FACIAL EMOTION DETECTION

Prepared For Image Processing (CSE4019) J COMPONENT
Winter Semester 20-21

Submitted To
Prof. Swathi J N
School Of Computer Science and Engineering

Submitted by
ZATAKIYA DEEP 18BCE0090
SOUMYA VISHWAKARMA 18BCB0086

Acknowledgment

This is to articulate my deepest gratitude and joy at being bestowed upon with an opportunity to get a project on “Facial Emotion Detection”.

We would like to thank all the authors of review reports who have previously worked on related research papers. We would avail this opportunity to express our profound gratitude and indebtedness to Prof. Swathi J N for giving us the opportunity to explore and learn about Image Processing and Automation under her guidance.

Lastly, we would also like to thank the sources studying through which we could learn the insights of dealing with image processing.

Table of Contents

List of Figures

List of Tables (add lit review)

1. Introduction

1.1 Objective

Realtime detection of the face and interpreting different facial expressions like happy, sad, angry, afraid, surprise, disgust, and neutral using image processing.

1.2 Motivation

Humans often have different moods and facial expressions changes accordingly. Human emotion recognition plays a very important role in social relations. The automatic recognition of emotions has been an active analysis topic from early eras. In this deep system user's emotions using its facial expression will be detected.

An automatic facial expression recognition system has to perform detection and site of faces during a cluttered scene, facial feature extraction, and facial expression classification. This system has ability to monitor people emotions, to discriminate between emotions and label them appropriately and use that emotion information to guide thinking and behavior of particular person.

1.3 Background

The study of facial expressions comes back to Darwin's research on evolution of the species which it appeared as a shape of nonverbal communication. In his studies, the facial behavior was categorized into several groups. This type of communication is faster than verbal and it brought more advantages to the human species than others.

Facial emotion stated as person's internal state, intentions and its feeling response to external stimulus. Emotions are an important property of humans and are essential for effective interactions among the society. Human communication can be either verbal of nonverbal, which it has been shown most of them refer to nonverbal communication.

In nonverbal communication, emotion plays an effective role because it conveys humans' feelings about the subject, and in psychology research it is proven that facial expressions are more effective than spoken word in conversation.

Emotion recognition has intersection of several areas of computer science, cognitive science and psychology, and it can be carried out by several methods such as body language, voice intonation and electroencephalography (EEG).

The easier and practical way is recognizing the emotion from facial expression. So, in the interaction environment, facial emotion recognition is more practical than recognizing emotion from EEG signal because EEG is suit for clinical application such as neurofeedback where the subjects are fixed.

There were shown on several psychological studies which culture and environment can influence the impact of emotion and the way of expressing feeling for human beings. In many of these studies shown that gender, cultural background, age have bias in expressing emotion while there is not clear evidence on importance of environment for tendency the emotion.

Sl no	***Title***	***Authors***	***Year***	***Dataset***	***Methodology Proposed***	***Pros***	***Cons***	***Future Work***
1	***Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition***	Diah Anggraeni, Ajeng Wulandari, T.Basaruddin, Dewi Yanti Liliana	2017	The experiments were performed using posed dataset JAFFE, CK+, and MUG.	There are two approaches in facial expression analysis, i.e. feature extraction and action unit detection from Facial action coding system. Feature extraction has two main techniques: geometric feature-based and appearance-based feature extraction Deep learning is a part of machine learning approaches that can be utilized as emotion recognition and facial expression analysis. However, its performance depends on the data size.	Performance of CNN increased and gave significant results in face detection and cropping phase. CNN can directly learn from region of interest because all background has been remove during pre-processing phase. Adding noise as augmentation techniques showed quite good impact although not as good as cropping phase.	The chance of the data to be overlap between training set and testing set increased since there are 2 version of the data: original and noisy data.	Involves exploring image synthesis techniques that may be considered as solution of augmentation data in deep learning. It aims to prevent data starvation and overfitting for small amount of data.
2	***Emotion recognition from facial expression using deep convolutional neural network***	D Y Liliana	2019	The dataset used in this research is the Extended Cohn Kanade database (CK + database). CK+ consists of 10.708 images from 123 different subjects.	The architecture of proposed CNN is depicted on Fig. 2. It has two convolutional layers, and two subsampling layers. The first convolutional layer used six masks, or so called c1 layer. The next layer is subsampling layer which has two layers (s1). The second convolutional layer (or c2) has 12 masks. The last subsampling neural network has two layers. The last is fully connected layer which resulting in the class classification.	the mean square error declines as the training data grows. the performance of the system reaches 92.81% of the accuracy rate	Each class has a misclassification results which indicates that the system needs further improvement.	concern on the design of CNN architect to gain the better result.
3	***Fuzzy Emotion Recognition Using Semantic Facial Features and Knowledge-based Fuzzy***	Dewi Yanti Liliana, T. Basaruddin, M. Rahmat Widyanto	2019	Indonesian Mixed Emotion Dataset (IMED) is used to measure the recognition performance of the proposed system since no dataset for mixed	Semantic facial feature is more meaningful than low-level feature (e.g. shape or texture) or the ordinary facial component features (e.g. eyes or mouth), since each facial component may have different condition. Whereas semantic facial feature describes the condition of facial component in a linguistic that is understandable by human. As an example, eye has three condition: narrow, normal, and wide. Hence, semantic facial feature is	This paper implements fuzzy emotion recognition as a natural human emotion which tolerates the ambiguity and multiple emotion occurrence in facial expression. Fuzzy emotion is a generalization of mixed emotion which is a real human emotion displayed in a daily life interaction.	Fuzzy emotion rules must be explicitly given 1. IF eyebrow is lower AND eye is normal AND mouth is widely open THEN angry is high 2. IF eyebrow is raise AND nose is wrinkle AND mouth is narrow THEN disgust is medium 3. IF eyebrow is normal AND	the study considers optimizing the parameter of fuzzy emotion inference system to increase the recognition results.

				emotion exists for public.	better representing the facial component's state (e.g. eye is wide). Facial points detection is a preliminary step prior to the semantic facial features extraction.		eye is narrow AND mouth is narrow THEN fear is low	
4	***Feature Selection on 2D and 3D Geometric Features to Improve Facial Expression Recognition***	Vianney Perez-Gomez, Homero V. Rios-Figueroa, Ericka Janet Rechy-Ramirez, Efren Mezura-Montes and Antonio Marin-Hernandez	2020	Consider the databases where a single face appears in each image and that3D landmarks are identified. We extracted the information from two different datasets: (a) Bosphorus and (b) UIBVFED.	4 stages: (i) data acquisition, (ii) feature extraction, (iii) feature selection, and (iv) classification.In the data acquisition stage, instances of facial expressions are collected from the Bosphorus orUIBVFED database. Then, feature extraction is carried out and, from 3D landmarks on a human face,89 geometric features are determined. In the third stage, relevant features are selected from the originalfeature set through two methods: PCA and a GA.Finally, a support vector machine with cubic kernel is applied for classification to the originalfeature set and the reduced feature sets to classify the six basic facial expressions. The expressionare (i) anger, (ii) disgust, (iii) fear, (iv) happiness, (v) sadness, and (vi) surprise	"Happiness" reported highest accuracy This proposal might be employed to improve the computational cognitive mechanism used to inferemotions in a static image; therefore, a computer might adapt its interaction with people based on theemotion detected.	Large scale dataset & massive computing power required, making it ill-suited for mobile platforms with limited resources.	Investigate the value of the proposed descriptors on moredatasets and possibly explore micro-expressions.
5	***Facial Emotion Recognition using Convolutional Neural Networks***	Akash Saravanan Gurudutt Perichetla Dr. K.S.Gayathri	2019	The dataset used was the FER-2013. It included 35,887 images with the expression such as angry, disgusted, fearful, happy, sad, surprised and neutral.	The network consists of six two-dimensional convolutional layers, two max pooling layers and two fully connected layers. Max pooling uses the maximum value from each of a cluster of neurons at the prior layer. This reduces the dimensionality of the output array. The input to the network is a preprocessed face of 48 x 48 pixels. The model was developed based on the observation of the performance of the previous models. It was decided to go with a deeper network over a wide one.	The advantage of using the model is that it prevents memorization	However, there was a bias toward the happy expression in the dataset. It was converted to a training set of 28709 images.	An indepth analysis of the top2 predicted emotions may lead to a much more accurate and reliable system. Further training samples for the more difficult to predict emotion of disgust will definitely be required in order to perfect such a system

6	***Affective state detection via facial expression analysis within a human – computer interaction context***	Anas Samara,Leo Galway,Raymond Bond & Hui Wang	2017	CK-8 and KDEF	Geometric-based techniques for facial expression analysis are based on locating the facial points and determining the location and the shape of associated facial components, including the eyebrows, eyes, nose, lips and mouth	a 95% confidence interval of classification results	Albeit a great deal of consideration was paid towards conveying outward appearance investigation in Affective Computing, the difficulties experienced in these undertakings may not exclusively be as far as specialized issues.	work is progressing in different directions within various disciplines to reinforce the link between the human and the computer.
7	***Facial expression analysis with AFFDEX and FACET: A validation study***	Sabrina Stöckli, Michael Schulte, Mecklenbeck, Stefan Borer & Andrea C. Samson	2017	For this three datasets were used for comparative analysis- WSEFEP, ADFES and RaFD.	The objective of this study is to validate the emotion classifying algorithms, namely- AFFDEX and FACET. A matching score was calculated based on the accuracy of the facial expressions that were realized by studying the basic emotions. It specified the percentage of the pictures that were classified correctly. Higher MS, higher the accuracy. The AFFDEX score was 73% across the three databases and FACET recognized 99% of the emotions in ADFES. While AFFDEX failed to detect 1% of the pictures, FACET had no failures.	FACET had no failures.	AFFDEX showed poor accuracy for fear an anger. It is fairly difficult to find a dataset which has equal number of images for each of the emotions. Therefore, it is a problem to achieve high accuracy for other emotions.	
8	***A Brief Review of Facial Emotion Recognition Based on Visual Information***	Byoung Chul Ko.	2018	Deep-Learning Based FER Approaches.	As we described, such approaches can be divided into two main streams: conventional FER approaches consisting of three steps, namely, face and facial component detection, feature extraction, and expression classification. The classification algorithms used in conventional FER include SVM, Adaboost, and random forest; by	Outperform conventional approaches with an average of 72.65% to 63.2%.	Large scale dataset & massive computing power required, making it ill-suited for mobile platforms with limited resources.	If emotional oriented deep-learning algorithms can be developed and combined with additional Internet-of-Things sensors in the future, it is expected that FER can improve its current recognition rate,

					contrast, deep-learning-based FER approaches highly reduce the dependence on face-physics-based models and other pre-processing techniques by enabling "end-to-end" learning in the pipeline directly from the input images.			including even spontaneous micro-expressions, to the same level as human beings.
9	***Emotion Recognition from Facial Expression using Deep Learning***	Nithya Roopya. S	2019	Kaggle's Facial Expression Recognition Challenge and Karolinska Directed Emotional Faces (KDEF) datasets were used, implemented using tensor-flow.	Inception Net v3 Model is used to train and build an emotion recognition model which can be used in any applications. Inception is evolved from GoogLeNet Architecture with some enhancements and advancements. Inception model is used for automatic image classification and image labelling. Inception-v3 is used for image classification in Google Image Search. Inception-v3 achieved top 5.6% error rate in ILSVRC 2012 classification challenge validation.	**1**) This paper presents application for the emotion in network teaching system. 2)Wearing glasses on the face area has no effect on emotion recognition.	1) Distance between the camera and face will have an impact on an area of face recognition. 2) Regional impact of the human face effect the performance of emotion recognition like Hear, Sitting postures, Light strength.	In future, real time emotion recognition can be developed using the same architecture
10	***Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality***	Dhwani Mehta, Mohammad Faridul Haque Siddiqui, and Ahmad Y. Javaid	2019	CK, JAFFE	Mixed Reality Device Microsoft Holo-lens is used for observing emotion recognition in Augmented Reality(AR)	Wide range of Output is produced thus removing the imitations of older versions	Requires complex De-bugging and Highly advance sensors. So it may not be suitable for everyday use.	Mixed Reality is the future. The device Microsoft HoloLens available can be used to improve the accuracy of the current experiment by exploiting the use of its high-quality sensors, and more robust testing can be done using an extensive database for real-time emotion recognition. Expression change should be detected as soon as a person's expression changes with time.

Table 3.1: Related Works

2. Project Description & Goals

Emotion recognition methods can be divided into two main groups:

- First group work on static images and second one work on dynamic image sequences.
  In the static approaches, temporal information is not considered and they just use current image information, while in the dynamic approaches images temporal information used in order to recognize expressed emotion in frame sequences.

- Automatic emotion expression recognition includes three steps: face image acquisition, feature extraction, and facial emotion expression recognition.
  In the optimal extracted features, with in class variations of expression should be minimum while between-class variations should be maximum. If the extracted features are not suite for task in hand and do not have enough information, even the best classifier may be unsuccessful to have best performance.

Feature extraction for emotion recognition can be divided into two approaches: Geometric feature-based methods and appearance-based methods.

In the first methods, location and shape of parts of the face such as eyes, mouth, eyebrows and nose are considered, while in the second methods, particular regions or whole of face are considered.

3. Technical Specification

This project runs on top of a Convolutional Neural Network (CNN) that is built with the help of Keras whose backend is TensorFlow in Python. We have taken spider environment to run the code.

In this project, we used a dataset provided by Kaggle website, which consists of about 37,000 wellstructured 48 × 48 pixel grayscale images of faces. The images are processed in such a way that the faces are almost centered and each face occupies about the same amount of space in each image. Each image has to be categorized into one of the seven classes that express different facial emotions.

These facial emotions have been categorized as: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, and 6=Neutral. Figure 1 depicts one example for each facial expression category. In addition to the image class number (a number between 0 and 6), the given images are divided into three different sets which are training, validation, and test sets.

4. Design Approach And Details

4.1 Materials & Methods

Our project code setup consists of the following steps:

- Detection of face through webcam. We have used OpenCV for image processing tasks where we identify a face from a live webcam feed.

- Eliminating the background for better performance.

- Detection the facial coordinates and happy it in the datasets.

- The facial emotions that can be detected and classified by this system are Happy, Sad, Anger, Surprise and Neutral.

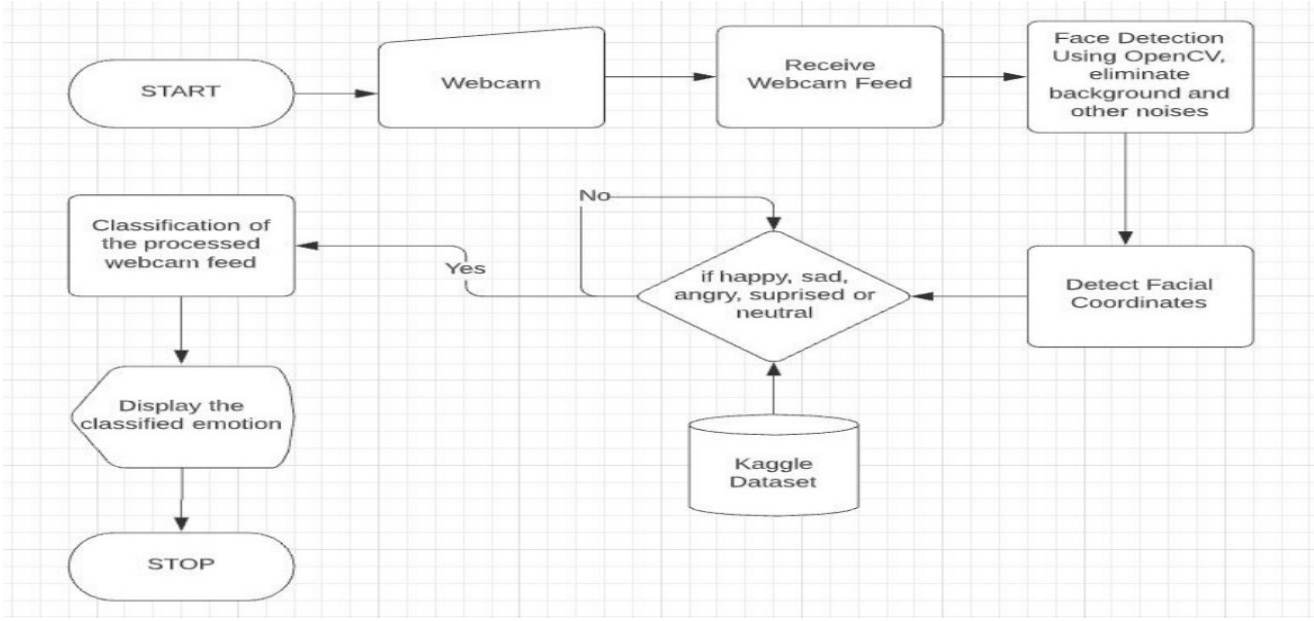

Figure 4.1: Model Employed for Facial Emotion Detection

4.2 Code

The following is the GitHub Code for the project:
https://github.com/DeepZatakiya/FacialEmotionDetection

4.3 Constraints, Alternatives & Tradeoffs

The accuracy of the facial detection system will grossly depend on the number of images in the training dataset. As explained before, we have chosen Kaggle Dataset consisting of 28,709 training images. Therefore, we expect a fairly accurate result.

The resolution of the camera used to extract the video of the user is also essential for the accuracy. Our project uses a webcam. Therefore, we have a pre-processing stage where the background noises will we removed to improve accuracy.

Even accurate facial emotion recognition may not give us an insight on the person's thought as sometimes they can be feigned

5. Project Demonstration

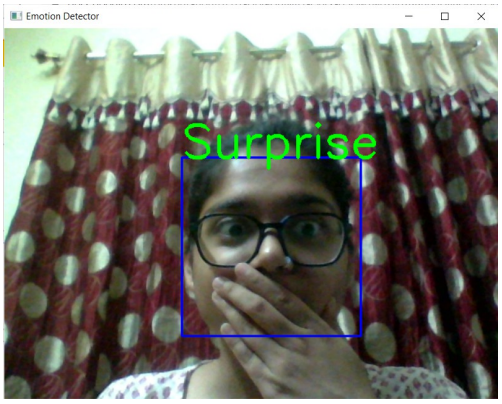

Figure 5.1 Facial Emotion Detected – Surprise

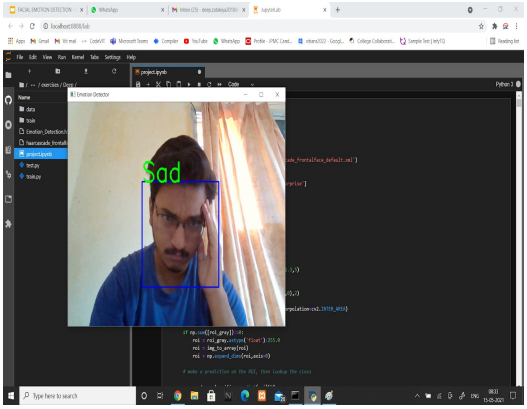

Figure 5.2 Facial Emotion Detected – Sad

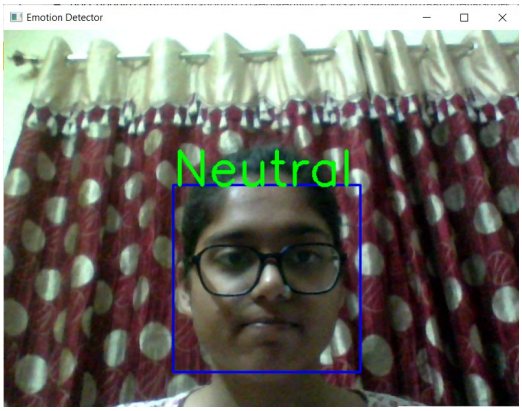

Figure 5.3 Facial Emotion Detected – Neutral

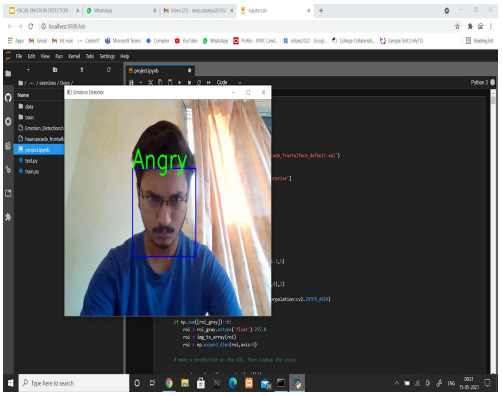

Figure 5.4 Facial Emotion Detected – Angry

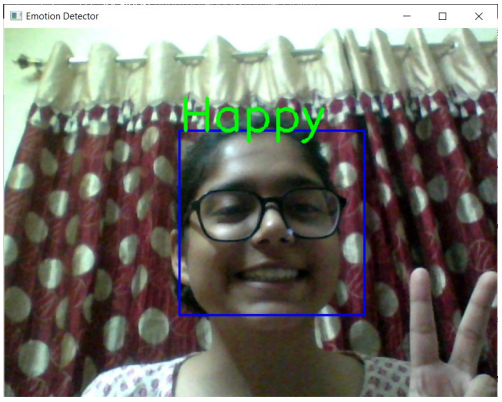

Figure 5.5 Facial Emotion Detected – Happy

6. Result

From the above demonstration presented, we have seen how the facial emotion is detected. To see how well our model has fared, we took into consideration the accuracy. In our project, we measured 3 accuracy methods. Accuracy is measured as correctly predicted class out of all the testing cases. This one is called the standard accuracy formula. It uses false positives, true positives, false negatives and true negatives. The second method we have used is called the prevalence accuracy which is another diagnostic test which uses sensitivity and specificity. The third method is called percent error which takes into consideration the observed values and the values that are accepted as truth. It informs us about the accuracy of the reading and how much the observed values derives from the truth.

The following images are the results observed for each of the images.

Standard method #1	
True positive	769
False negative	37
False positive	24
True negative	170
Accuracy	93.9 %

Figure 6.1: Accuracy, Standard Method

Prevalence method #2	
Prevalence	93.9 %
Sensitivity	95.4 %
Specificity	87.63 %
Accuracy	94.93 %

Figure 6.2: Accuracy, Prevalence Method

Percent error method #3	
Observed value	1,000
Accepted value	939
Percent error	6.496 %

Figure 6.3: Accuracy, Percent Error Method

Over all, our conclusion is that we have built a considerably accurate model based on the percentages that are all 90+ for the first two methods. For the percentage error method, the smaller the percentage derived, the better the accuracy. Therefore, even by this measure our model is fairly accurate.

In this project user's emotions using its facial expression will be detected. Real-time detection of the face and interpreting different facial expressions like happy, sad, angry, afraid, surprise, disgust, and neutral etc.

An automatic facial expression Recognition system will perform detection and site of faces during a cluttered scene, facial feature extraction, and facial expression classification. It provides a proposed model to solve the problems of emotion recognition based on facial recognition in virtual learning environments, and the efficiency and accuracy are considered at the same time.

This system has ability to monitor people emotions, to discriminate between emotions and label them appropriately and use that emotion information to guide thinking and behaviour of particular person

7. **Reference**s

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5856132/
- https://www.researchgate.net/publication/227031714_Facial_ Expression_Recognition
- http://cs231n.stanford.edu/reports/2016/pdfs/023_Report.pdf
- https://www.ijert.org/research/effective-emotion-recognitionof-expressions-from-facial-features-IJERTCONV5IS06002.pdf
- https://www.mdpi.com/1424-8220/18/2/40
- https://www.ijeat.org/wp-content/uploads/papers/v8i6S/F10190886S19.pdf
- https://www.irjet.net/archives/V3/i2/IRJET-V3I284.pdf
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5856132/
- https://arxiv.org/pdf/1910.05602.pdf
- https://link.springer.com/article/10.1007/s12652-017-0636-8
- https://link.springer.com/article/10.3758/s13428-017-0996-1
- https://www.researchgate.net/publication/332518418_Emotion_recognition_from_facial_expression_using_deep_convolutional_neural_network
- https://www.researchgate.net/publication/320402070_Enhancing_CNN_with_Preprocessing_Stage_in_Automatic_Emotion_Recognition
- https://www.researchgate.net/publication/332783468_Fuzzy_Emotion_Recognition_Using_Semantic_Facial_Features_and_Knowledge-based_Fuzzy
- https://www.researchgate.net/publication/332783468_Fuzzy_Emotion_Recognition_Using_Semantic_Facial_Features_and_Knowledge-based_Fuzzy
- https://www.researchgate.net/publication/343921384_Feature_Selection_on_2D_and_3D_Geometric_Features_to_Improve_Facial_Expression_Recognition
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5856132/
- https://www.researchgate.net/publication/227031714_Facial_ Expression_Recognition
- http://cs231n.stanford.edu/reports/2016/pdfs/023_Report.pdf
- https://www.ijert.org/research/effective-emotion-recognitionof-expressions-from-facial-featuresIJERTCONV5IS06002.pdf
- https://www.mdpi.com/1424-8220/18/2/40