



## Emotion Analysis using Facial Expressions in Video

Kubra Ekmekci<sup>1\*</sup>, Serkan Ozbay<sup>2</sup>

<sup>1\*</sup> Gaziantep University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Gaziantep, Turkey, (ORCID: 0000-0001-5597-7749), [kubraekmekci1994@gmail.com](mailto:kubraekmekci1994@gmail.com)

<sup>2</sup> Gaziantep University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Gaziantep, Turkey, (ORCID: 0000-0001-5973-8243), [sozbay@gantep.edu.tr](mailto:sozbay@gantep.edu.tr)

(2nd International Conference on Access to Recent Advances in Engineering and Digitalization (ARACONF)-10–12 March 2021)

(DOI: 10.31590/ejosat.926478)

**ATIF/REFERENCE:** Ekmekci, K., Ozbay, S., (2021). Emotion Analysis using Facial Expressions in Video. *European Journal of Science and Technology*, (24), 523-527.

### Abstract

The topic of human computer interaction is one of the increasingly popular topics in recent times. Human facial expression and emotion analysis with the computer is one of the complex problems as well as interesting. In this paper, emotion analysis was made on human images. In the study, 5 different emotional states, being happy, angry, sad, surprised and neutral, were analyzed. The proposed algorithm basically consists of 3 steps. The first is the preprocessing of the images required for the SVM model. The second is the creation of the SVM model that could perform emotion analysis. The final step is to assign facial expressions to the relevant emotion class. In this study, JAFFE dataset and many images available from Google were used. The recognition success rates of 5 different emotions determined for the study were found between 80% and 100%.

**Keywords:** Emotion Analysis, Support Vector Machines (SVM), Facial Expressions.

## Videodaki Yüz İfadeleri Üzerinden Duygu Analizi

### Öz

İnsan ve bilgisayar etkileşimi konusu, son zamanlarda giderek daha popüler hale gelen konulardan biridir. Bilgisayar ile insan yüz ifadesi ve duygu analizi ilginç olduğu kadar karmaşık sorunlardan biridir. Bu çalışmada insan görüntüleri üzerinden duygu analizi yapılmıştır. Çalışmada mutlu, kızgın, üzgün, şaşırılmış ve nötr olmak üzere 5 farklı duygu durumu analiz edilmiştir. Önerilen algoritma temelde 3 adımdan oluşmaktadır. Birincisi, SVM modeli için gerekli görüntülerin ön işlemesidir. İkincisi, duygu analizi yapabilen SVM modelinin oluşturulmasıdır. Son adım, yüz ifadelerin ilgili duygu sınıfa yönlendirilmesidir. Bu çalışmada, JAFFE veri seti ve Google'da bulunan birçok görsel kullanılmıştır. Çalışma için belirlenen 5 farklı duygu için başarı oranları % 80 ile % 100 arasında değişmektedir.

**Anahtar Kelimeler:** Duygu Analizi, Destek Vektör Makineleri (SVM), Yüz İfadeleri.

\* Sorumlu Yazar: [kubraekmekci1994@gmail.com](mailto:kubraekmekci1994@gmail.com).

## 1. Introduction

Facial expressions are one of the non-verbal signs that play an important role in ensuring integrity of meaning in human relationships. While human beings do not have any difficulty in grasping facial expressions, this is not the case for machines. Therefore, developments on this subject continue. Human face detection and facial emotion analysis are based on digital image processing. These processes start with the capture of the image and continue with the use of different techniques for the specific purposes. These processes, which include mathematics and computer science, are used in areas such as design, manufacturing, security, medicine, electronics, machinery and architecture.

In facial expression analysis, images must be transformed into numerical data that can be processed by a computer. In order to make sense of digital data, digital data should be passed through the basic steps of image processing. Still images can be analyzed using image processing methods, and meaningful information can be extracted from these images. The fact that real-time image processing hardware is more accessible in terms of costs and the increase in processor capacities have enabled face recognition systems to be automated and integrated with other software networks [1]. Valsar M. made geometric feature extraction using the Support Vector Machine (SVM) and Hidden Markov Model (HMM) in his study in 2012. He reached a consistency of 50.4% in his study [2]. In this way, the emotions of the detected faces are separated in the coordinate system with geometric feature extraction. Gudi A., on the other hand, used the CNN (Convolutional Neural Network) architecture based on machine learning in 2015, and in his study on two separate data sets as SEMAINE and BP4D, he reached a consistency rate of 52.2% and 34.1% in emotional analysis over a hundred [3]. Pooya K reached a consistency of 50.5% by using CNN (Convolutional Neural Network) and RNN (Recurrent Neural Network) architectures together in his study in 2016 [4].

In the study conducted by Nguyen et al. in 2018, the model they proposed is a multi-level 18-layer CNN (Convolutional Neural Network) model and is similar to the VGG (Visual Geometry Group) model. This model also has a high level of features. The multi-level CNN (Convolutional Neural Network) model was proposed and reached 69.21% sensitivity. Fer2013 data set was used in this model to increase the rate they reached in sentiment analysis and achieved a sensitivity of 73.03% [5]. Another recent study is the study that Cao et al. conducted in 2019. This study created a model by combining K-Means, clustering and the Support Vector Machine (SVM) classifier. In this model, K-Means CNN determines the initial value of the model. SVM layers take features from the trained CNN (Convolutional Neural Network) model to classify images in the Fer2013 dataset [6].

In [7], facial expression detection applications were applied under two headings. The first was to analyze partial facial images with classical machine learning methods and the second was to analyze the whole face images. The proposed system was shown that emotion could be detected only by looking at the eyebrows and eye areas. Different datasets and methods were used. The success rate of 67.35% was achieved in the study using the JAFFE dataset for both testing and training. In [8], Tenekeci was aimed to get information about the psychological state of the faces on the image. 3 different psychological

situations were emphasized. Face detection was made using the OpenCV Haar class. Then, a conclusion was reached by using the features on the face picture. The application was tested on 15 pictures and the correct result was achieved in 11 pictures. In [9], facial expression recognition from pictures was performed in the study. Seven different emotions were tried to be recognized: happy, sad, surprised, fear, anger, disgusted and neutral. In the application, Viola-Jones face detector using AdaBoost algorithm was used to find the face location. Haar Filter was used to find the eyes and mouth. The recognition success rates of 7 different emotions determined for the study vary between 50% and 97%.

In this paper, emotion analysis based on facial expression in videos is studied. The proposed algorithm obtains information about the instantaneous psychological conditions of the people in the videos taken in real time through the images of the individuals. The paper is organized as follows. In Section 1, a brief introduction and then studies on facial expression recognition were mentioned. The steps required for the proposed method are described in Section 2. In Section 3, the databases in which the study was applied and the results of the study were mentioned. In the last section, there are suggestions for the development of the study, emphasizing the importance of the study.

## 2. Material and Method

As it is stated, the paper proposes an algorithm to capture information about the instantaneous psychological conditions of the people in the videos taken in real time through the images of the individuals.

For this, first of all, human faces in the videos are detected and the data are then pre-processed. Facial landmarks are detected as a result of the procedures performed on facial images. A support vector machine algorithm is created through these points, and finally, it is estimated to which emotion class (happy, sad, angry, surprise and normal) the emotion belongs to.

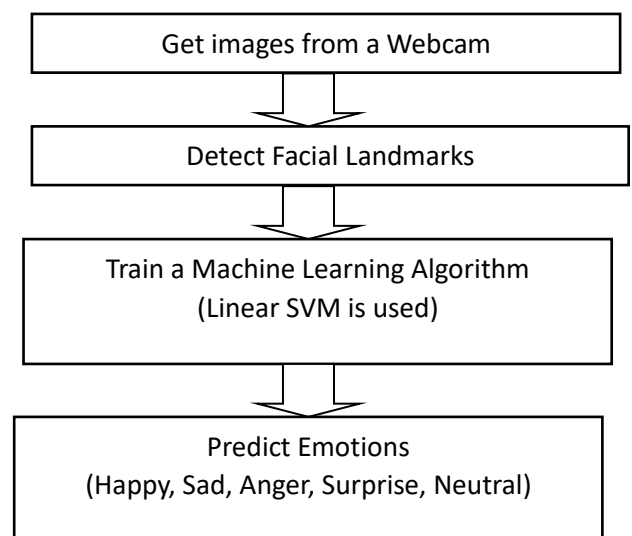


Figure 1. The system structure

The result of all experiments performed in this study was obtained by using computer with the following properties and software installed:

- Processor: Intel(R) Core(TM) i7-4720HQ.
- RAM 16GB.
- GPU 2.6GHz.
- 64-bit Operating System, X64 based processor.
- Python (2020.2.3) is an object oriented, interpretative, modular and interactive high level programming language. Its modular structure supports system and any data field entry. With Python, which can work on almost any platform, software can be developed in many areas such as system programming, user interface programming, network programming, application and database software programming.
- Anaconda is an open source, free and open source dispersion that points to encourage bundle administration in logical computing utilize of Python and R programming dialects.
- OpenCV (Open Computer Vision) is an open source library created by Intel, composed in C dialect that can run on Windows, Linux, MacOS X, PSP (Playstation Versatile) stages.
- Scikit-Learn (sklearn) could be a broadly utilized open source machine learning library. It incorporates direct relapse, calculated relapse, choice tree, arbitrary timberland and numerous fake learning strategies.
- Dlib library is an open source library created with C ++ programming dialect, which incorporates machine learning, profound learning and computer vision calculations.
- Visual Studio (2017) is an integrated development environment (IDE) developed by Microsoft.

## 2.1. Preprocessing

A video image taken at this stage was pre-processed. If the studies are to be listed:

- A video requested from the database as input was transferred to the Python environment.
- This entry was converted to gray scale.
- Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied on the images in order to increase the image quality, to eliminate the noise on the image and to make the human face detection more accurate in order for the application to give better results. Histogram Equalization was connected to the outlines to alter the differentiate and get on picture clearer than the initial one beneath the circumstances of different lighting conditions. However, Adaptive Histogram Equalization (AHE) included a propensity to overamplify clamar in moderately homogeneous districts of an picture. A variation of versatile histogram equalization called CLAHE anticipates this by constraining the enhancement.

## 2.2. SVM Model

At this stage, the points on the detected faces were determined first. Dlib library functions were used for this

process and the data was made available after preprocessing. A learning process was performed in order to detect facial landmarks on the images taken. 68 facial landmarks were detected on the detected faces, on the x - y coordinate plane. These detected facial landmarks were reflected on the screen on the face in the video.

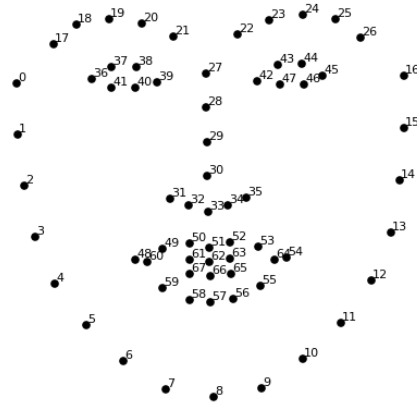


Figure II. Facial Landmarks [10]

- Jaw points: 0-16
- Right eyebrow points: 17-21
- Left eyebrow points: 22-26
- Nose points: 27-35
- Right eye points: 36-41
- Left eye points: 42-47
- Mouth points: 48-60
- Lip points: 61-67 [10]

Second, these points placed on the face were transformed into shapes to feed the classifier. Properties were small pieces of information that describe the object or object state that we were trying to categorize. Facial signs from the same image material described the position of all moving parts of the depicted face, which we used to express an emotion. The coordinates of the points on these moving parts were extracted. The positions of the points relative to each other were calculated on these coordinates, which means the process of calculating the point coordinates of the "center of gravity" information of all face marks. The positions of all points were calculated according to this central point obtained as a result of the calculation.

Here, the following steps were made:

- The center of gravity calculations were performed on the data.
- The distances of the points on the face with respect to this calculated center of gravity point were found.

These found positions, that was, the lines expressed both the size (i.e. the distance between two points) and a direction (angle

relative to the image with horizontal = 0). This means that these data obtained represent a vector.

As a result, the angle values made by the points with both distance and center of gravity were calculated.

The aim was to divide the existing data set into training and prediction sets with corresponding tags, consequently to train the classifier. (Support Vector Machines from SKLearn)

### 2.3. Emotion Recognition

The classification process, which was the last process in the SVM model, was carried out in two steps. First, 10 different data segmentation was made, training and test file was run, secondly, prediction probabilities were evaluated.

As a result of the study, Support Vector Machine learned to classify the model successfully. Because it was observed that the proposed algorithm performed emotion analysis through facial expressions in the real-time images provided to it.

## 3. Results and Discussion

In the approach, facial landmarks were first removed from the images. 80% of the data into a random training set and 20% into a test set were split, then they were trained in the training set to feed the classifier. Finally, in order to see how the model handles unknown data, the resulting model that was predict what happened in the test set was evaluated.

At the end of this evaluation, it was seen that different success rates were obtained on 5 emotional states (happy, sad, anger, surprised, neutral). JAFFE face database was used within the scope of the study.

JAFFE: Ten Japanese women modeled on JAFFE [11] dataset. There are 5 different emotional states of each person in these images. While these feelings were recorded as happy, sad, surprised, angry, neutral, care was taken to keep the hair away from the face and the ambient light to be the same. The database only contains front-facing face images of the respective model. It contains 150 images in total. These pictures are of equal numbers of each emotion. Table I shows the number of images and recognition accuracy with each emotion. There are 37 misrecognized facial emotion images; 4 images for happy, 3 images for sad, 17 images for neutral, 13 images for anger and any images for surprise.

Table 1. The number of images and recognition accuracy with each emotion of the JAFFE data set

| Emotions  | Number of Images | Accuracy(%) |
|-----------|------------------|-------------|
| Happy     | 30               | 95          |
| Anger     | 30               | 90          |
| Sad       | 30               | 84          |
| Surprised | 30               | 98          |
| Neutral   | 30               | 86          |

The proposed method was tested on approximately 2500 images found on Google. Different numbers of images were

collected for 5 different emotional states. Table I shows the number of images, recognition accuracy and precision with each emotion state.

The total number images of for happiness was 700. The number of correct recognized was 600 and the number of misrecognized was 100.

The total number images of for anger was 400. The number of correct recognized was 300 and the number of misrecognized was 100.

The total number images of for sadness was 400. The number of correct recognized was 290 and the number of misrecognized was 110.

The total number images of for surprise was 400. The number of correct recognized was 320 and the number of misrecognized was 80.

The total number images of for neutral was 600. The number of correct recognized was 450 and the number of misrecognized was 150.

Table 2. The number of images, recognition accuracy and precision with each emotion of tested images

| Emotions  | Number of Images | Accuracy (%) | Precision (%) |
|-----------|------------------|--------------|---------------|
| Happy     | 700              | 91           | 85            |
| Anger     | 400              | 89           | 75            |
| Sad       | 400              | 91           | 72            |
| Surprised | 400              | 91           | 75            |
| Neutral   | 600              | 92           | 80            |

A comparison with a previous work in which the JAFFE dataset was used utilized in Table III and Figure III. Note that, in the previous work, commit tests were performed for each level of concentrated for five feelings.

Table 3. The comparison with previous studies according to the SVM results the false acceptance rate (FAR) and false reject rate (FRR)

| Emotions  | Proposed System |       | Öznel (2018) [7] |       |
|-----------|-----------------|-------|------------------|-------|
|           | FAR             | FRR   | FAR              | FRR   |
| Happy     | 0,033           | 0,103 | 0,009            | 0,074 |
| Anger     | 0,099           | 0,105 | 0,056            | 0,037 |
| Sad       | 0,029           | 0,437 | 0                | 0,111 |
| Surprised | 0,016           | 0     | 0,009            | 0,037 |
| Neutral   | 0,127           | 0,235 | 0,019            | 0,111 |

As shown below, although the proposed system has shortcoming, it is more successful for some emotions.

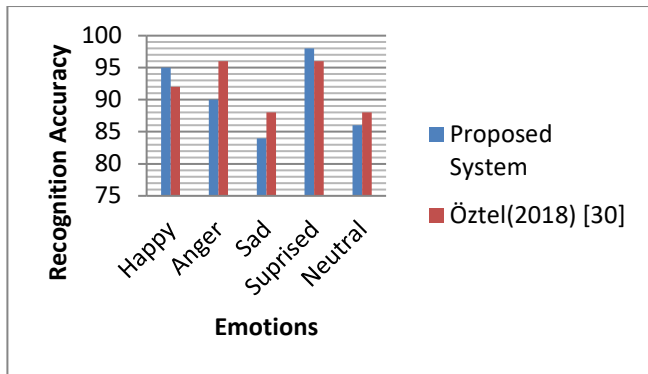


Figure III. Comparison with previous studies for each emotion

#### 4. Conclusions and Recommendations

As a summary, in this study, emotion analysis was made over real time videos. The emotional states of the people in the videos were detected instantly. The advantage of the SVM classifier used during the run is its probability. So this means that probabilities are assigned to each category in which it is trained. It returns the emotion class with the most votes as a result. If we look at where the model has flaws, we get a single answer as a result, but a single image may be happy with 90% probability and angry with 10% probability. Different methods can be developed so that the classifier can return such things and achieve higher success rates.

The reason why emotion analysis is a difficult field is that many expressions of emotions can be found simultaneously and similarly on the face. The success rates obtained in the current and limited data set are higher than the average values examined in the literature. In order to increase these rates, more data is needed as in every machine learning study. However, there are many other constraints. However, due to the emotional complexity and rapid changes in mood, it is almost impossible to completely separate certain emotions.

#### References

- [1] Romero, M. , Pears N. (2009). Landmark localisation in 3D face data. 6th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2009, 73-78.
- [2] Valstar, M.F. & Pantic, M. (2012). Fully automatic recognition of the temporal phases of facial actions. IEEE Trans. Syst. Man Cybern. Part B Cybern. 42(1), 28-43.
- [3] Gudi, A. , Tasli, H.E. , Den Uyl, T. M. & Maroulis A. (2015). Deep learning based FACS action unit occurrence and intensity estimation.

- [4] Khorromi, P.K. , Paine, T.L. , Brady, K. , Dagli, C. & Huang, T.S. (2016). How deep neural networks can improve emotion recognition on video data. 1-5.
- [5] Nguyen's Yearn, HD., Lee, G-S. , Yang, H-J. , Na, L. & Kim, H. (2018). Facial emotion recognition using an ensemble of multilevel convolutional neural networks. International Journal of Pattern Recognition and Artificial Intelligence.
- [6] Cao, T. & Li, M. (2019). Facial expression recognition algorithm based on the combination of CNN and K-Means. 11th International Conference on Machine Learning and Computing.
- [7] Oztel, I. (2018). Facial expression detection with machine learning methods on partial and full face images. PhD Thesis, Sakarya University.
- [8] Tenekeci, M. E. , Gumuscu, A., Baytak, A., Aslan, E. (2014). Emotion analysis from image with OpenCV. Academic Informatics'14 - XVI. Academic Informatics Conference Proceedings.
- [9] Ozmen, G. (2012). Facial expression recognition with cubic bezier curves. Master's Thesis, Trakya University.
- [10] Martinez, J. C. "<https://livecodestream.dev/post/detecting-face-features-with-python/> 21.02.2021".
- [11] Lyons. M. , Akamatsu, S. , Kamachi, M. & Gyoba, J. (1998). Coding facial expressions with Gabor wavelets. Third IEEE International Conference on Automatic Face and Gesture Recognition, 200-205.