A Project/Dissertation Final Project Report

on

Face Expression to Emoji

Submitted in partial fulfillment of the requirement for the award the degree of B.TECH CSE

Under The Supervision

Mr.Ravi Sharma

School of Computing Science and Engineering, Galgotias University

Assistant Professor
Greater Noida, Uttar Pradesh,
India.

Submitted By

Group ID – BT3097

S.NO	NAME	ENROLLMENT NO	ADMISSION NO	PROG/BRAN	SEM
1	Kapil Tyagi	19021011360	19SCSE1010165	B.TECH/CSE	5
2	Abhishek Kumar	19021011329	19SCSE1010132	B.TECH/CSE	5

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GALGOTIAS UNIVERSITY, GREATER NOIDA

INDIA

DEC, 2021

SCHOOL OF COMPUTING SCIENCE AND
ENGINEERING GALGOTIAS
UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled **Face Expression to emoji** in partial fulfillment of the requirements for the award of the B.TECH CSE submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of

SEP, 2021 to DEC,2021, under the supervision of **Mr. Ravi Sharma** Designation, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

Kapil Tyagi , 19SCSE1010165

Abhishek Kumar , 19SCSE1010132

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

Supervisor Name

Designation

CERTIFICATE

The Final This/Project/ Dissertation Viva-Voce examination of Kapil Tyagi – 19SCSE1010165 ,Abhishek Kumar 19SCSE1010132 has been held on 16/10/2021 and his/her work is recommended for the award of B.TECH SCSE.

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: DEC, 2021

Place: Greater Noida

Acknowledgement

I would like to express my special thanks of gratitude to my Group mate as well as my teacher **Mr. Ravi Sharma** who gave me the golden opportunity to do this wonderful project on the topic **Face expression to emoji** which also helped me in doing a lot of research and I came to known about so many new things I am really thankful to them.

Secondly, I would also like to thank my parents and friends who helped me a lot in finalizing this project within the limited time frame.

Abstract

Emoji characters are becoming more and more popular for communicating emotions along with written messages. However, emoji characters are restricted to predefined characters. These characters lack originality and creativity, so in order to make emoji characters personalized and creative, this study explored and finalized algorithms that allows the user to "exotify" any image. We used localized segmentation to detect and isolate the desired object of the user. The algorithm then applies a bilateral and cartooning effect to make the image more abstract and smooth. At the end, Gaussian pyramids were used to resize the image to a smaller dimensions appropriate for communication platforms. This study further recognized the limitations and possible optimizations that could improve the application for a wider range of users.

Existing Problem:

Communication is an important part of everyday life. Verbal or non-verbal communication allows one to engage in conversations. Today is the era of communication technologies. The internet and other communication devices have made it possible to engage in the fast, dynamic, and affective communication. The emojis are being used for the visual depictions of human emotions. This paper presents the visual expressions of humans using emojis. Therefore, the efforts have always been made to enhance the experience of communication in human beings. The present study has three major objectives to be accomplished. These three major objectives are to investigate emotional recognition using facial expression by emojis in real time. Moreover, another objective is to develop the parameters of measuring the facial expression by emoji texting, as well as to understand the facial emotion recognition in real time.

Therefore, the present study in that realm focuses on the use of the emojis as an important technique that can be used in the field. Hence, the present research will be designed to assess human emotions and their visual depiction in the form of six emotions. These emotions include neutral, fear, anger, happy, sad, and surprise emotions. These are important human expressions and, thus, are quite important. The channels show the convolution of 3D delineation models. These models are prone to the use of layers, which records clamor in process. Moreover, these applications will be developed in the study to assess its impacts on the emojis.

Proposed Solution:

This research study presents how the facial looks and discovery for the facial communication is delivered. Therefore, it can be said that the facial expression is important for the review of the investigation of the facial activity coding framework for the sake of the communication that is being watched. Moreover, the facial expressions are critical as they help incorporate and communicate feelings to discover the activity units. Therefore, this study develops an application that impacts communication channels, which are designed for human conduct, as well as to identity someone.

Therefore, the use of the emoji provides the feelings that have a profound impact that has to be delivered, and with the expression and the facial recognition, which is being conducted, our general identity is restored through the emoji. These are important things that must be considered and investigated in current research fields. The emojis will be developed, and the human expressions will be tested through the use of such emojis.

Tools and Technology Used :

Tools and techniques that we will be using are Open CV , Dlib, pyhthon **,**

Scikit learn, jupyter notebook, Database

Results:

. In the above figure, i.e., Sample output is the window where the user expressions are captured by the web cam and the respective emotions are detected. On detecting the emotion, respective emoticon is shown on the left side of the screen. This emoticon changes with the change in the expression of the person in front of the web cam. Hence changes with the change in the expression of the in front of the web cam. Hence this real time application is very beneficial in various fields like psychology, computer science ,linguistics ,neuroscience and related disciples.

Conclusion and Future Scope :

Proposed is a human emotion detector using emoticon using machine learning, python to predict emotions of the people and represent them using emoticon.These include image acquisition, preprocessing of an image, face detection, feature extraction, classification and then when the emotions are classified the system assigns the user particular music according to his emotion. Our system focuses on live videos taken from the webcam.
The main aim of this project is to develop automatic facial emotion recognition system in which an emoticon is used for giving the output for individuals thus assigning them various therapies or solutions to relief them from stress. The used for the experiments include happiness, , Surprise, Fear, Disgust, and Anger that are universally accepted.

Table of Contents

1. Introduction

Facial emotion recognition and analysis has been gaining a great attention in the advancement of human machine interface as it provides a natural and efficient way to communicate between humans. Some application are as related to face and its expressions include person identification and access control, video call and teleconferencing, forensic applications, human-computer interaction, automated surveillance, cosmetology, and so on. But the performance of the face expression detection certainly affects the performance of all the applications. Many methods have been suggested to detect human face in pictures and videos, they can be divided into four types: knowledge-based methods, feature-based methods, template based methods and appearance-based methods. When these methods are used separately, they cannot solve all the problems of face detection like pose, expression, orientation. Hence it is recommended to operate with several successive or parallel methods. Most of the existing facial expression recognition methods which are popularly used till today are focused on recognition of five primary expression categories such as: happiness, sadness, fear, anger and disgust.

1.2 Formulation of Problem

1.2.1 Tools and Technology Used

OpenCV

OpenCV is the library we will be using for image transformation functions such as converting the image to grayscale. It is an open source library and can be used for many image functions and has a wide variety of algorithm implementations. C++ and Python are the languages supported by OpenCV. It is a complete package which can be used with other libraries to form a pipeline for any image extraction or detection framework. The range of functions it supports is enormous, and it also includes algorithms to extract feature descriptors.

Dlib

Dlib is another powerful image-processing library which can be used in conjunction with Python, C++ and other tools. The main function this library provides is of detecting faces, extracting features, matching features etc. It has also support for other domains like machine learning, threading, GUI and networking.

Python

Python is a powerful scripting language and is very useful for solving statistical problems involving machine learning algorithms. It has various utility functions which help in pre- processing. Processing is fast and it is supported on almost all platforms. Integration with C++ and other image libraries is very easy, and it has in-built functions and libraries to store and manipulate data of all types. It provides the pandas and numpy framework which helps in manipulation of data as per our need. A good feature set can be created using the numpy arrays which can have n-dimensional data.

Scikit-learn

Scikit-learn is the machine learning library in python. It comprises of matplotlib, numpy
and a wide array of machine learning algorithms. The API is very easy to use and understand. It has many functions to analyze and plot the data. A good feature set can be formed using many of its feature reduction, feature importance and feature selection functions. The algorithm it provides can be used for classification and

regression problems and their sub-types.

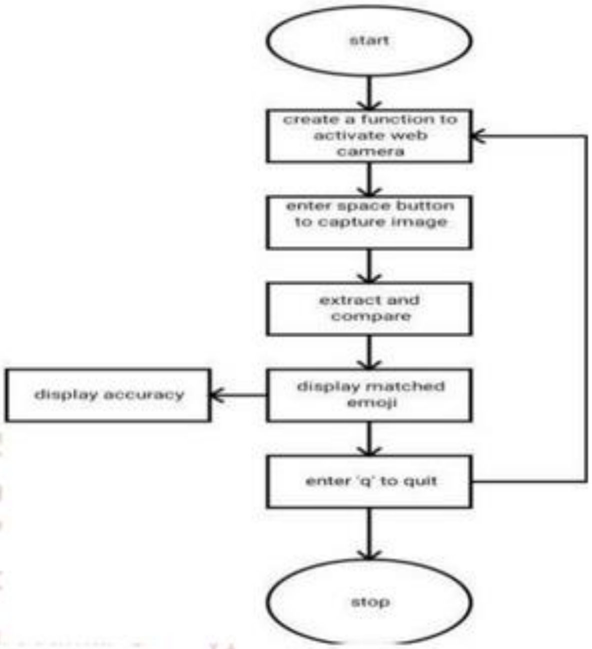
Jupyter Notebook

Jupyter Notebook is the IDE to combine python with all the libraries we will be using in our implementation. It is interactive, although some complex computations require time to complete. Plots and images are displayed instantly. It can be used as a one stop for all our requirements, and most of the libraries like Dlib, OpenCV, Scikit-learn can be integrated easily.

Database

We have used the extended Cohn-Kanade database (CK+) and Radbound Faces database(RaFD). CK+ has around 593 images for 123 subjects. Only 327 files have labeled/identified emotions. It covers all the basic human emotions displayed by the face. The emotions and codes are as follows: 1 – Angry, 2 – Contempt, 4 – Fear, 5 – Happy, 6 – Sadness, 7 – Surprise. The database is widely used for emotion detection research and analysis. There are 3 more folders along with the images. FACS contains action units for each image. Landmark contains AAM tracked facial features of the 68 facial points. Emotion contains emotion label for the 327 files. Radbound Faces Database [19] is a standard database having equal number of files for all emotions. It has images of 67 subjects displaying 8 emotions: Neutral included. The pictures are taken in 5 different camera poses. Also, the gaze is in 3 directions. We are using only front 22 facing images for our experiment. We have a total of 536 files with 67 models displaying 8 different emotion

Methods-



Facial expressions can be described as the arrangement of facial muscles to convey a certain emotional state to the observer in simple words. Emotions can be divided into six broad categories—Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Neutral. In this, train a model to differentiate between these, train a convolutional neural network using
the FER2013 dataset and will use various hyper-parameters to fine-tune the model.

2. Literature Survey

The role of emotion is evident in our daily lives. Human beings use different kinds of emotions to show compassion and establish relationships with others (Seiter, 2016). These emotions express the emotional conditions in our daily lives. The comprehensive list of emotions range from anger to happiness, wondering, suspicion, skepticism, sorrow and grief. However, they are frequently witnessed in our daily lives. Therefore, it is quite easy to understand the inner feelings of a person with the use of the facial expressions that are quite visible. Thus, the facial expressions and emotional recognitions are interrelated with each other (Beswick, 2014). Using facial expressions with ideograms and smileys, is the emoji. The Japanese word, "emoji" consist of two parts: "the *e* means "picture" and *moji* means "letter"" (Zhou, Hentschel, & Kumar, 2017, p. 2). Emoticons were used before emoji as "symbolic representations for facial expressions based on punctuation marks that could be covered using a standard keyboard (e.g. :))" (Zhou et al., 2017, p. 2). Both emojis and emoticons are frequently used in the text messaging, emails, and other electronic forms of communication.

Emojis are a part of the life which was first introduced by Japanese mobile phone companies, such as Vodafone and NTT DocoMo. An early nineties was the period when Japanese companies enabled the use of the emoji in their communication via electronic devices. They were the pioneers in the use of the emoji. Through these companies, the trend enhanced and the other companies also came forward, and used these emojis to make the communication better (Guinness, 2015; Lee, 2012). Emojis became popular worldwide and are widely being used in the world at an international level (Danesi, 2016). The emoji was adopted by Apple Inc.; the corporation recognized the use of the smileys and other electronic pictorial symbols to show what the sender is feeling. Besides the text meaning, the pictorial smileys and other expressional symbols were important because they provide the opportunity to show the inner feelings of the sender. After the adaptation of iPhones, the other phones such as Samsung also used these methods. Now, it is used worldwide.

The uses of the smileys are important in day-to-day lives to show facial expression. The use of smileys and other pictorial images in every platform is common. These platforms include Android and Windows. The meaning of the word emoji in original form is pictograph (Gamble & Gamble, 2013). Emoji is now available in colorful forms. For the time being, it has progressed and now multiple forms and types are available through the internet and communication devices. In the beginning, the emoji was only available in the form of black and white shapes and it was also in a basic format. However, at the current time, they have been developed effectively and they are available in variety of shapes. Using emojis has increased the effectiveness of the use of the symbols. The use of symbols is also seen in other communication formats. Besides with electronic devices, the symbols are widely used on the internet. Today is the era of internet as well as communication and information technologies (Goss, Anthony, Stretch, & Nagel, 2016). Thus, the use of this communication innovation is evident. No one can deny its importance and the organizations engaged in the communication are well aware of the importance of all the techniques that are suitable for enchasing the effectiveness of the communication. Therefore, the use of the emoji is quite common and has developed over time. The first international conference on the emoticon was conducted in 2016. This is an important progress that has been noted or seen on the topic (Goss et al., 2016). It is expected that, for the time being, a lot of other initiatives must be taken to make sure that the emojis are incorporated with emoticons to enhance their effectiveness. Thus, that was the brief history about the emoji and its use in communication via electronic devices. In addition, emojis are used in publications and social media as well (Highfield & Leaver, 2016).

The importance of the emojis is twofold. Emojis show feelings because they express emotions which makes them important. We often times witness in our day to day lives the message conveyance. Emojis suggest that an emotional communication is as significant as using words. By using smileys, readers can understand the sender's sentiments. The use of the smileys and other symbols are imperative in our everyday lives to demonstrate effective communication. That is why more often the smiley signs are used in a time of happiness, an important advancement in communication. For example, in the face-to-face communication, if people do not use gestures or expressions in their conversations, his or her conversation becomes less meaningful. So, it can be said that these are the important considerations in the realm of the development of the symbols and their ultimate use in the communication patterns. Besides this, the emoji are considered as non-verbal tools. The emoji are the most powerful tools to facilitate the communication and allows people to express their linguistic capabilities (Lingard et al., 2005). An investigation by was also done on the use of the emoji in real time communication to make sure whether the emojis are effectively communicating. For this purpose, the scientists experimented and analyzed that when people use the sign of anger in their day to day communication, the receiver perceives that the person to whom they are in talking is angry and they apologize in that context. They also use the sign of the apology. Thus, this implies that the use of the emoji in relation to strengthen the language is important. The people understand their linguistic cues, and they use emoji in that context. Hence, the emojis are the powerful tools to demonstrate the freedom of expression and the nonverbal cues in our daily textual language.

The symbols and the pictures that are used in the emoji have different colors, and these colors provide the best support for the representation of the emotions and the facial expressions. The usage of the appropriate colors and the style is an important indication of the effectiveness of the emoji (Seiter, 2016). The body and skin colors are used in those symbols and the pictures are related to the human body. The use of the skin color is significant and promising because it must match with the reflection of facial expressions. Both the Windows and the Android platforms are used by millions of the people in the world. To ensure an effective communication, these platforms allows users to use symbols. The Windows 8.1

and onward are using smileys and the symbols. Any platform in the world supports different fonts. Even MS Word provides different kinds of the facial expressions. As mentioned earlier, the Japanese organizations have additionally delivered the images and the smileys that are especially identified with the way of life of the Japan, and they are not subject to the impression of the feelings of the entire world. Therefore, they are operable in any platform.

Furthermore, the use of the emoji has great cultural influence in terms of the facial expressions that are expressed by the emojis (Iemoji, 2015; Emojipedia, 2016). The Oxford dictionary also considered the year 2015 as the most influential in terms of the development of the emoji and its impact on the culture. The emoji of the year was the smiley with tears of joy. This emoji best illustrated and expressed the emotions of the humans when they express their love and joy for life. The uses of the emoji is worldwide and the representation of the emotions by the people is evident in every culture and the environment. The emoji is also the reflection of the specific culture of the country. Emojis provide the appearance of the facial expressions and emotions that are specific to each culture. For instance, the Japanese companies have also produced the symbols that are particularly related to the culture of the Japanese and they are not subject to the reflection of the emotions of the whole world. Therefore, the role of the diversity of the culture is essential and the emojis have a profound impact on the life of the people.

.

Image Features

We can derive different types of features from the image and normalize it in vector form.We can employ various types of techniques to identify the emotion like calculating the ellipses formed on the face or the angles between different parts like eyes, mouth etc. Following are some of the prominent features which can be used for training machine learning algorithms

FACS

Facial Action Coding System is used to give a number to facial moment. Each such number is called as action unit. Combination of action units result in a facial expression. The micro changes in the muscles of the face can be defined by an action unit. For example, a smiling face can be defined in terms of action units as 6 + 12, which simply means movement of AU6 muscle and AU12 muscle results in a happy face. Here Action Unit 6 is cheek raiser and Action Unit 12 is lip corner puller. Facial action coding system based on action units is a good system to determine which facial muscles are involved in which expression. Real time face models can be generated based on them.

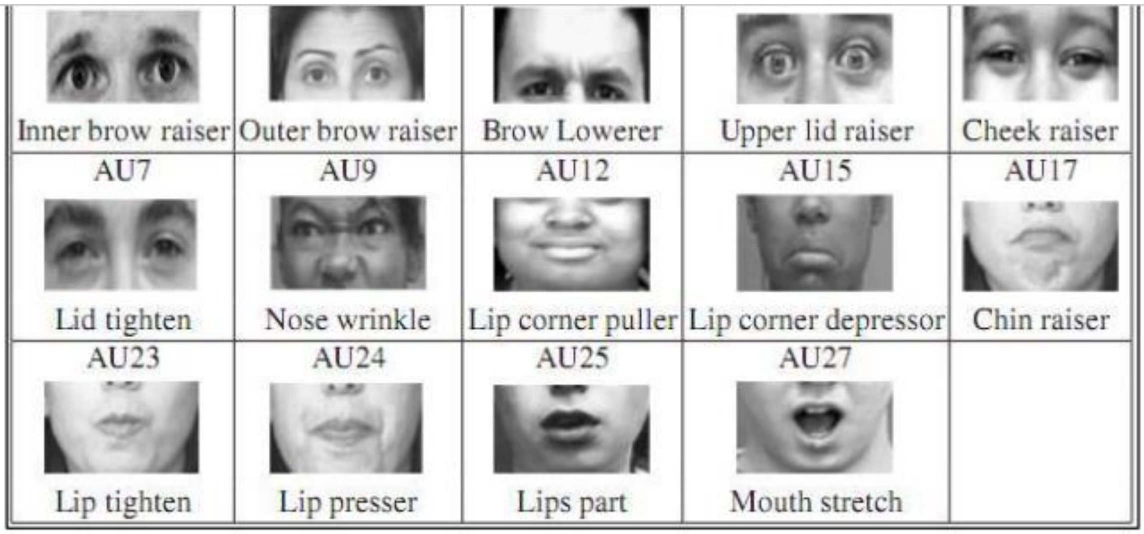


Figure 1: Action Units corresponding to different movements in face [15]

3

Landmarks

Landmarks on the face are very crucial and can be used for face detection and recognition. The same landmarks can also be used in the case of expressions. The Dliblibrary has a 68 facial landmark detector which gives the position of 68 landmarks on the face.

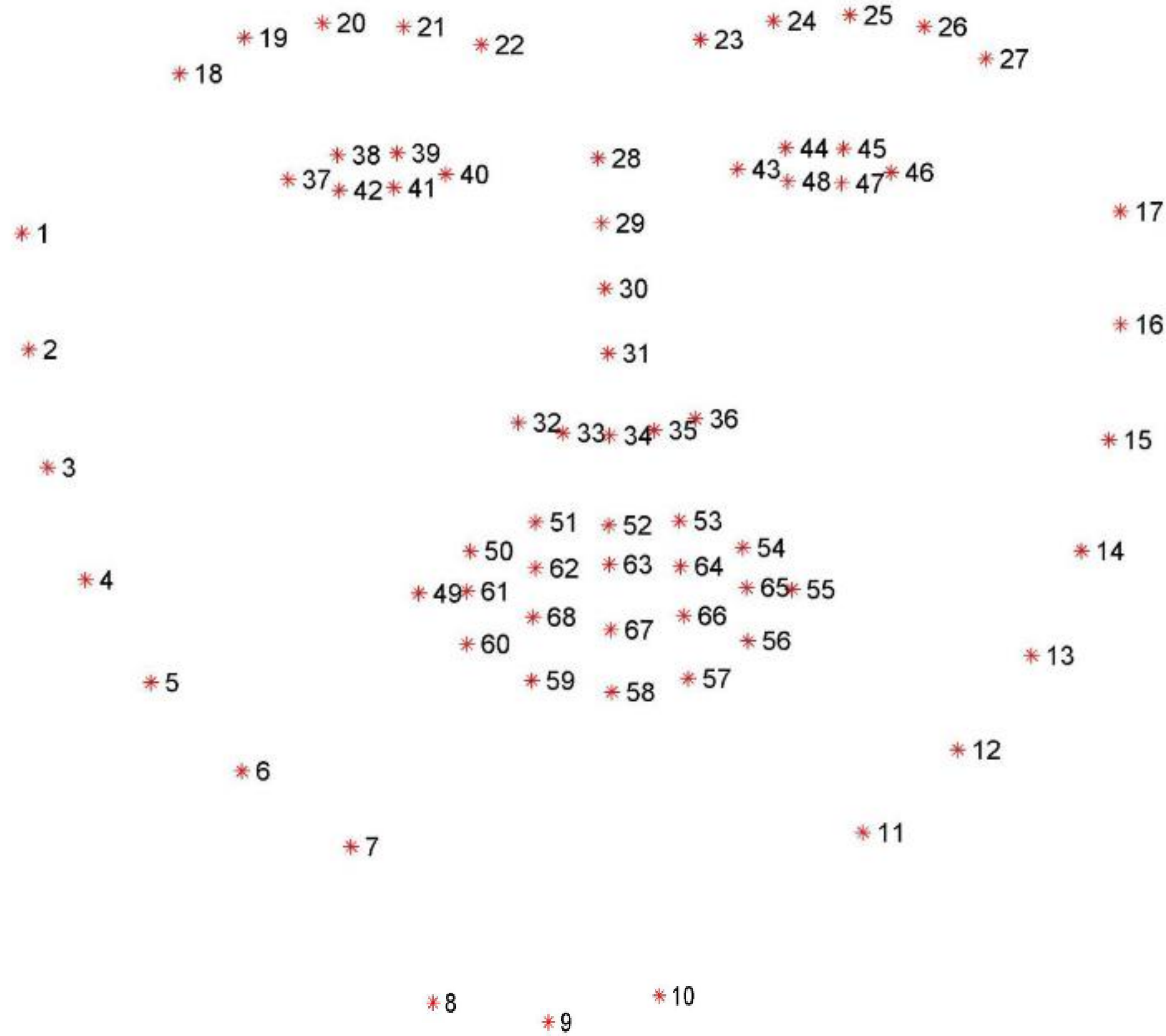


Figure 2: Landmarks on face [18

Figure 2 shows all the 68 landmarks on face. Using dlib library we can extract the co-ordinates(x,y) of each of the facial points. These 68 points can be divided into specific areas like left eye, right eye, left eyebrow, right eyebrow, mouth, nose and jaw.

Feature Descriptors

Good features are those which help in identifying the object properly. Usually the images are identified on the basis of corners and edges. For finding corners and edges in images, we have many feature detector algorithms in the OpenCV library such as Harris corner detector. These feature detectors take into account many more factors such as contours, hull and convex. The Key-points are corner points or edges detected by the feature detector algorithm. The feature descriptor describes the area surrounding the key-point. The description can be anything including raw pixel intensities or co-ordinates of the surrounding area. The key-point and descriptor together form a local feature. One example of a feature descriptor is a histogram of oriented gradients. ORB (based on BRIEF), SURF, SIFT etc. are some of the feature descriptor algorithms.

Related Work

Feature Extraction Techniques
Ensemble of regression trees

This method uses cascaded regression trees and finds the important positions on the face using images. Pixel intensities are used to distinguish between different parts of the face, identifying 68 facial landmarks [1]. Based on a current estimate of shape, parameter estimation is done by transforming the image in the normal co-ordinate system instead of global. Extracted features are used to re-estimate the shape parameter vectors and are recalculated until convergence [5].

3. Project Design

A static approach using extracted features and emotion recognition using machine learning is used in this work. The focus is on extracting features using python and image processing libraries and using machine learning algorithms for prediction. Our implementation is divided into three parts. The first part is image pre-processing and face detection. For face detection, inbuilt methods available in dlib library are used. Once the face is detected, the region of interest and important facial features are extracted from it. There are various features which can be used for emotion detection. In this work, the focus is on facial points around the eyes, mouth, eyebrows etc.We have a multi-class classification problem and not multi-label. There is a subtle difference as a set of features can belong to many labels but only one unique class. The extracted facial features along with SVM are used to detect the multi-class emotions. The papers we have studied focus on SVM as one of the widely used and accepted algorithms for emotionclassification. Our database has a total of 7 classes to classify. We have compared our results with logistic regression and random forest to compare the results of different algorithms. The processing pipeline can be visualized

Setting up the database

The image files for the CK+ database are in different directories and sub-directories based on the person and session number. Also, not all the images depict emotion; only 327 files have one of the emotion depicted from 1-7. All the files were of type portable networks graphic file(.png). The emotion labels are in the different directory but with the same name as image files. We wrote a small utility function in java which used the emotion file name to pick up the correct image from the directory and copy it in our final dataset folder. We also appended the name of the emotion file to the image file name. Thus, while parsing the file in our program we will have the emotion label for that file.

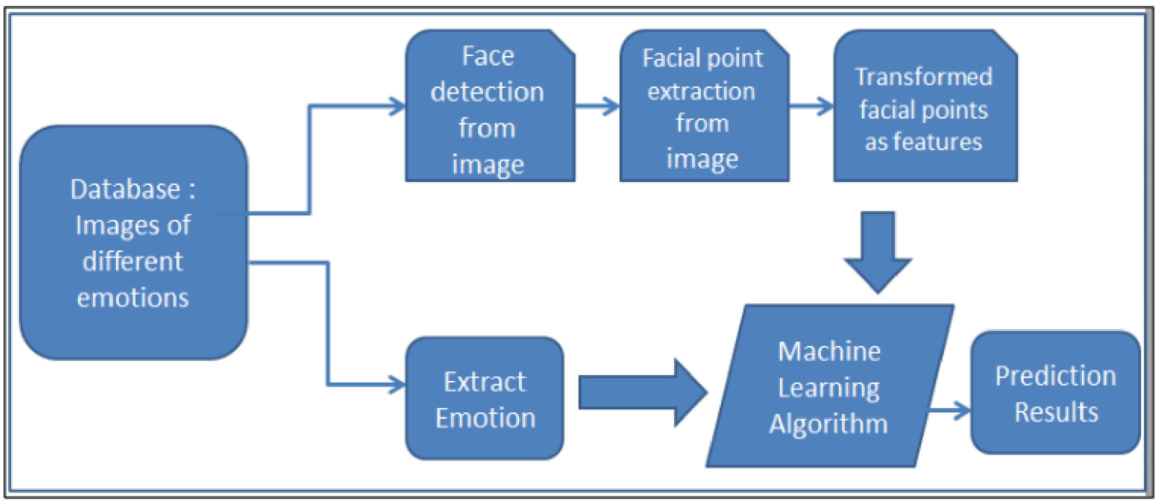


Figure 13: Implementation Pipeline

4. Module Description

The facial expression is specifically based on the emoticon identification system. It is identified as an open source extension to the tracker.js (A modern approach for Computer Vision on the web) framework that converts into expression of human facial to match best emoticon (Sashikar et al., 2015, p. 1). The emoticons are the major parts towards the digital communication system. Emoticons, are also called, emoji's that "are used to express the emotions of a person through text in a way that is not possible with just words" (Sashikar et al., 2015, p. 1). As mentioned by Sashikar et al. (2015), "the importance of emoji has become so huge that they [also have been specifically] annotated with the WordNets" (p. 1). The vision of computer has grown at large scale that use as commercial product for the benefits of high speed image (Sashikar et al., 2015, p. 1).

The key task at hand ensures the face expression and face detection for the recognition of emotion (Russell, 1994, p. 103). The formed task behind the face detection and emotional recognized gained from the eigen faces, fisher faces, and viola jones are the detections framework that prologs the hausdorff distance (Sashikar et al., 2015, p. 1). The study exploration of the facial action coding system that observed the facial muscles, that plays an important role in the compiled and expressing emotions to find out the action units. Such action units are raising from the outer eyebrow and inner eyebrow, which is important for the quantification of human expressions (Sashikar et al., 2015, p. 1).

5.Result

We applied support vector machines to our dataset and predicted the results. The results were interpreted using confusion matrix and accuracy metric. The train:test split was 75:25. We also did cross-validation on the dataset to remove any biases. Value of split was chosen as 4 because the resultant splits will have same number of images as our 25% test set. The results are as follows

Table 6: Accuracy for 75:25 split and cross-validation

SVM kernel	Accuracy (%)	Cross-Validation Accuracy Score (cv=4)
linear	78.05	0.78(+/- 0.07)
rbf	21.95	0.25(+/- 0.01)
poly	75.61	0.76(+/- 0.06)

Table 6: Accuracy for 75:25 split and cross-validation

In our experiment SVM with linear kernel performed better than other kernels. Rbf gave us the worst performance, whereas poly was as good as linear kernel. We tried to keep the test set % same for both split and cross-validated data so as to have uniformity in results. The mean cross-validation score was also approximately equal to the accuracy score achieved by the split. Figure 17 shows the heat-map of the confusion matrix from our multi-class classification results. On further analysis of the confusion matrix, we see that the diagonals have higher weights; thereare a few misclassifications in every class except class 4: Fear

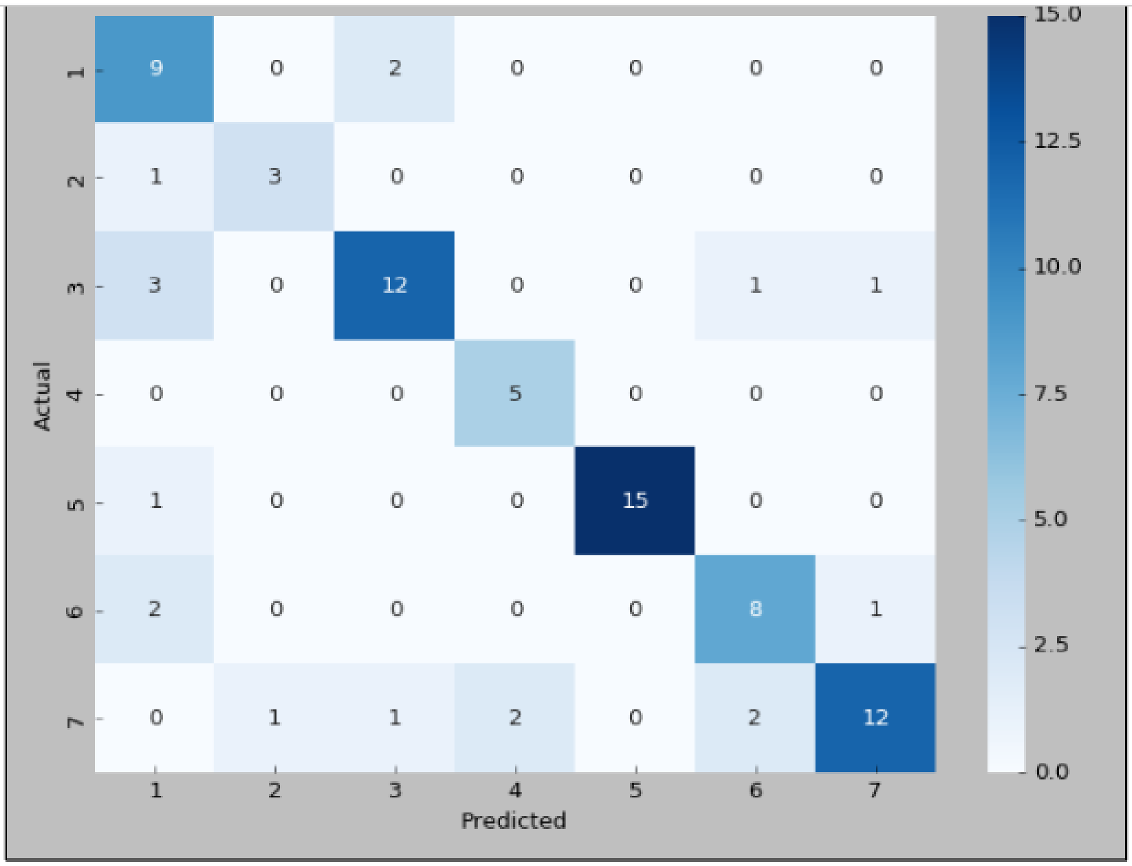


Figure 17: Heat map of actual and predicted values

Table 7: Report of predicted values Vs actual values

Predicted / Actual	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise	All
Anger	**9**	0	2	0	0	0	0	**11**
Contempt	1	**3**	0	0	0	0	0	**4**
Disgust	3	0	**12**	0	0	1	1	**17**
Fear	0	0	0	**5**	0	0	0	**5**
Happy	1	0	0	0	**15**	0	0	**16**
Sadness	2	0	0	0	0	**8**	1	**11**
Surprise	0	1	1	2	0	2	**12**	**18**
All	**16**	**4**	**15**	**7**	**15**	**11**	**14**	**82**

Table 7 shows the predicted values and actual values in report format. From this, we can inferthe correct number of emotions predicted for each class.
1: Anger – 9/11 – 82%
2: Contempt – 3/4 - 75%
3: Disgust – 12/17 - 70%
4: Fear – 5/5 - 100%
5: Happy – 15/16 - 93%
6: Sadness – 8/11 - 72%
7: Surprise – 12/18 – 66%
Many of the samples were misclassified into other classes such as contempt, disgust, fear and sadness. Fear was detected accurately for all the samples. The samples for disgust and sadness were misclassified as anger in some cases. This can be due to the similarity of features between these emotions. For our next part, we considered all the 68 facial point features and determined which of these features actually help to determine the emotion. Jaw was not adding any difference in all of the micro-expressions. Hence, we decided to ignore the jaw facial landmark positions andcontinue with eyes, eyebrows, mouth and nose for our analysis. Our feature set was still the same
with dimensions 327*68*2, but all the jaw facial landmarks had co-ordinates as 0. Also, we kept

the split ratio of test-train dataset same as above. Following are the results for this experiment.

SVM kernel	Accuracy (%)	Cross-Validation Accuracy Score (cv=4)
linear	82.93	0.80(+/- 0.04)
rbf	21.95	0.25(+/- 0.01)
poly	80.49	0.78(+/- 0.07)

Our focus is on the cross-validation score as it takes entire database into consideration for training, thus removing any biases. After removing jaw-line features the cross-validation results also increased from 78% to 80%. For this experiment, we used the cross-validation fold value as 4. Further we analyzed the predicted Vs actual results report for this scenario as in Table 9. From Table 9 we see that the detection rate for anger, disgust and surprise has increased, whereas contempt and fear has been misclassified. Detection rate of sadness has been same as the above experiment. A logical assumption will be that the sadness emotion of the subjects, closely resemble that of anger. To improve the results, we did some fine tuning. We changed the data split to 70:30 and cross validation folds to 5. We use stratified sampling for the split. We also extracted morefeatures from the available 68 facial point landmarks.

Table 9: Report of predicted values Vs actual values without jaw features

Predicted / Actual	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise	All
Anger	**10**	1	0	0	0	0	0	**11**
Contempt	1	**2**	0	0	0	0	1	**4**
Disgust	2	0	**14**	0	0	0	1	**17**
Fear	0	0	1	**4**	0	0	0	**5**
Happy	1	1	0	0	**14**	0	0	**16**
Sadness	2	0	0	1	0	**8**	0	**11**
Surprise	0	0	1	0	0	1	**16**	**18**
All	**16**	**4**	**16**	**5**	**14**	**9**	**18**	**82**

Distances :

As we had the 68 co-ordinates we could easily calculate the distance between any two facial points on the face using the distance formula:

$$\sqrt{(x2 - x1)^2 + (y2 - y1)^2}$$

In Figure 18 we show some of the horizontal and vertical distances calculated for the face. In all,
we calculated 25 such distances as shown in Table 10. Paper [1] use displacement ratios
calculated using the facial landmark instead of directly using the distances.

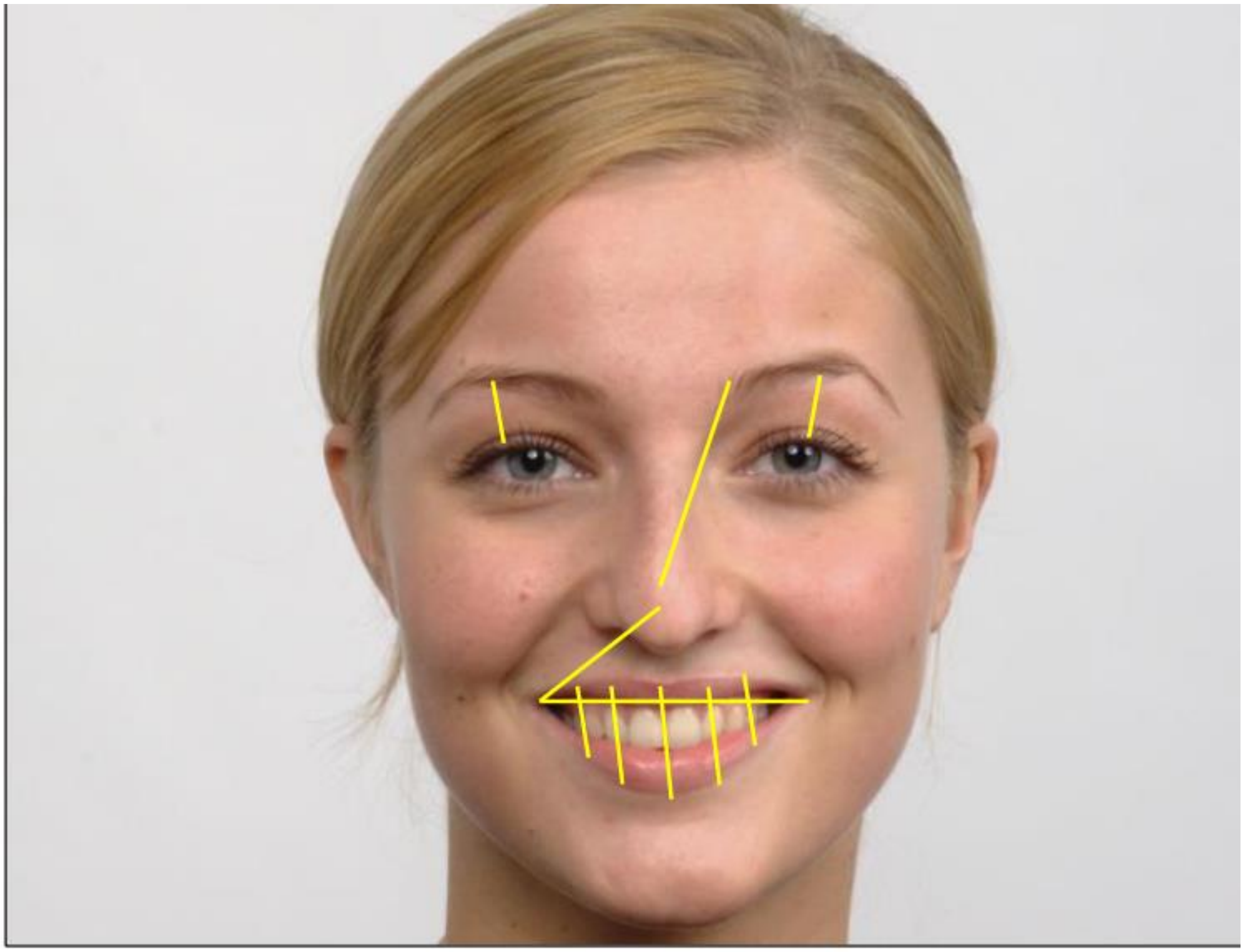


Figure 18: Distances calculated on the face

6.Conclusion and Future Scope:

Our implementation can roughly be divided into 3 parts:
1) Face detection
2) Feature extraction
3) Classification using machine learning algorithms

Feature extraction was very important part of the experiment. The added distance and area features provided good accuracy for CK+ database (89%). But for cross-database experiment we observed that raw features worked best with Logistic Regression for testing RaFD database and Mobile images dataset. The accuracy was 66% and 36% for both using CK+ dataset as training set. The additional features (distance and area) reduced the accuracy of the experiment for SVM as seen in Table 13 and Table 15.The algorithm generalized the results from the training set to the testing set better than SVM and other algorithms. The results of

the emotion detection algorithm gave average accuracy up to 86% for RaFD database and 87% for CK+ database for cross-validation=5. RaFD dataset had equal number of classes; hence, cross-validation did not help in improving the accuracy of the model.
Table 22 shows our performance as compared to different papers. When compared to Paper[1], which used ORB feature descriptors our method performed better, using only the 68 facial landmark points and the distances and area features [14]. Paper [1] achieved an accuracy of 69.9% without the neutral emotion whereas we achieved average accuracy of 89%. Paper [14] had similar feature extraction technique as ours and their accuracy was slightly(0.78) better than us. Paper [20] used large number of iterations to train the layers and achieved an accuracy of 98.15% with 35000 iterations. Paper [21] also used similar concept of angles and areas and achieved an accuracy of 82.2% and 86.7% for k-NN and CRF respectively.

We did not focus on face detection in this paper. Our main focus was on feature extraction and analysis of the machine algorithm on the dataset.But accurate face-detectionalgorithm becomes very important if there are multiple people in the image. If we are determining the emotion of a particular person from a webcam, the webcam should be able to detect all the faces accurately.

Future Work:

For future work, a more robust face detection algorithm coupled with some good features can be researched to improve the results. We focused on only some distances and areas, there can be many more such interesting features on the face which can be statistically calculated and used training the algorithm. Also, not all the features help to improve the accuracy, some maybe helpful with the other features. Feature selection and reduction technique can be on the created feature to improve the accuracy of the dataset. We can experiment with facial action coding system or feature descriptors as features or a combination of both of them. Also, we can experiment with different datasets amongst different races. This will give us an idea if the approach is similar for all kinds of faces or if some other features should be extracted to identify the emotion. Applications such as drowsiness detection amongst drivers [1]can be developed using feature selection and cascading different algorithms together. Algorithms like logistic regression, linear discriminant analysis and random forest classifier can be fine-tuned to achieve good accuracy and results. Also, metrics such as cross-validation score, recall and f1 score can be used to define the correctness of model and the model can be improved based on these metric results.

Refrences:

[1] W. Swinkels, L. Claesen, F. Xiao and H. Shen, "SVM point-based real-time emotion detection," 2017 IEEE Conference on Dependable and Secure Computing, Taipei, 2017.

[2] Neerja and E. Walia, "Face Recognition Using Improved Fast PCA Algorithm," 2008Congress on Image and Signal Processing, Sanya, Hainan, 2008

[3] H. Ebine, Y. Shiga, M. Ikeda and O. Nakamura, "The recognition of facial expressions with
automatic detection of the reference face," 2000 Canadian Conference on Electrical and Computer Engineering. Conference Proceedings. Navigating to a New Era (Cat. No.00TH8492), Halifax, NS, 2000, pp. 1091-1099 vol.2.

[4] A. C. Le Ngo, Y. H. Oh, R. C. W. Phan and J. See, "Eulerian emotion magnification for
subtle expression recognition," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016

[5] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH,2014

[6] M. Dahmane and J. Meunier, "Emotion recognition using dynamic grid-based HoGfeatures," Face and Gesture 2011, Santa Barbara, CA, 2011

[7] K. M. Rajesh and M. Naveenkumar, "A robust method for face recognition and face emotion detection system using support vector machines," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques