Human Emotion Detection using Machine learning

Avishek Gupta1, Sudeshna Das2 and Sohini Banerjee3

1,2,3 Assistant Professor, Department of Computer Science and Engineering, AIEM, Mogra

1 avishekgupta.aiem@jisgroup.org, 2 sudeshnadas645@gmail.com, 3 banerjeesohini88@gmail.com

Abstract

Human emotions are a common mental condition that can significantly affect both a people's daily life and mental health. Now a days, mental illness and human emotions are major issues. There are many emotions of facial expressions which incorporate natural, happy, sadness, disgust, fear, anger, surprise, and contempt. In this context we introduced monitoring system that is based on technology which recognizing emotions from our sample images. This topic evaluates the effectiveness of social media platforms' emotional catboats, messages, photos, and facial expressions in recognizing emotional cues.

Keywords: *machine learning, artificial intelligence, decision tree Regression, facial emotion recognition*

1. Introduction

Through voice, text analysis, gestures, facial expressions, and other means, emotions can be recognized. As an illustration, an AI-based driving application that can warn the driver if he falls asleep at the wheel could save a person's life. All of this is possible thanks to facial expression recognition technology, which uses a camera to photograph a person's face and determine whether or not they are asleep. Similar to this, using image and video processing, emotion detection systems may identify various moods such as happiness, sadness, neutral, etc. by various eye, lip, nose, and hand gestures. By analyzing the words and emoticons users exchange with chatbots, chatbots may also identify emotions. If a person is depressed, the system will automatically create a joke or start playing music to lift their spirits. Techniques from ML, AI, and data mining are being employed for this. The program for emotion recognition stores information about user responses from chatbots in a database. This can aid with stress reduction. Applications for emotion detecting systems are numerous. In market research, emotions are identified in order to understand client sentiment, which is crucial for organizations. Through consumer reviews of various products, emotion detection systems can be utilized to determine the customer's emotions. Employing emotion analytics in the hiring process can quickly locate potential employment prospects.

Based on the candidate's responses, AI algorithms analyze facial expressions, personality features, and emotions in a video interview, resulting in an objective interview process and simplifying the interviewer's task.

Machine learning methods very useful for categorization and pattern identification. In this context we will evaluate the extraction and modification of various features such as SVM (i.e.; Support Vector Machine. We will contrast feature extraction methods and algorithms from various studies. The human emotion dataset can serve as an excellent model for examining the nature and robustness of classification algorithms as well as their performance on various dataset types. Typically, face detection techniques are used to the image or video frame prior to feature extraction for emotion detection. The following is a generalization of the emotion detection steps:

1) Preprocessing the dataset
2) Recognizing faces
3) Extraction of features
4) Categorization according to characteristics

The primary focus of this work is on the technique of emotion detection and feature extraction using the extracted features.

2. Image Attributes

The image can yield a variety of features, which we can then normalize into vector form. We can measure the angles between the ellipses on the face or look at the different emotions on the face, eyes, and other features to determine the emotion. The following are some noteworthy qualities that can be used to instruct machine learning algorithms:

2.1 CHARTS

Human facial movements can be classified according to how they appear on the face using the Facial Action Coding System (FACS). It dissects facial expressions into discrete muscle movement components known as Action Units (AUs). In terms of action units, a smiling face, for instance, the movement of the Action Unit (AU) i.e.; AU6 ACTION: The cheek raiser squeezes the lateral eye corners and tightens the outer rings of the eye orbit. AU7 ACTION: The lower eyelid skin is pushed toward the inner eye corners by the lid tightener, which also tightens the rings surrounding the eyelids. It is possible to create real-time face models depending on human characteristics.

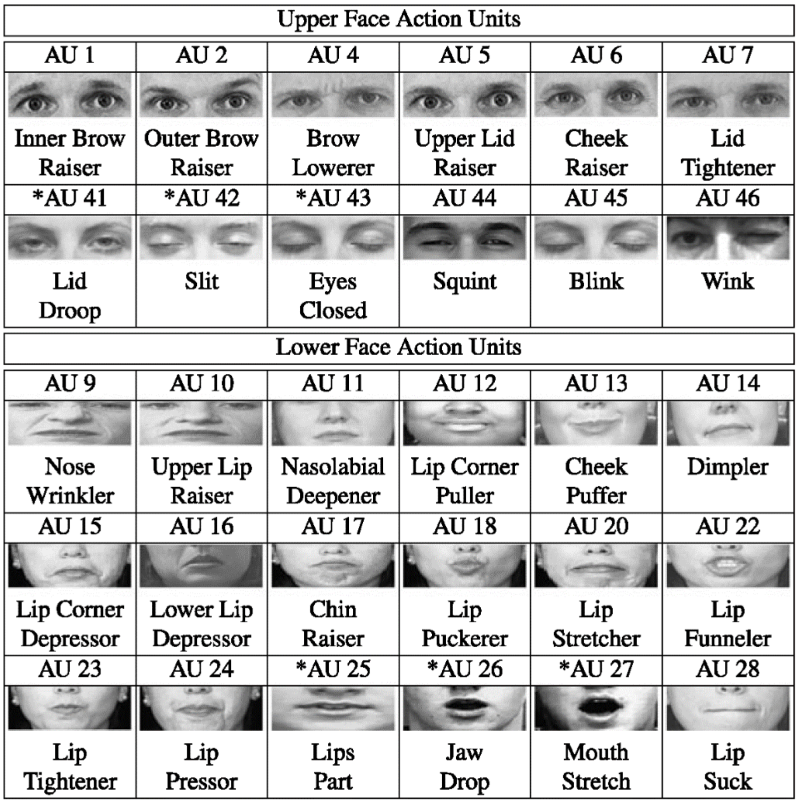

Figure 1: Action Units: corresponding to various facial expressions [3]

2.2 Critical Point

Critical points of Face are very important and can be used for facial recognition and detection. Here we present all total 68 facial critical point discoverer in the dlib.PyPI (a set of tools for developing machine learning and data analysis applications in the real world.)

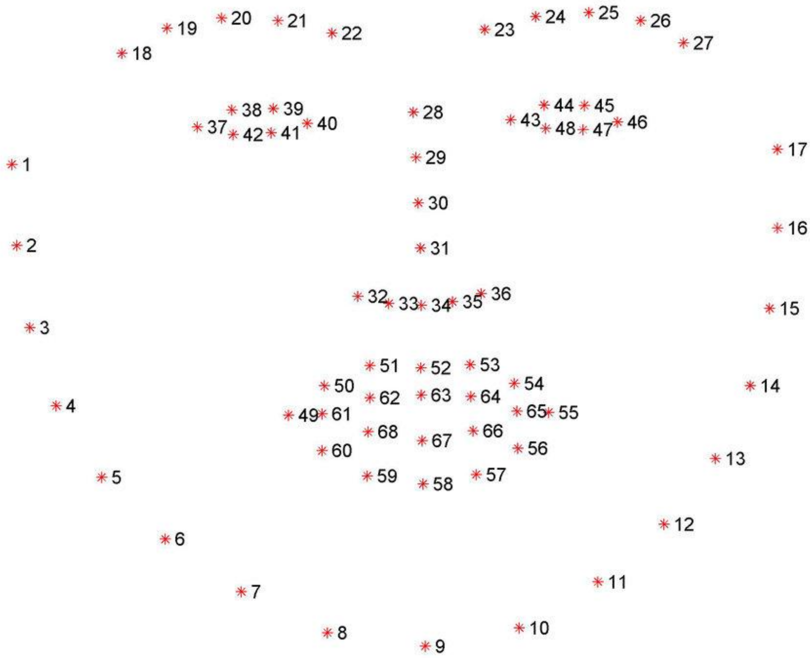

Figure 2: Face Critical point [3]

All these 68 crtical point on the face are depicted in Figure 2. The (x,y) coordinates of every facial point can be retrieved by using dlib.PyPI tools. All these points can be broken down into categories such as face, nose, eyebrow and jaw.

3. Connected Work

3.1Methods of Feature Extraction

3.1.1Decision tree Regression ensemble

This method uses images to identify the critical locations on the face using cascaded regression trees. 68 facial landmarks are identified by using pixel intensities to differentiate between various facial parts [1]. The shape parameter vectors are re-estimated using the extracted features, and this process is repeated until convergence is reached [5].

Figure 3: 19 Feature points in an image [1]

the author [1] uses 19 features Out of the 68. In Figure 3, focusing only on the region surrounding the eyes, eyebrows, and face.

Regression tree ensemble was incredibly quick and reliable, yielding 68 features in about 3 milliseconds. Table 1 [1] displays the algorithm's tuned parameters:

Parameter	Value
Waterfall level	14
Tree depth	3
Waterfall on every stage	490
Instruct multiple split tests	16

Table 1: Adjusting the unit of decision tree regression algorithm parameters [1]

3.1.2 Ratios of transition

Table 2 [1] displays the calculations for the twelve different types of distances.The 593 video sequences in the Extended Cohn-Kanade (CK+) dataset come from 123 distinct subjects, ages 18 to 50, representing a range of genders and ethnicities.

Range	Significance of the ranges
R1 and R2	The separation between the left and right eyes' upper and lower lids
R3	The separation between the inner corners of the right and left eyebrows
R4 and R5	The separation in between inner point of right and left eyebrows and also the nose
R6 and R8	The separation between right and left mouth corners
R7 and R9	The separation in between middle of upper and lower lips and also the nose
R10	The separation between the mouth corners on the right and left
R11	The separation in between lower and upper lips midpoints
R12	diameter of the face

Table 2: Determine the ratios of the displacement between several facial features [1]

3.2 Machine Learning Algorithms

Introducing a more flexible classification algorithm in the next stage after creating the dataset with the important features. One or more feature extraction techniques are combined with them [8].

3.2.1 SVM

SVM stands for Support vector machine, is one of the most popular supervised learning algorithms for both regression and classification issues. However, machine learning classification problems are the main application for it.

For this reason, the SVM algorithm looks for the best line or decision boundary that can split n-dimensional space into classes, making it easier to classify new data points in the future.

Generally, a multi-class SVM is used instead of a binary when detecting emotions like faer, happiness, surprise, and contempt [1].

Using K-fold cross-validation, one can eliminate Principal component analysis (PCA) was utilized by Loconsole et al. [8] to reduce the feature set before feeding the smaller feature set into SVM.

In this case, gamma aids in decision boundary optimization and C is the misclassification penalty function.

4. Implementation

4.1 Simplifying Facial Landmark Detection Using OpenCV

Different facial points such as the mouth, eyebrows, and area around the eyes. Rather than a multi-label problem, we have introduced multi-class classification problem. The multi-class emotions are detected using SVM in conjunction with the extracted facial features. The articles we've read center on SVM and well-defined algorithms for classifying emotions. In our database there are seven classes which are categorized. we have compared our results with those obtained from random forest and logistic regression. Face landmark detection done using Python, OpenCV, and dlib.

4.2 Setting up the database

Depending on the user and session number, the image files for the Cohn-Kanade Dataset (CK+) are located in various directories and subdirectories. File saved in (.jpg) image format. The image files have the same name as the emotion labels, but they are located in a separate directory. In order to pick up the emotion file name, we created a small utility function in Java.

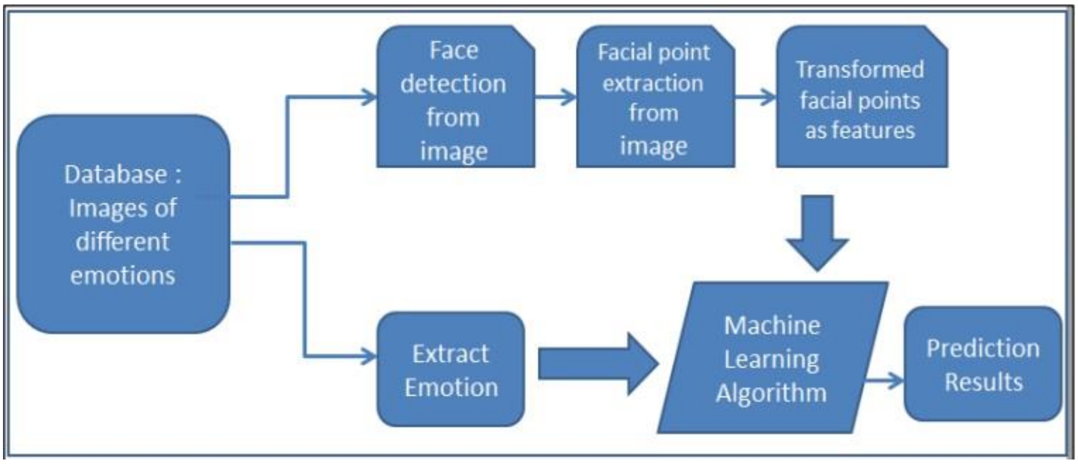

Figure 4: Implementation Pipeline

There was only frontal face images in the dataset we created; no files with no emotions, or neutral emotions, were included. There were colored images. Regardless of the lighting, the same processing pipeline was used for every image.
We simply used the name of the image file name with (.jpg) format, to determine the emotion name, and entered it into the RaFD database.

Feelings	Number of pictures that capture the feeling
1: Fury	43
2: Disdain	17
3: Disgust	57
4: Anxiety	23
5: Joy	67
6: Melancholy	26
7: Unexpected	81

Table 3: Count of pictures in each class for CK+ database

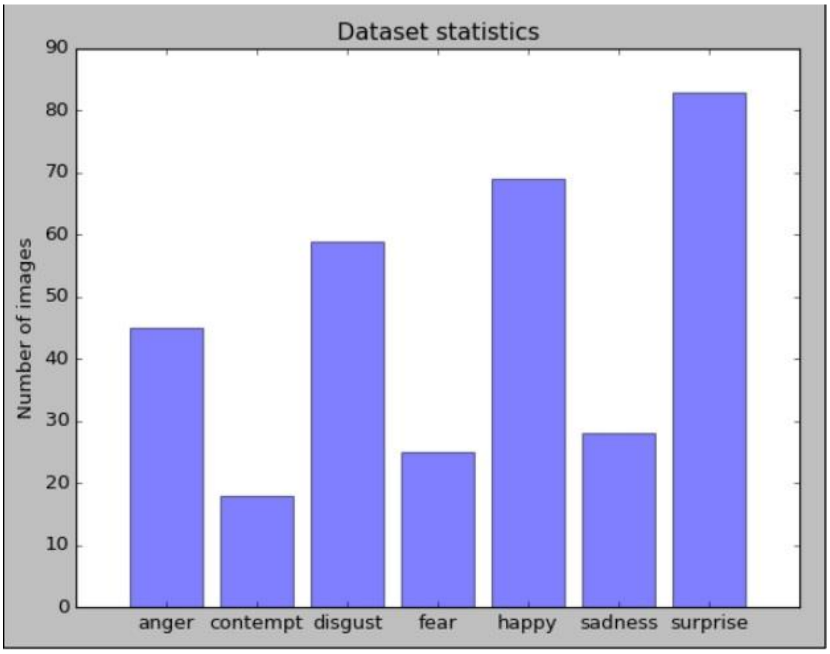

Figure 5: Bar graph showing the quantity of samples in each class

Figure 14 illustrates the unequal number of phases, Furthermore, the class may be regarded as an outlier because there are fewer training samples. The algorithm may start to favor the emotion of surprise and identify the majority of photos as such. There was no bias towards any particular class in the RaFD database.

4.3 Pipeline for image processing

4.3.1 Identifying faces

The most vital step in the processing pipeline was face detection. Detecting the face was necessary before any further processing could begin, with focus the facial expression data.

Figure 6: Original photo from the database and facial recognition software

Experimented with various face detection algorithms, including OpenCV's Haar-cascades. Ultimately, we decided to use the Dlib library's histogram of oriented gradients as the basis for our face detector. SVM and HoG descriptors are utilized to extract the face from the picture. To ensure uniformity, images are resized and converted to grayscale.

4.3.2Extraction of facial features

Edge detectors are frequently used to extract facial features like skin tone, eyebrows, eyes, nose, mouth, and hairline. A statistical model is constructed based on the extracted features in order to explain their relationships and confirm the presence of a face.

Figure 7: Recognizable landmarks on the face

We computed polygon areas and distances for a few of the facial landmarks as well.

4.3.3 The Python workflow

Each of the 327 files in the dataset was processed to produce the feature set after being placed in a directory. The emotion label was retrieved from the file name as soon as it was grabbed. Features prediction and face detection were applied to the image. The target classes situated in a numpy array. The RaFD database underwent the same procedure.

4.3.4 Machine Learning

In many situations, machine learning algorithms can be trained to accurately identify emotional expressions from photos of people's faces. On the other hand, implementation can be a challenging and intricate task. The state of technology is still in its infancy. Finding high-quality datasets can be challenging. The dataset was initially split into 30% for testing and 70% for training. Numerous other splits, including 80:20 and 70:30, were tried.

5. Results

Using our dataset, we trained support vector machines to forecast the outcomes. The accuracy metric and confusion matrix were applied to interpret the outcomes. The test ratio was 75:25. To get rid of any biases in the dataset, we also performed cross-validation. The split value of 4 was selected.
The following are the outcomes:

Support vector Machine kernel	Precision (%)	Cross-Validation Accuracy Score (cv=4)
In line	77.04	.77
Radial Basic Function	20.94	.24
Poly	76.51	.75

Table 4: Cross-validation and 75:25 split efficiency

The efficiency rating attained by the split was roughly matched by the mean cross-validation score. Our multi-class classification results is displayed in Figure 6.

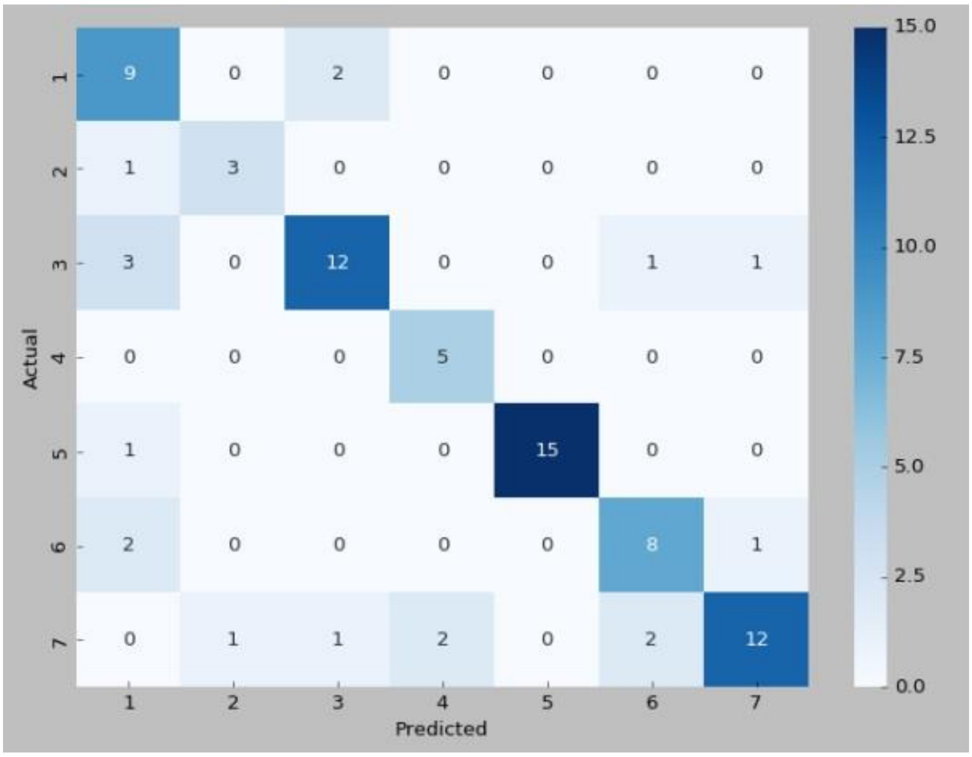

Figure 7: Actual and predicted values on a heat map

Predicted / Actual	Fury	Disdain	Disgust	Anxiety	Joy	Melancholy	Unexpected	All
Fury	**8**	0	3	0	0	0	0	**11**
Disdain	1	**3**	0	0	0	0	0	**4**
Disgust	4	0	**11**	0	0	1	1	**17**
Anxiety	0	0	0	**5**	0	0	0	**5**
Joy	2	0	0	0	**14**	0	0	**16**
Melancholy	2	0	0	0	0	**8**	1	**11**
Unexpected	0	1	2	2	0	2	**11**	**18**
All	**17**	**4**	**16**	**7**	**14**	**11**	**13**	**82**

Table 5: Comparison of anticipated values Vs reality values

The actual values and predicted values are displayed in report format in Table 5. This allows us to determine how many emotions are appropriate.
1: Fury – 8/11 – 72%
2: Disdain – 3/4 - 75%
3: Disgust – 11/17 - 65%
4: Anxiety – 5/5 - 100%
5: Joy – 14/16 - 88%
6: Melancholy – 8/11 - 73%
7: Unexpected – 11/18 – 61%
Numerous samples were incorrectly assigned to other classes, including fear, disgust, contempt, and sadness. For every sample, fear was accurately detected.

6. Conclusion

Three sections simply define our approach to execution:
1) Face recognition
2) Extraction of features
3) Utilizing machine learning algorithms for classification

An essential component of the experiment was feature extraction. For the CK+ database, the additional area and distance features offered good accuracy (89%). This was observed in a cross-database experiment. The accuracy was 36% and 66% for both when the training set was the CK+ dataset.

References

[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," 2008 Congress 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 HoG features," 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 (ICEECCOT), Mysuru, 2016

[8] C. Loconsole, C. R. Miranda, G. Augusto, A. Frisoli and V. Orvalho, "Real-time emotion recognition novel method for geometrical facial features extraction," 2014 International Conference on Computer Vision Theory and Applications (VISAPP), Lisbon, Portugal, 2014

[9] J. M. Saragih, S. Lucey and J. F. Cohn, "Real-time avatar animation from a single image," Face and Gesture 2011, Santa Barbara, CA, USA, 2011

[10] G. T. Kaya, "A Hybrid Model for Classification of Remote Sensing Images With Linear SVM and Support Vector Selection and Adaptation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 4, pp. 1988-1997, Aug. 2013