FACIAL EMOTION RECOGNITION BASED ON BIORTHOGONAL WAVELET ENTROPY AND SUPPORT VECTOR MACHINE

V SEMESTER MINI PROJECT REPORT

FOR THE DEGREE OF

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

BY

PRABHAT SINGH (IIT2015057)
YASH GUPTA (IIT2015072)
AKASH KUMAR (IIT2015084)
SAURABH SINGH (LIT2015013)
RAHUL KUMAR (LIT2015023)

UNDER THE SUPERVISION OF

Dr. ANUPAM AGRAWAL
Professor
IIIT-ALLAHABAD

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, ALLAHABAD

November 2017

CANDIDATES' DECLARATION

This is to certify that this project report entitled **"Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy and Support Vector Machine"** which is submitted by our group in partial fulfillment of the requirement for the evaluation of Mini Project for 5th semester B.Tech. in Information Technology (July 2017 - Nov 2017), is an authentic record of our original work carried out under the guidance of **Prof. Anupam Agrawal** and due acknowledgements have been made in the text of the report to all other materials used. This report work was done in full compliance with the requirements and constraints of the prescribed curriculum.

Date:

Prabhat Singh (IIT2015057)

Yash Gupta (IIT2015072)

Akash Kumar (IIT2015084)

Saurabh Singh (LIT2015013)

Rahul Kumar (LIT2015023)

CERTIFICATE FROM SUPERVISOR

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date:

Place : Allahabad

Declaration Period : July 2017 – Nov 2017

Prof. Anupam Agrawal
Project Supervisor

Abstract

Human emotions are not only important in relations but also play an important role in the way which we interact with the computer. Many algorithms have been developed for recognizing human emotions but they have not been up to the mark in terms of accuracy.

In this project we aim to develop a Facial Emotion Recognition model based on Biorthogonal Wavelet Entropy as a Feature Extractor and Support Vector Machines as a classifier, and test it on various datasets including benchmark datasets and a dataset created on our own.

We trained and tested our software on JAFFE dataset and accuracy was found to better than earlier emotion recognition techniques.

Contents

1.Introduction

Human emotions are complicated, and very difficult to envision. However, psychological studies have proven that emotions play a critical role in logical and intellectual behavior. Emotions are a complex mental state that drives a person to react positively or negatively to a stimulus. Emotions help in different types of communications, like, verbal or non-verbal, peer to peer or community. But, emotions are very much short-lived as compared to other effects, namely behavior or mood. Nevertheless, emotions have been considered as signs of intelligence. By adding meaning to syntax, emotions make communication interactive and understandable. Even though it is very difficult for a machine to understand the human emotional state. However, since the onset of artificial intelligence there have been differences between the researchers defining this term. This is the reason that till now the machines are not as smart as human beings in terms of synergy.

Emotional recognition is the process of identifying human emotions, largely facial expressions. Both of these things are done by the person himself but computational methods have also been developed. Recently, Human-Robot interaction have come into the limelight. Socializing robots with humans, comprehending facial expressions and visual hints of a human have been in great demand. It permits the robot to understand the manifestations of humans, which in turn can increase its effectiveness in fulfilling various tasks. This acts as a measurement system for behavioral science. Socially intelligent software tools can be achieved. Emotions can be reflected through voice, gestures, facial expression etc. A person's physical condition may also change, as he expresses different feelings. There are six major emotions i.e. happiness, sadness, fear, hatred, anger and wonder. Research in this area is still evolving, because so far in the developed systems, the feelings are not detected since factors and thus we have less accuracy.

One possible way to understand this complicated behavior is to understand human emotions. Human emotions can be reflected by many patterns such as facial expressions, gestures, muscular movements, speech tone, dermal activity etc. In spite of these important settings, psychological indicators were given more attention to the identification of human emotions among stimulating computing researchers.

1.1 Motivation

Due to technological advancement in the last decade due to the development of more powerful computers, artificial intelligence has reached new heights. Artificial Intelligence Research is particularly specialized with the highly fundamental goal to maximize the potential of success by taking action after understanding the environment. Beyond today's applications, artificial intelligence is in the heart of many new technologies that will shape our future. It has ignited intelligent agents to work more efficiently.

The inspiration for this project is that although many methods exist for solving the facial emotion recognition problem, the performances of these methods have never been satisfying, besides their method can’t handle noise. Thus we propose a new method with much higher accuracy.

Psychological studies have shown that emotions play an important role in rational and intelligent behavior. To make intelligent agents, we need to first understand human emotions, as well as we need to identify the object. For these reasons, we got motivation to create an algorithm that helps the machine identify and classify human emotions.

1.2 Problem definition

This project aims to come up with a solution to the “Facial Emotion Recognition” problem. It uses Biorthogonal Wavelet Entropy for feature extraction. This method is composed of the advantages of both Biorthogonal Wavelet Transform and Shannon Entropy. Moreover, this project uses a compelling variant of Support Vector Machine (SVM) for classification of the extracted features. Finally, Stratified Cross Validation was engaged as a strict validation model for the results

1.3 Objective

Our objective is to make a system that will use computer vision techniques to automatically detect and analyze the emotions from the digital images. The emotion will be recognized based on Biorthogonal Wavelet Entropy as a feature extractor. Once we have the confusion matrix of our dataset, we plan to make decision by comparing and analyzing them. Fig 1.1 shows the objective of our project.

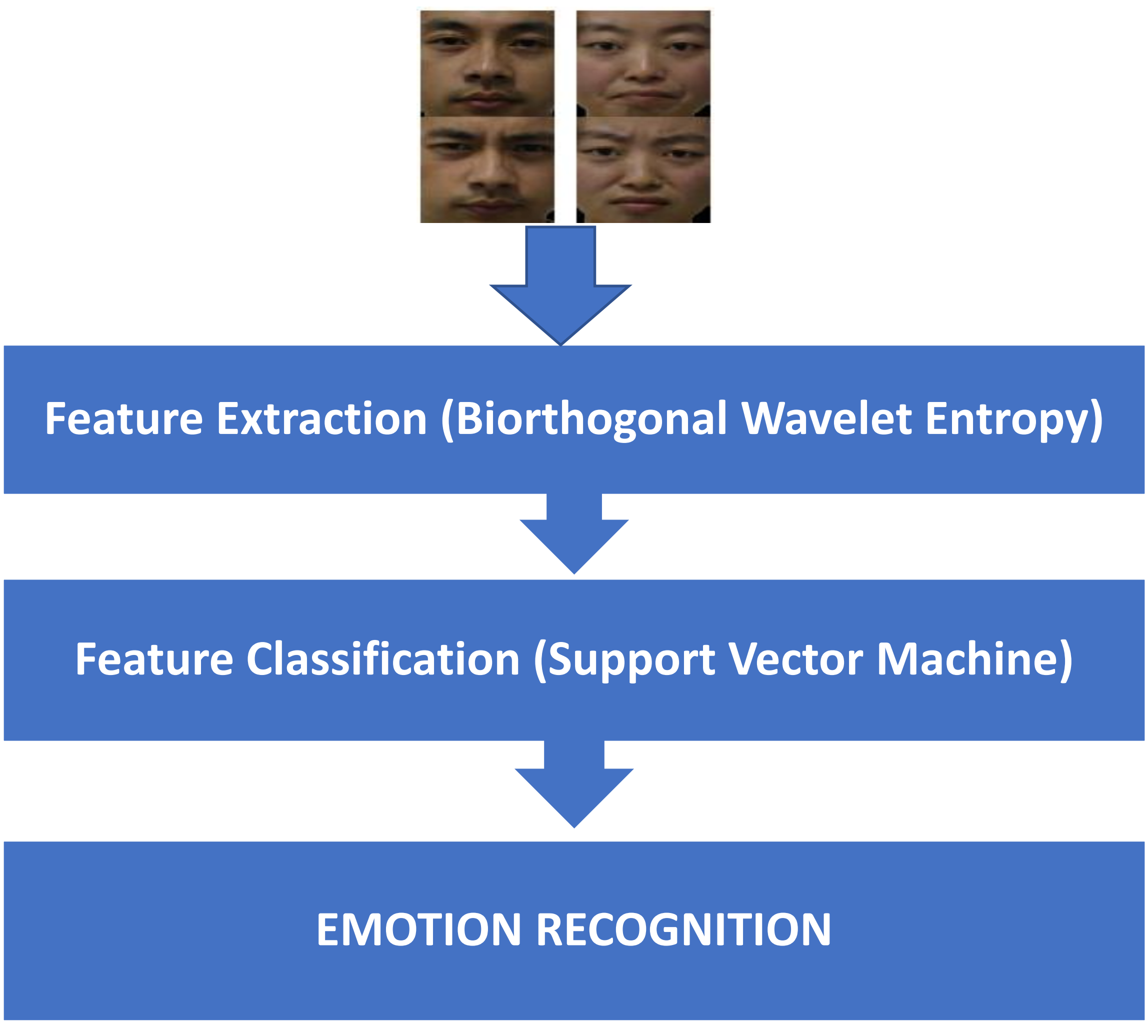

Figure 1.1: Objectives

1.3.1 JAFFE Dataset

The JAFFE dataset [13] used for training process in the mini project is a collection of 213 images consisting of 7 different emotions namely anger, disgust, fear, happy, sad , neutral and surprise.

1.3.2 Feature Extraction

The next task is to extract features from the data of selected images.

1.3.3 Feature classification

After feature extraction the next task would be to classify it. Support Vector Machines(SVM's) has been chosen as a classifier, which gives optimal classification results.

1.3.4 Recognition of emotion

Based on the classification results the emotions are classified into their respective class.

2. LITERATURE SURVEY

Sr. No.	Title of Paper	Author Name	Journal/ Conference Name (Year)	Publisher	Main Technique	Data Set	Merits	Demerits	Future Scope
1	Face Recognition and Facial Expression Identification using PCA[1]	Suknaya Sagarika Meher and Pallavi Maben	International Advance Computing Conference (2014)	IEEE	PCA	ATT, CSU and MPI	Quick Speed and Simple Calculation	The trade-offs between the number of ngen faces necessary for unambiguous classification	This work can be extended to surveillance application.

2	Automatic Face Expression Recognition System[2]	Jiquean Li And M. Oussalah	IEEE 9th International Conference (2010)	IEEE	1.Negative Matric Factorization(NMF) 2.Principle component analysis(PCA) 3.K Nearest Neighbor (KNN)	1.Indian Face Database 2.TFEID Taiwanese dataset	Higher recognition rate over 75% average , NMF better performs reduction of data redundancy .	NMF takes more time in comparison to PCA	Minimum distance qualifier algorithm can be used to select out high intensity facial expression images which can increase accuracy
3	Emotion Recognition Using Electroencephalography (EEG) Images[3]	Priyanka A. Abhang and Bharti Gawali	International Journal of Advanced Research in Computer Science (2014)	IJARCS	Electroencephalography (EEG)	Radboud faces Database	Distinguishing between positive and negative emotion.	Classify Expression as happy and sad only.	Can be used to find emotion of audio less video.

4.	Multimodal Emotional Recognition using Deep Learning Architectures [4]	Hiranmayi Rangnathan, Shayok Chakrabotry, Sethuraman Panchanathan	IEEE Winter Conference on Applications of Computer Vision (2016)	IEEE	1.Gaussian RBM 2. Bernoulli-Bernoulli RBMs for training of deep layers	emoF-BVP dataset which consist of facial and vocal expressions, body gestures and physiological features.	Used primary expressions of emotion of the lowest intensity from the lowest intensity from emoFBVP, mind Reading.	In real time scenario , data from one or more modalities may be absent.	Future prospect is to design real time emotion recognition system that gives perfect results even if one or more parameters are absent.
5.	A Real Time Facial Expression Classification System Using Local Binary Patterns[5]	S L Happy, Anjith George, and Aurobinda Routray	4th International Conference on Intelligent Human Computer Interaction, Kharagpur ,India (2012)	IEEE	Haar classifier , PCA , multi local binary pattern features	Average confusion matrix	Recognition accuracy is more than 97%.	To classify frontal image only (rotation of face degrades the performance of system)	It help to the patient when they are in movement.

6.	Emotion Detection Algorithm[6]	Namrata Mahajan and Harshad Mahajan	International Journal of Electrical and Electronics Research (2014)	IJEER	Face Detection + Lip Detection + Skin Segmentation	Roipoly function to create template (template data base generator)	Simple and fast method as compared to EEG or by Speech	Face Image consist of Single Image (more than one people can't use simultaneously)	The fall detection device can ported as hand hold portable device with WIFi facility.
7.	Multimodal emotion recognition based on peak frame selection from video[7]	Zahid Akhtar And Cigdem Erdem	Signal, Image and Video Processing (Springer) (2015)	Springer	1.Maximum dissimilarity-based method (MAX DIST) 2. Clustering-based method (DEND-CLUSTER) 3. Emotion intensity-based method (EIFS) 4. Support Vector Machine(SVM)	1.eNTERFACE And 2.BAUM-1a	Achieved subject independent audio visual emotion recognition at rate of 78.26% on eNTERFACE database.		The investigation of the effects of lip motion due to speech on the facial expression recognition performance is another direction for future research

8.	Facial Expression Recognition with fusion features extracted from salient facial areas[8]	Yanpeng Liu, Yibin Li, Xin Ma and Rui Song	Sensors (MDPI) (2017)	MDPI	Gamma Correction, LBP, PCA And HOG	CK+ database and JAFFE database.	The gamma features correction method is firstly applied on the LBP features and this significantly improves the recognition result in our algorithm frameworks	1)if the landmarks are not accurate our recognition result will be influenced 2) if there is not enough image data, our gamma correction can not improve the recognition a lot	1)this technology can help doctors to monitor patients, which will save hospitals much time and money. 2) facial expression technology can be applied in a car to identify whether the driver has fatigue, and this can save many lives.
9.	Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector	YU-DONG ZHANG, ZHANG-JING YANG, HUI-MIN LU, XING-XING ZHOU,	IEEE Access (Volume : 4)	IEEE	biorthogonal wavelet entropy, support vector machine, fuzzy logic.	Self Data-Set	method achieved an overall accuracy of 96.77±0.10 %		Recommender system , cloud computing , big data And Sparse representati

	Machine, and Stratified Cross Validation[9]	PREETHA PHILLIPS, QING-MING LIU, AND SHUI-HUA WANG							on is another potential research direction

Table 2.1: Literature Survey

3.Theory

3.1 Wavelets:

Wavelet is a small portion of a wave, which is continuously decaying. It can be defined over a finite interval and having an average value of zero.

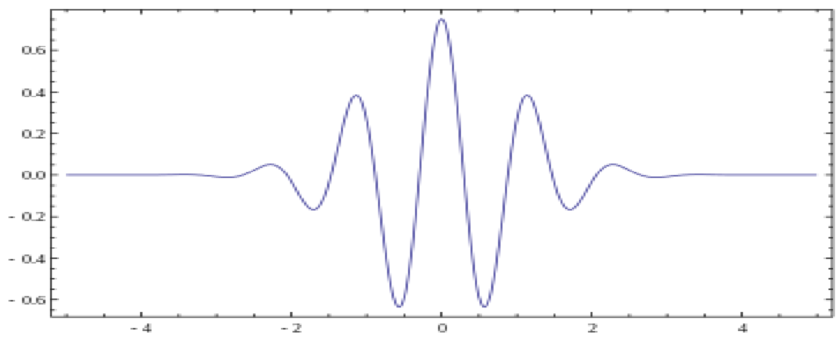

Fig 3.1: An example of Wavelet [10]

3.2 Wavelet transform:

Wavelet transform decomposes signal into a set of basic functions (wavelets). Wavelets are obtained from a prototype wavelet ψ(t).

3.3 Why wavelet transform:

Fourier transform is a powerful tool for data analysis because it does not represent abrupt changes efficiently which is required in the image analysis, Therefore to analyze abrupt changes we need a function (wavelet) which are well localized in time and space. So, we use wavelet transform.

3.4 Biorthogonal Wavelet Transform:

Biorthogonal Wavelet Transform is generalized from Orthogonal Wavelet Transform by using an extra set of low pass and high pass filters for decomposition and reconstruction. Biorthogonal Wavelets allows more freedom than orthogonal and help in creating symmetric wavelet functions.

3.5 Support Vector Machines (SVMs) :

Support Vector Machines are machine learning algorithms which take a set of data (training data) as input and outputs the choicest hyperplane which classifies the given input data into two classes, moreover, the generated hyperplane is used to categorize new data. The support vectors are a data points of the training data, that are closest to the optimal hyperplane, if these data points are removed then they will alter the position of the dividing hyperplane. Because of this, these support vectors are considered critical elements of the data set which help in building the hyperplane. While generating this hyperplane the margin between these support vectors is maximized, so that the new data has the maximum likelihood of falling in the correct class. The generated hyperplane is n-dimensional, where n is the number of features in the training vectors. Some use cases of SVMs are classification, regression, outlier detection and clustering. We use one of these functions namely, regression, which identifies which side of the hyperplane the new input data will occupy.

Support Vector Machines are a relatively new form of supervised machine learning, compared to neural networks. SVMs are great for small data sets that contain few outliers. Whereas, other machine learning algorithms such as deep neural networks, use large sets of data, but always give very robust models. The decision of which classifier to use depends upon the size of the dataset and the general complexity of the problem.

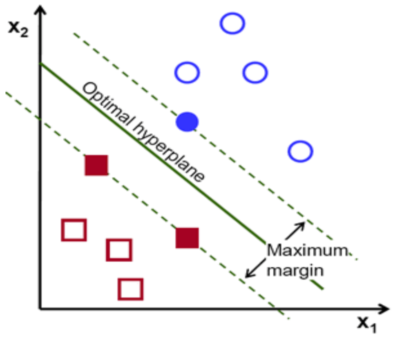

Fig 3.2: Optimal Hyperplane dividing the 2 classes [12]

In machine learning, we generally optimize an objective function, and in order to optimize an objective function we should minimize the error function. In SVMs, the use hinge loss as the error function.

4. Methodology

Firstly, we extract images from a benchmark dataset. This Facial Emotion Recognition System uses Biorthogonal Wavelet Entropy as a multiscale feature extraction technique. This system uses Support vector machine as a classifier. After extracting the features from the image, it trains the SVM. recognize the facial expression of the individual in the image. The SVM classifies the images into seven different classes: happy, sad, anger, fear, disgust, surprise and neutral.

4.1 Feature Extraction:

We use a multiscale feature extraction technique named as Biorthogonal
Wavelet Entropy (BWE). In this technique it first extracts the image's biorthogonal wavelet transform. This gives the approximation and detail coefficient sub bands. This is called 1-level BWT. Then it applies BWT again on an approximation coefficient sub band giving its own detail and approximation coefficient sub bands. This is called 2-level BWT. Then it applies Shannon Entropy over all the seven coefficient sub bands. We use Wavelet Transform Toolbox of MATLAB for extraction of the Biorthogonal Wavelet Transform of the image and went ropy library for calculation of the Shannon Entropy of each coefficient sub band.

The following figure shows the whole process of feature extraction.

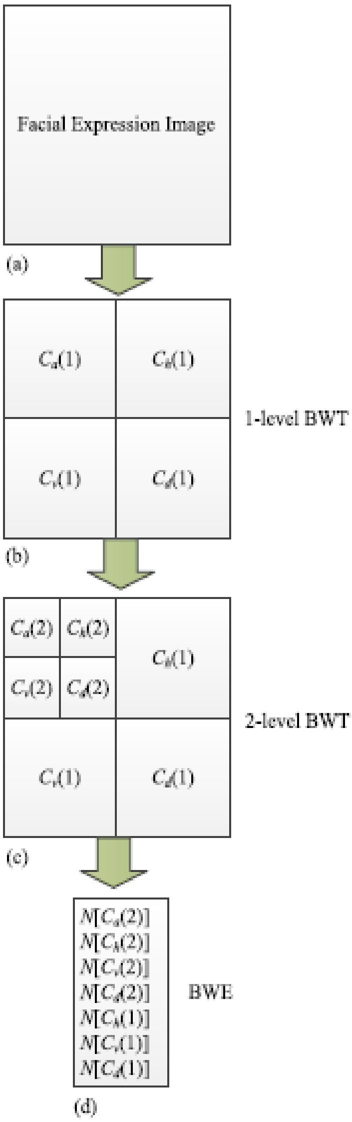

Fig 4.1: Block Diagram of a 2-level BWE: (a) A facial expression image; (b) 1-level BWT; (c) 2-level BWT; (d) Shannon entropy calculation. In this figure, (j) represents the j -level decomposition. [9]

4.2 Training the SVM:

This system uses a multiclass Support Vector Machine for classification of the features extracted from the images. A SVM can classify only between 2 classes, thus, this system uses a 7 class SVM where each class represents an emotion.

Once the SVM is structured for a particular classification, it is ready to be trained accordingly. To start this process, initial weights are chosen randomly. Then the training begins. In the training process, first both inputs are outputs are provided.

The SVM then processes the inputs and compares its resulting outputs against the desired outputs. As the SVM is trained, the optimal hyperplane is changing with each input and output. So, the more the SVM is trained the more will be the possibility of correctly identifying the emotion of the new input data.

4.3 Test and Validation:

Proposed approach will be tested and validated using Benchmark datasets and also one of our own datasets which will contain 140 images of 10 individuals with all 7 different emotions: happy, sad, angry, fear, disgust, surprise and neutral.

5. Hardware and Software Requirement

5.1 Software requirement:

- MATLAB 2012 or above
- Wavelet Toolbox
- Image Acquisition Toolbox

5.2 Hardware requirement:

- Minimum 4 GB RAM
- Webcam

5.3 Dataset:

+

5.3.1 Benchmark dataset: JAFFE[13]

Fig 5.1 : Seven Emotions on JAFFE Dataset

5.3.2 Dataset created by our group

Fig 5.2: Seven Emotions of own Dataset

6. Activity Chart

	Mid - Sem	End-Sem

Activity	Phase - I	Phase - II	Phase - III	Phase - IV
	August	September	October	November
Collection and reading of Literature Survey	14^{th} – 25^{th}			
Feature Extraction using BWT		1^{st} – 12^{th}		
Training			4^{th} –10^{th}	
Improvement and analysis			11^{th} – 25^{th}	
Test and validation of proposed approach using Benchmark datasets				1^{st} – 15^{th}

Table 6.1 : Activity Chart Table

7. Experimental Training And Results

7.1 Training on dataset

We started with training our system on all seven emotions on JAFFE dataset. We have used tenfold stratified cross validation to validate our system.
213 image models contain 30 images for ANGER, 29 images for DISGUST, 32 images for FEAR, 31 images for HAPPY, 30 images for NEUTRAL, 31 images for SADNESS and 30 images for SURPRISE.

Total classes of emotions in our dataset are 7. They are

- Anger
- Disgust
- Fear
- Happy
- Sad
- Surprise
- Neutral

7.2 Results

Using JAFFE dataset, we first trained our software on all seven emotions. Then there were similar results for disgust and anger so we trained on six emotions without disgust.

Following are the results when we train our software on JAFFE dataset and also on our dataset.

7.2.1 Feature Extraction

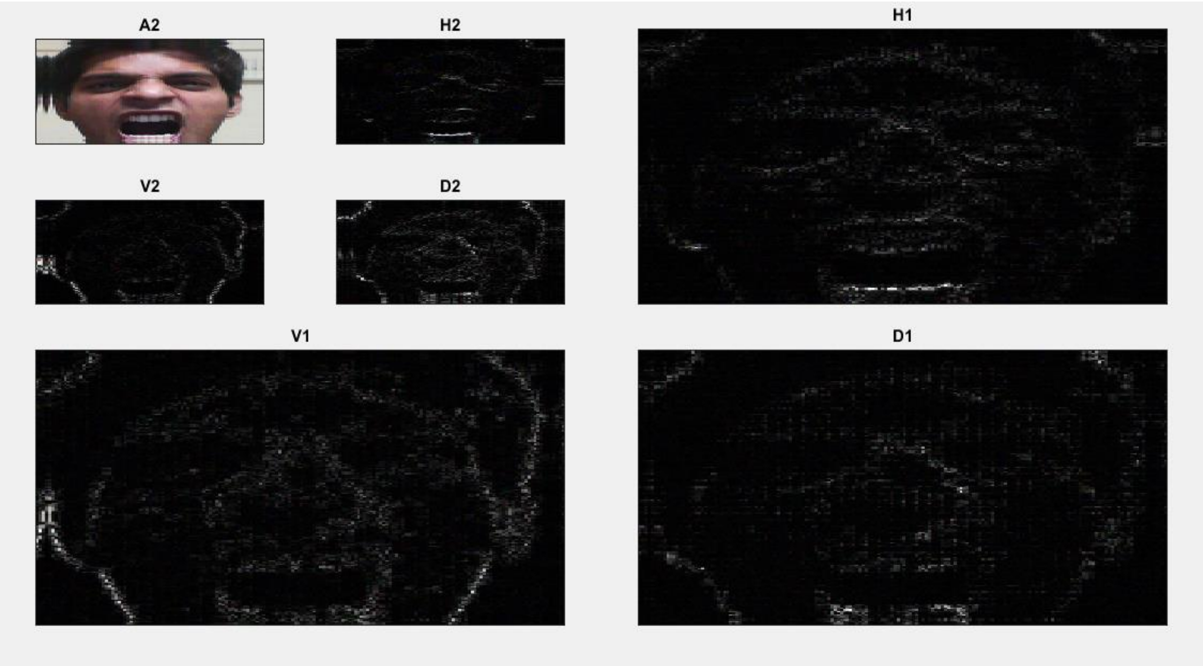

Fig 7.1: Feature Extraction using BWT

7.2.2 Results on all Seven Emotions

Emotion	Overall Accuracy
Anger	77.4%
Disgust	79.2%
Fear	70.3%
Happy	85.2%
Neutral	89.3%
Sad	88.9%
Surprise	69.8%
All 7 Emotions	**79.3%**

Table 7.1: Each emotion accuracy on JAFFE dataset

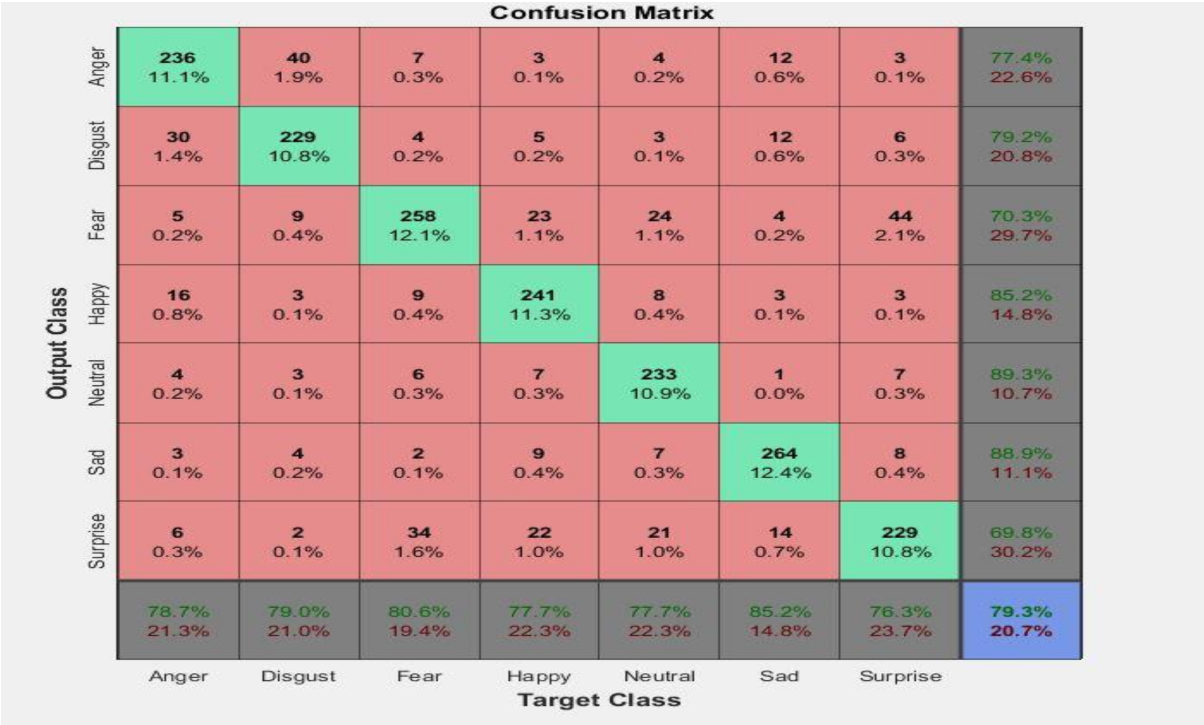

Fig 7.2: Confusion matrix for training on JAFFE dataset with 7 emotions (1. Angry 2.Disgust 3.Fear 4.Happy 5.Neutral 6.Sad 7.Surprise)

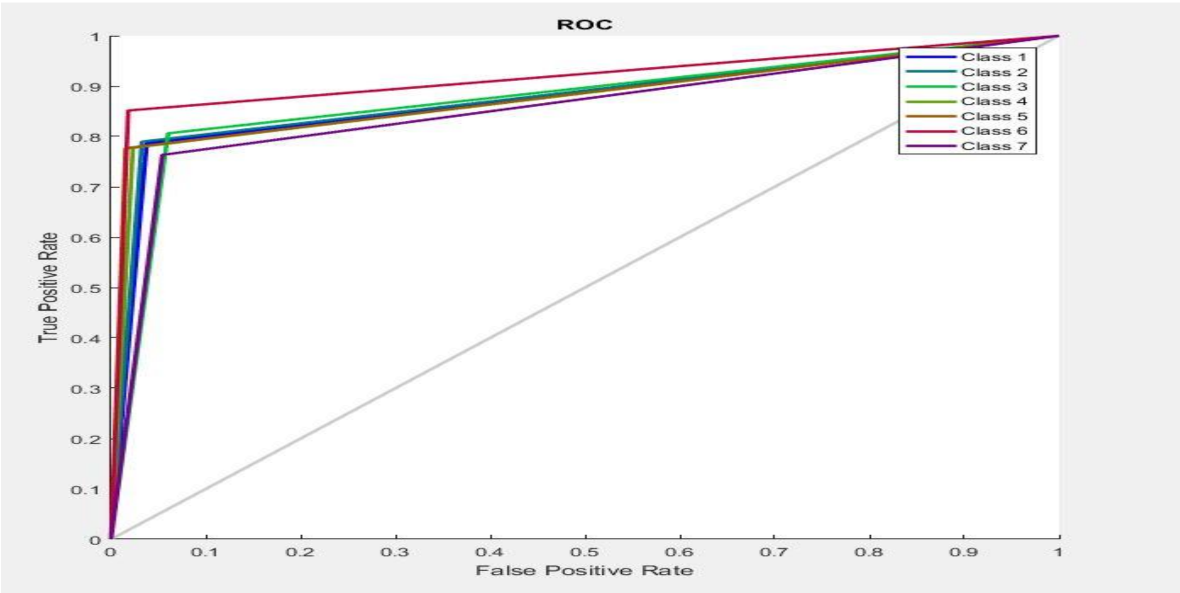

Fig 7.3 : ROC(Receiver Operator Characteristic) for training JAFFE dataset with 7 emotions (1.Angry 2.Disgust 3.Fear 4.Happy 5.Neutral 6.Sad 7.Surprise)

7.2.3 Results on 6 Emotions (Without Disgust)

Emotion	Overall Accuracy
Anger	86.9%
Fear	75.0%
Happy	89.3%
Neutral	81.1%
Sad	84.1%
Surprise	75.5%
All 6 Emotions	**81.7%**

Table 7.2 :Each emotion accuracy on JAFFE dataset (Without Disgust)

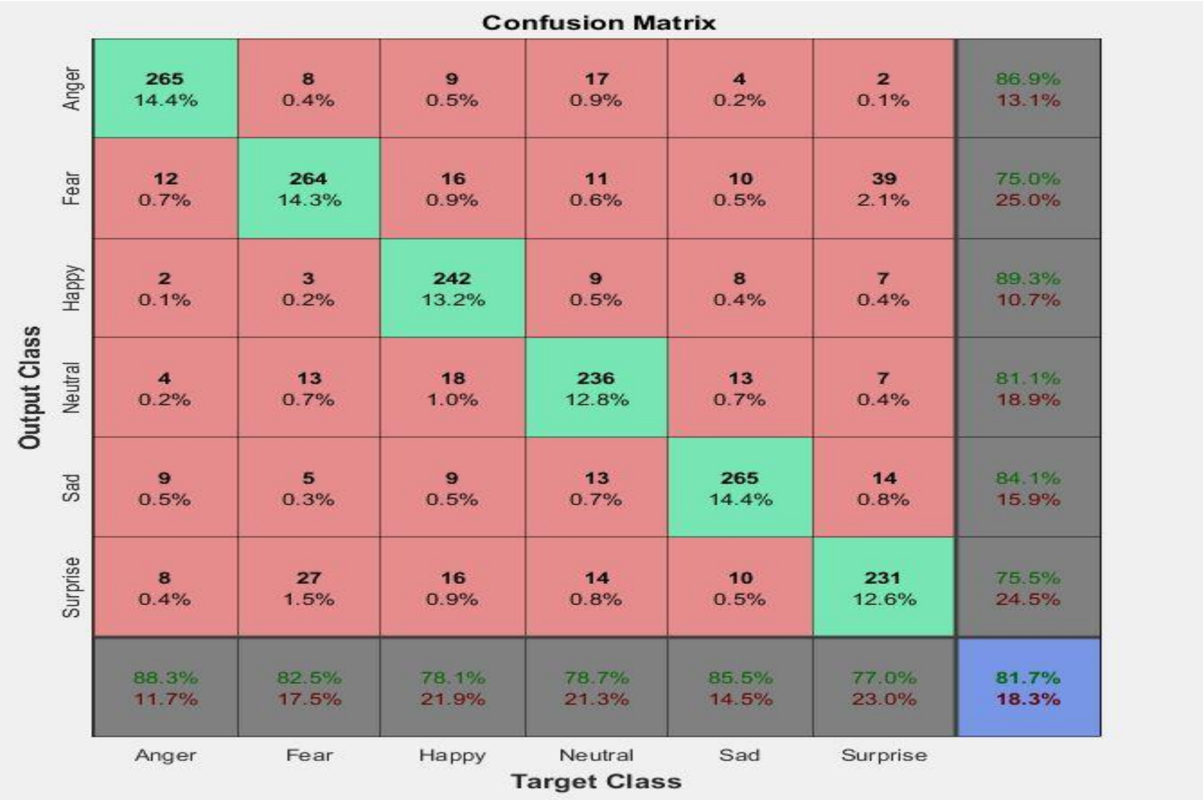

Fig 7.4 : Confusion matrix for training on JAFFE dataset with 6 emotions
(1.Angry 2.Fear 3 Happy 4 .Neutral 5.Sad 6.Surprise)

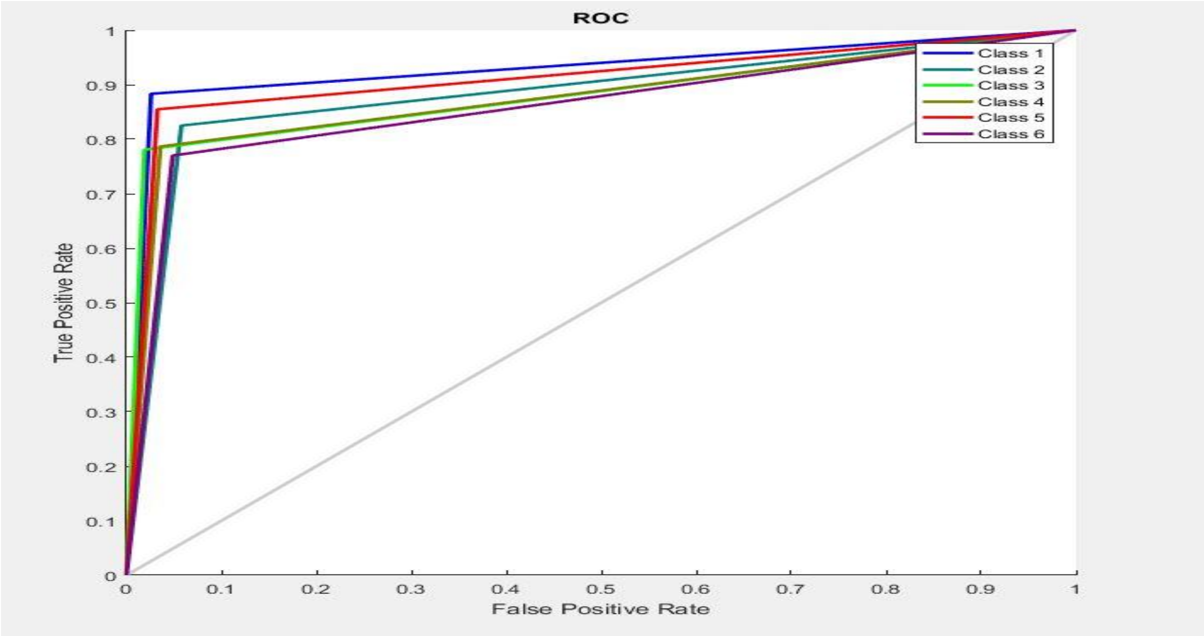

Fig 7.5 : ROC(Receiver Operator Characteristic) for training JAFFE dataset with 6 emotions (1.Angry 2.Fear 3 Happy 4 .Neutral 5.Sad 6.Surprise)

7.2.4 Results on own dataset

We created our own dataset by clicking images of our group members. Total number of images we used are 140 consisting of all seven emotions. Testing for our own dataset on our software lead to accuracy of 88.7%.

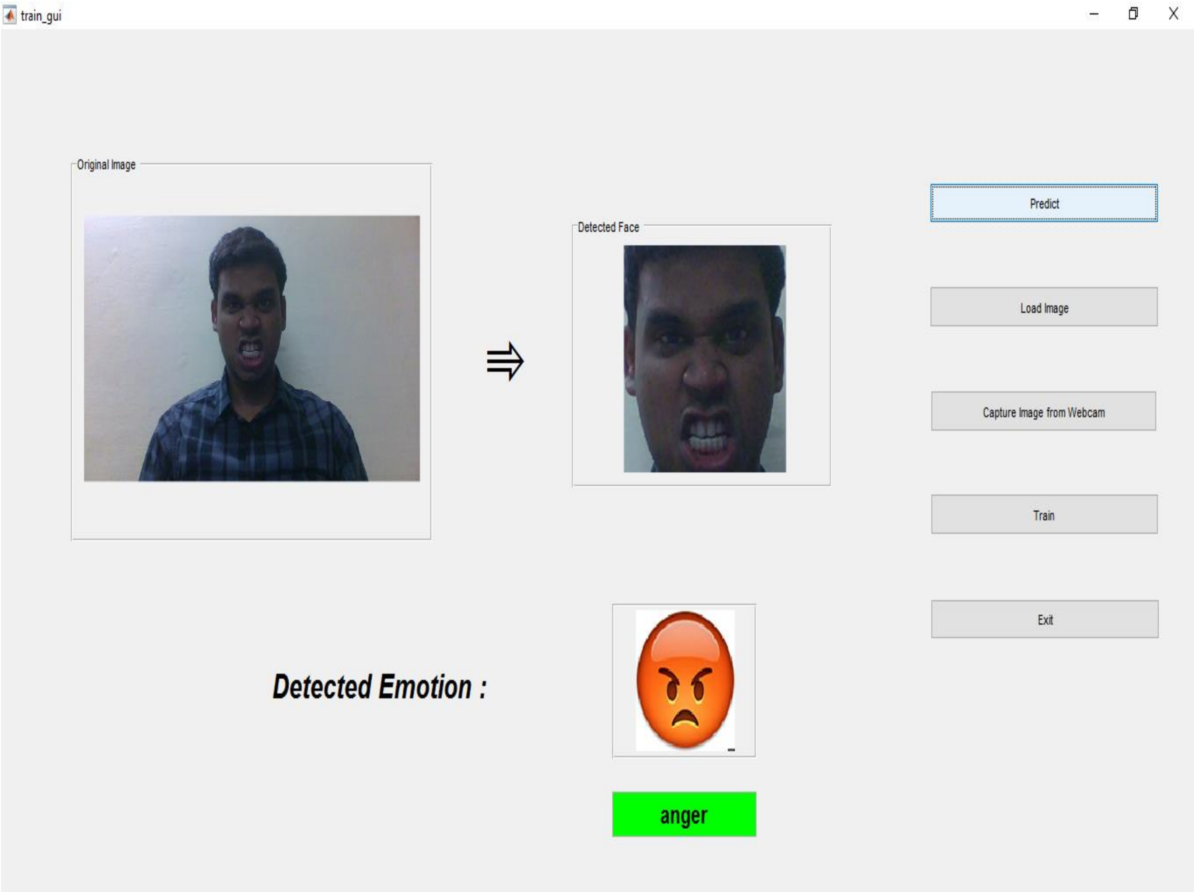

Fig 7.6 : GUI for project

7.3 Inferences

- When we train our software on all 7 emotions, the combined accuracy is 79.3% (refer Table 7.1)

- As we decrease the number of target emotions as in section 7.2.3 we observe a rise in target accuracy. This jump in accuracy is due to removal of overlapping emotional set which when removed increases the classification accuracy of our classifier.

- Increased accuracy after excluding disgust emotion is 81.7% (refer Table 7.2)

- The diagonal of Confusion matrix gives the number of images for which software predicts correct emotion.

- Receiver Operating Characteristic **(ROC)** curve the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points. The Curve is plotted for all emotions.

8. Comparison

Most applications of emotion recognition examine static images of facial expressions. One such technique use Convolutional Neural Networks (CNN)[14] for emotion recognition in real time with a video input stream. It's accuracy was found to be 57.1% on JAFFE Dataset.

Whereas the technique we used for classification is Support Vector Machine (SVM) which leads to accuracy of 79.3% on same JAFFE dataset.

9. Conclusion

In this work we have been able to analyze feature extraction from an image using Biorthogonal Wavelet Transform. We came to a conclusion that emotion recognition is dependent on the chosen Support Vector Machine classifier.

Furthermore we test our software on Standard and Benchmark dataset JAFEE, We have been able to achieve accuracy of 79.3% .

Further we decrease the number of target emotion and we found that accuracy increases. If we remove disgust emotion the accuracy increases to 81.7%.

.

10. Future Scope

1) Sometimes user emotion is not unique, it is mixture of two or more emotions. Fuzzy logic can be implemented and the algorithm and result is Mixed Emotion.

2) Our Project can be implemented in a music system which detect person's emotion and then play a suitable song accordingly.
For example if a person is sad, it will play a happy song and if a person is happy, a party song is played.

3) It is observed that physiological features like heart-beat rate can classify fear better than other features. Blood pressure is generally reduces in case of sadness. Hence inclusion of physiological features may lead to better accuracy.

11. References

[1] S. Meher and P. Maben, "Face Recognition and Facial Expression Identification using PCA", *IEEE International Advance Computing Conference*,2014.

http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6779478

[2] J. Li and M. Oussalah, "Automatic Face Expression Recognition System", *IEEE 9th International Conference*, 2010.
http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5898118

[3]P. Abhang and B. Gawali, "Emotion Recognition Using Electroencephalography (EEG) Images", *International Journal of Advanced Research in Computer Science*, 2014.
http://www.ijarcs.info/index.php/Ijarcs/article/download/2247/2235

[4] H. Rangnathan, S. Chakrabotry and S. Panchanathan, "Multimodal Emotional Recognition using Deep Learning Architectures", *IEEE Winter Conference on Applications of Computer Vision*, 2016.
http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7477679

[5] S. Happy, A. George and A. Routray, "A Real Time Facial Expression Classification System Using Local Binary Patterns", *4th International Conference on Intelligent Human Computer Interaction, Kharagpur ,India*, 2012.
http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6481802

[6] N. Mahajan and H. Mahajan, "Emotion Detection Algorithm", *International Journal of Electrical and Electronics Research*, 2014.
http://www.researchpublish.com/download.php?file=Emotion%20Detection%20Algorithm-222.pdf&act=book

[7] Z. Akhtar and C. Erdem, "Multimodal emotion recognition based on peak frame selection from video", *Signal, Image and Video Processing (Springer)*, 2015.
https://link.springer.com/article/10.1007/s11760-015-0822-0

[8] Y. Liu, Y. Li, X. Ma and R. Song, "Facial Expression Recognition with fusion features extracted from salient facial areas", *Sensors (MDPI)*, 2017. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5421672/

[9] Y. Zhang, Z. Yang, H. Lu, X. Zhou, P. Phillips, Q. Liu and S. Wang, "Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation", *IEEE Access*, vol. 4, 2016. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7752782

[10] "Wavelet", *En.wikipedia.org*, 2017. [Online]. Available: https://en.wikipedia.org/wiki/Wavelet. [Accessed: 13- Sep- 2017].

[11]"Artificial Intelligence Fuzzy Logic Systems", *www.tutorialspoint.com*, 2017.[Online].Available: https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_fuzzy_logic_systems.htm. [Accessed: 13- Sep- 2017].

[12]"Introduction to Support Vector Machines — OpenCV 2.4.13.3 documentation", *Docs.opencv.org*, 2017. [Online]. Available: http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html. [Accessed: 13- Sep- 2017].

[13] "Facial Expression Database: Japanese Female Facial Expression (JAFFE) Database", Kasrl.org, 2017. [Online]. Available: http://www.kasrl.org/jaffe.html. [Accessed: 20- Nov- 2017].

[14] D. Duncan, G. Shine and C. English, "Facial Emotion Recognition in Real Time", 2017. http://cs231n.stanford.edu/reports/2016/pdfs/022_Report.pdf