prediction**

5.1.2 Class Diagram



**Fig.8:Class Diagram**

5.1.3 Sequence Diagram



**Fig.9: Sequence Diagram**

5.2 Phase in Facial Expression Recognition

The supervised learning method is used to train the facial expression recognition system, which collects photos of various face emotions. The system starts with a training and testing phase before acquiring images, detecting faces, analyzing those images, extracting features, and classifying them. Face recognition and feature extraction are performed on face photos, followed by classification into six groups that correspond to the six primary expressions listed below:

5.2.1 Image Acquisition

Static photos or image sequences are utilized to detect face expressions. A camera can record pictures of faces.

5.2.2 Face Detection

In order to identify facial images, face detection is helpful. Face detection is done in the training dataset using the OpenCV implementation of the Haar classifier Voila-Jones face detector. Haar-like features, which are composed of linked rectangles of black and white and whose value is the difference in the sum of the pixel values in the black and white areas, encode the difference in average intensity in various portions of the picture.

5.2.3 Image Pre-Processing

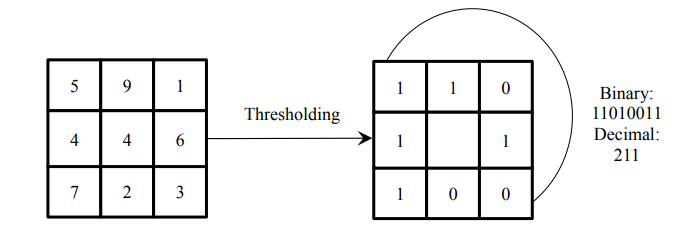
Noise is eliminated during image pre-processing, and brightness or pixel location variations are normalized. Normalization of the histogram and color.

5.2.4 Feature Extension

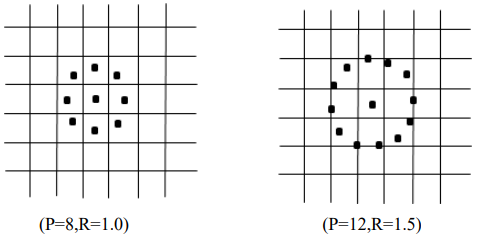
The choice of the feature vector in a pattern classification issue is crucial. After pre-processing, the facial picture is utilized to extract the key characteristics. Scale, attitude, translation, and changes in light level are some of the fundamental issues with picture categorization. The LBP algorithm is used to extract the significant characteristics and is outlined below:

5.2.4.1 Local Binary Pattern

The feature extraction method is called LBP. The original LBP operator encodes the local structure around each pixel by pointing the pixels in an image with decimal numbers, often known as LBPs or LBP codes. By removing the value of the center pixel, each pixel is compared to its eight neighbors in a 3 X 3 neighborhood. In the final code, negative values are represented by 0 and all other values by 1. A binary number is created for each given pixel by adding all of these binary values clockwise, beginning with the top-left neighbor. The specified pixel is then given a label based on the produced binary number's equivalent decimal value. The LBPs or LBP codes are the terms used to describe the resulting binary integers.



**Fig. 10: The Basic LBP Operator**

****

**Fig. 11: Two examples of extended LBP**

The basic LBP operator's drawback is that big scale structures cannot be used to capture the dominating characteristics in its tiny 33 neighborhood. As a result, the operator was later expanded to employ neighborhoods of various sizes to deal with the texture at various scales. Any radius and number of pixels in the neighborhood are possible when using circular neighborhoods with bilinear interpolation of the pixel values. In examples of the extended LBP are displayed, where (P, R) designates sampling locations on a circle with a radius of R.

LBP is further expanded to include consistent patterns. When the binary string is thought of as circular, an LBP is referred to as uniform if it has no more than two bitwise transitions from 0 to 1 or vice versa. Such patterns include 00000000, 001110000, and 11100001, for instance. A labelled image's histogram, f1(x, y), may be described as

Hi = ∑ I (fl x,y (x, y) = i), i = 0, … , n − 1 ……(1)

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

I(A) = { 1 A is true

0 A is false ….(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

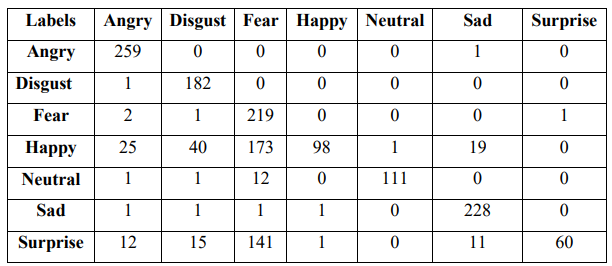
Hi = ∑ I (fl x,y (x, y) = i)I((x, y)ϵRj …..(3)

**Chapter 6**

6.1 Experimentation and Results

The development of a comprehensive face expression recognition system is the goal of this research endeavor. For the experiments, COHN KANADE and JAFFE datasets were employed. First, supervised learning was used to train the system using various random samples from each dataset. Every dataset has its data split into two halves for training and testing. Each dataset contains entirely unique samples that are uniformly drawn at random from the pool of the specified dataset. The COHN KANADE datasets had 585 folders of both subject and session, with 97 subject directories and 8795 picture files overall. The datasets were divided in an 8:2 ratio, with 6481 (80%) of the train dataset being used for testing and 1619 (20%) of the test. Similar to this, the JAFFE dataset has 213 pictures that were split into 160 (75%) train images and 53 (25%) test images in a ratio of 7.5:2.5.

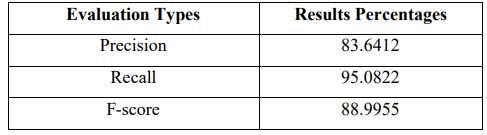
The confusion and accuracy evaluation results of COHN-KANADE and JAFFE datasets are as below:



**Table 2: Confusion matrix of COHN-KANADE**

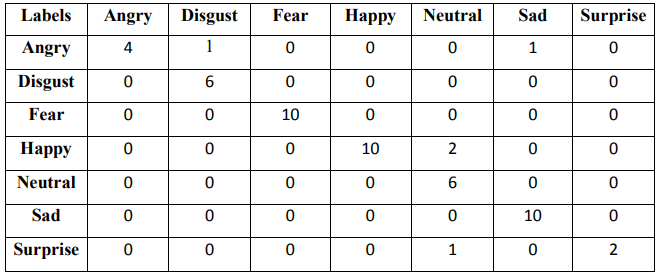
Row in the table above displays actual classes, whereas column displays anticipated classes.

The classifier made a total of 1619 predictions, out of which angry was predicted 25,300 times, disgust was predicted 239 times, fear was predicted 545 times, joyful was predicted 99 times, neutral was predicted 112 times, sad was projected 259 times, and surprise was predicted 61 times. While in fact, 260 incidents involved anger, 183 disgust, 223 fear, 356 happiness, 125 neutrality, 228 sadness, and 240 surprise.



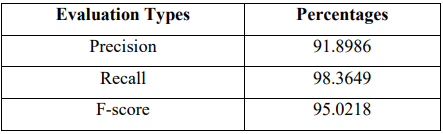
**Table 3: Accuracy of COHN-KANADE**

According to the table above, 95.0822% of the expressions were accurately allocated, while 83.6412% of the expressions were anticipated. Precision and recall had a harmonic mean of 88.9955%.



**Table 4: Confusion matrix of JAFFE**

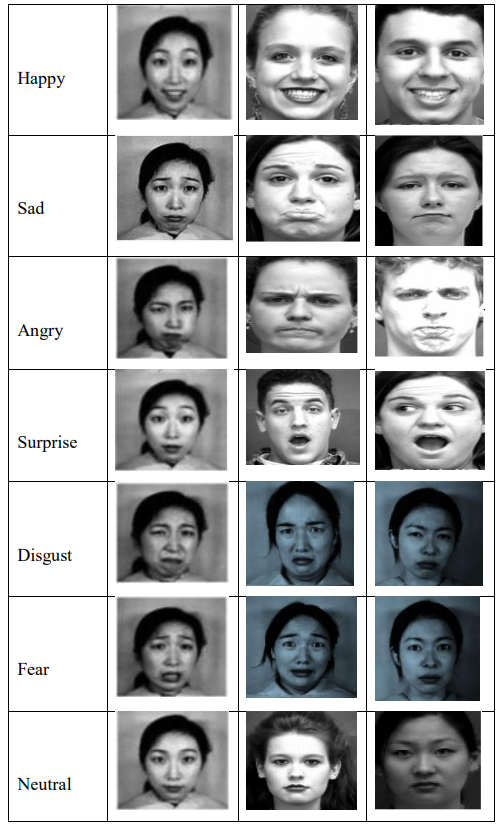
Row in the table above displays actual classes, whereas column displays anticipated classes. A total of 53 predictions were produced by the classifier, of which 4 were for anger, 7 were for disgust, 10 were for fear, 10 were for happiness, 9 were for neutral, 11 were for sadness, and 2 were for surprise. While in actuality there were six incidents of anger, six of disgust, ten of fear, twelve of happiness, six of neutrality, ten of sadness, and three of astonishment.



**Table 5: Accuracy of JAFFE**

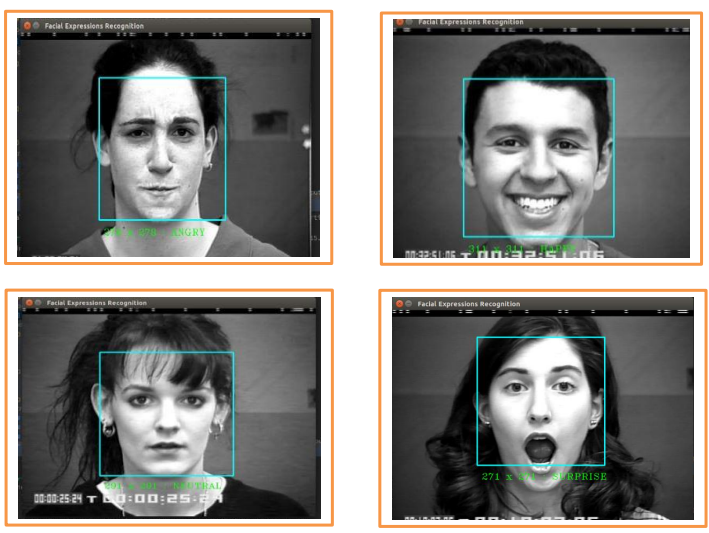
According to the aforementioned table, 91.8986% of the expressions were accurately predicted, and 98.3649% of the expressions were allocated appropriately. Precision and recall have a harmonic mean of 95.0218%.

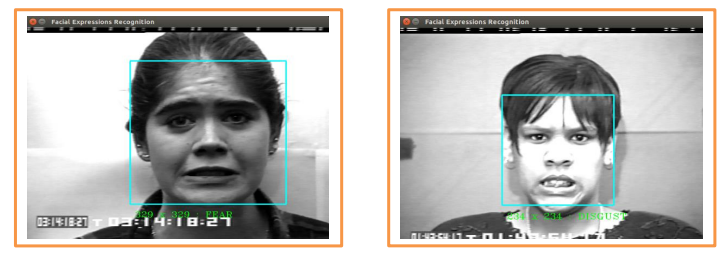
**Datasets Collection**

****

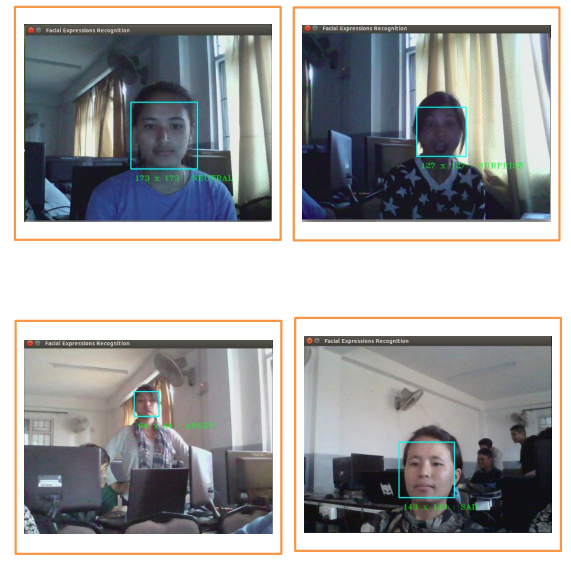
**Table 6: Dataset images of facial recognition**

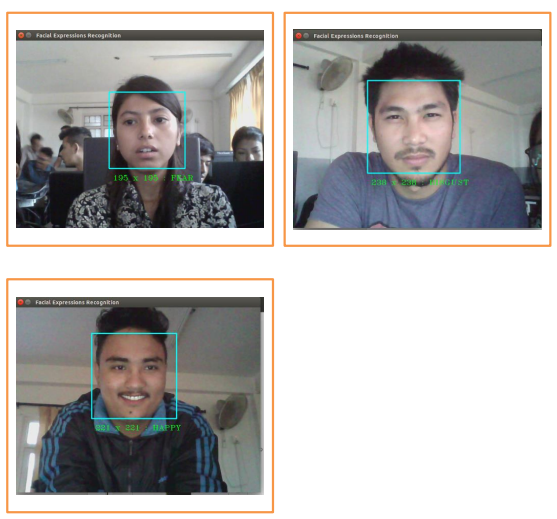
**Experimental Demonstration**





**Figure 12: Experimental Demonstration from Image File**





**Figure 13: Experimental Demonstration**

**Chapter 7**

7.1 Conclusion and Recommendation

In this study, a method for classifying face expressions is proposed. Many applications, including robotic vision, video surveillance, digital cameras, security, and human-computer interaction, benefit from face detection and extraction of facial emotions. The goal of this research was to create a facial expression recognition system that utilized computer visions and improved face expression recognition's advanced feature extraction and classification.

Seven different facial expressions of photos of various people from various datasets have been evaluated for this research. In this research, facial expressions are preprocessed from collected face photos before features are extracted using Local Binary Patterns and facial expressions are classified using training datasets for Support Vector Machines on facial image datasets. This research uses the JAFFE, COHN-KANADE face database to detect additional facial emotions. The system has been assessed using Precision, Recall, and Fscore in order to gauge the effectiveness of the suggested methodology and approaches and verify the correctness of the findings. By splitting the datasets into training samples and testing samples in the ratios of 8:2 for COHN-KANADE and 7.5:2.5 for JAFFE, the identical datasets were utilized for both training and testing. The COHN-KANADE dataset's Precision, Recall, and Fscore were 83.6142%, 95.0822%, and 88.9955%, respectively, whereas the JAFFE dataset's values were 91.8986%, 98.3649%, and 95.0218%.

Results of experiments conducted on the COHN-KANADE dataset and the JAFFE database demonstrate the potential of the suggested strategy. Recognition of facial expressions is an extremely difficult topic. More work has to be put into enhancing categorization performance for critical applications. Our future work will concentrate on enhancing the system's functionality and generating more accurate classifications that might be helpful in a variety of real-world applications.

**References**

* <https://www.javatpoint.com/>
* <https://www.geeksforgeeks.com/>
* Bettadapura, V. (2012). Face expression recognition and analysis: the state of the art. arXiv preprint arXiv:1203.6722.
* 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.
* 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.
* 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).
* 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.
* 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.
* 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.
* Chibelushi, C. C., & Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. CVonline: On-Line Compendium of Computer Vision, 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.
* Michel, P., & El Kaliouby, R. (2005). Facial expression recognition using support vector machines. In The 10th International Conference on Human-Computer Interaction, Crete, Greece.
* 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

**Course Outcome**

1. The basics of machine learning system.
2. How to use machine learning to make decision and predictions in a system.
3. Understand what is learning and why it is essential to the design of intelligent machines.
4. Know how to fit models to data.
5. Understand numerical computation, statistics and optimization in the context of learning.
6. Have a good understanding of the problems that arise when dealing with very small and very big data sets, and how to solve them.
7. Understand the basic mathematics necessary for constructing novel machine learning solutions.
8. Learnt the different frameworks and concepts of data visualization.
9. Analyzing data using different functions.