Image-Based Facial Emotion Recognition

Rishabh Aggarwal
2017eeb1161
Sakib Shaukat
2017eeb1163

IIT Ropar
Punjab, India

Abstract

Facial Expression conveys non-verbal cues, which plays an important role in interpersonal relations. The Facial Expression Recognition system is the process of identifying the emotional state of a person. In this system, the captured image is compared with the trained data set available in the database and then the emotional state of the image will be displayed. This system is based on image processing and machine learning. In this project, we investigate ways to leverage the representational power of convolutional neural networks to distinguish between seven emotions from pictures of facial expressions.

1. Introduction

Emotions often mediate and facilitate interactions among human beings. Thus, understanding emotion often brings context to seemingly bizarre and/or complex social communication. Emotion can be recognized through a variety of means such as voice intonation, body language. However, the easier, more practical method is to examine facial expressions. There are seven types of human emotions shown to be universally recognized across different cultures: anger, disgust, fear, happiness, sadness, surprise, contempt. Therefore a utility that detects emotion from facial expressions would be widely applicable. The task of emotion recognition is particularly difficult for two reasons:
1) There does not exist a large database of training images and
2) classifying emotion can be difficult depending on whether the input image is static or a transition frame into a facial expression.
The latter issue is particularly difficult for real-time detection where facial expressions vary dynamically. The system classifies the facial expression of the same person into the basic emotions namely anger, disgust, fear, happiness, sadness and surprise. The main purpose of this system is efficient interaction between human beings and machines using eye gaze, facial expressions, cognitive modeling, etc. Here, detection and classification of facial expressions can be used as a natural way for the interaction between man and machine. And the system intensity varies from person to person and also varies along with age, gender, size, and shape of the face, and further, even the expressions of the same person do not remain constant with time.

1.1. Project Statement

Human emotions and intentions are expressed through facial expressions and deriving an efficient and effective feature is the fundamental component of facial expression systems.

Face recognition is important for the interpretation of facial expressions in applications such as intelligent, man-machine interface and communication, intelligent visual surveillance. An automatic Facial Expression Recognition system needs to solve the following problems: detection and location of faces in a cluttered scene, facial feature extraction, and facial expression classification.

1.2. Objectives

1. To develop a facial expression recognition system.
2. To experiment with machine learning algorithms.
3. To detect emotion.

1.3. Scope and Applications

The scope of this system is to tackle the problems that can arise in day to day life. Some of the scopes are:

1. The system can be used to detect and track a user's state of mind.
2. This system can be used for lie detection amongst criminal suspects during interrogation
3. The system can be used in mini-marts, shopping centers to view the feedback of the customers to enhance the business.
4. The system can be installed at busy places like airports, railway stations or bus stations for detecting human faces and facial expressions of each person. If there are any faces that appear suspicious like angry or fearful, the system might set an internal alarm.
5. The system can also be used for educational purposes such as one can get feedback on how the student is reacting during the class.

2. Literature survey

The problem of classifying emotions from facial expressions in images is widely studied. Over the last two decades, researchers have significantly advanced human facial emotion recognition with computer vision techniques. The major challenge that the researchers face is the non-availability of spontaneous expression data. Capturing spontaneous expressions on images and videos is one of the biggest challenges ahead. Many attempts have been made to recognize facial expressions. Historically, there have been many approaches to this problem, including using pyramid histograms of gradients, AU aware facial features. However, recent top submissions to the 2017 Emotions in the Wild contest for static images all used deep convolutional neural networks (CNNs). A recent development by G. Levi showed significant improvement in facial emotion recognition using a CNN. The authors addressed two salient problems:

1) a small amount of data available for training deep CNNs and
2) appearance variation usually caused by variations in illumination. They used Local Binary Patterns (LBP) to transform the images to an illumination invariant, 3D space that could serve as an input to a CNN.

A notable implementation of a CNN to real-time detection of emotions from facial expressions is by S. Oullet. The author implemented a game, where a CNN was applied to an input video stream to capture the subject's facial expressions, acting as a control for the game. This work demonstrated the feasibility of implementing a CNN in real-time by using a running average of the detected emotions from the input stream, reducing the effects of variation and noise.

3. Methodology

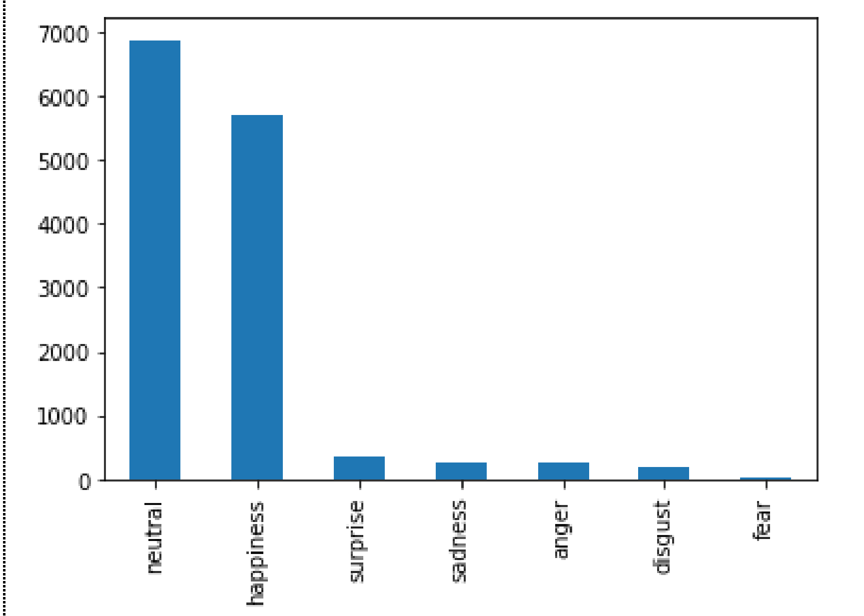
We are using a dataset having approx 13000 images that can be downloaded from [here](#).

METHOD 1: Facial expression recognition using CNN, these are the following steps:

1. Data preparation

Below plots show stats of the dataset, the frequency of particular emotion:

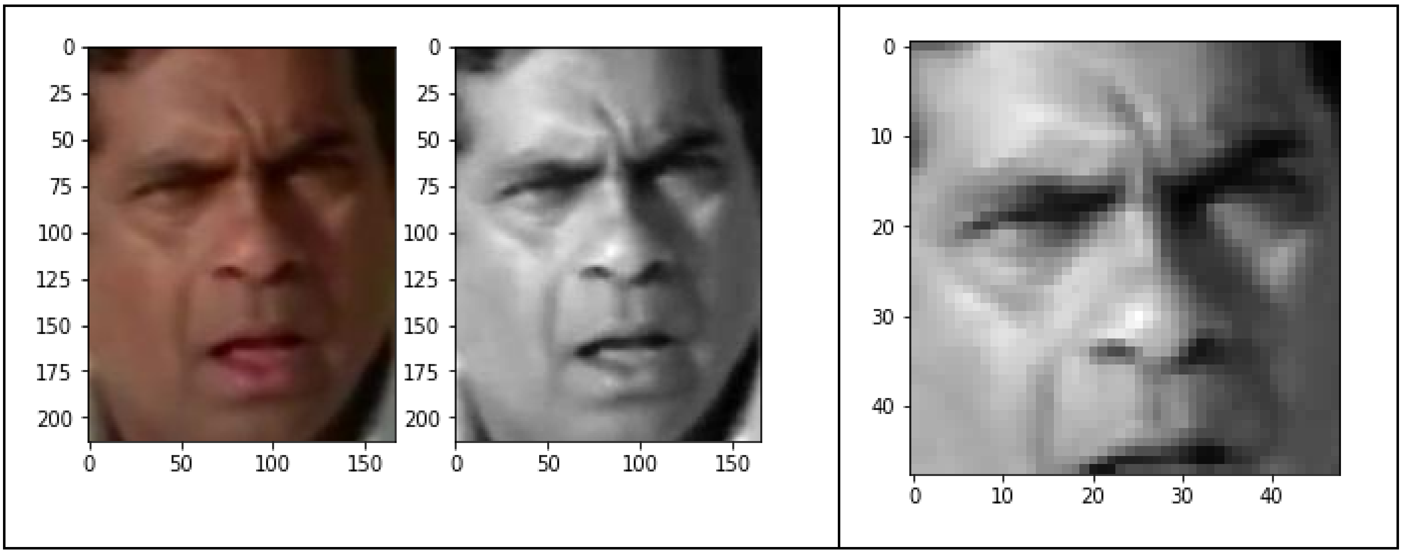
anger	252	252
contempt	9	9
disgust	208	208
fear	21	21
happiness	5696	5696
neutral	6868	6868
sadness	268	268
surprise	368	368



As we can see in the plot above, the contempt emotion doesn't have many records in the dataset. Therefore, it will not be as useful as others. That's why we will remove it by adding its images into the angry category :

2. Image Processing

In this step, we will make our data more accurate and useful.

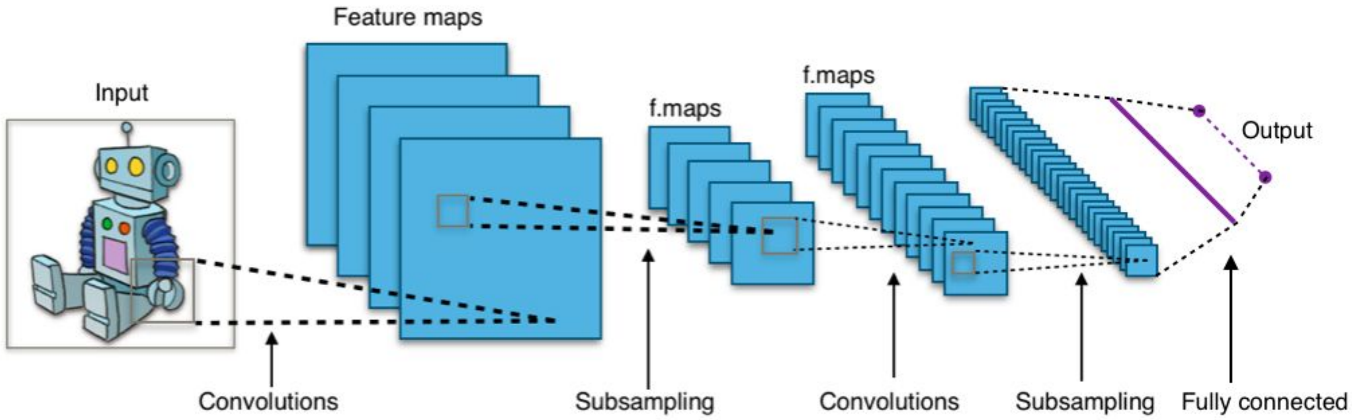


- Colored images contain some information that is considered as noise in the image processing work. That's why we need to convert them to grayscale.
- Sometimes images contain other types of noise like background information and so on. That's why we also need to perform cropping on images in order to keep only the regions of interest.

In our case, we want to have images of size (48,48). But before cropping them we need to make sure that we keep the face region in the cropped image. That's why we need to proceed by detecting the face first and then crop the image.

3. Build a Convolutional Neural Networks Model
4. Test the model on another Dataset

These are the basic steps for CNN.



By Aphex34 - Own work, CC BY-SA 4.0,
https://commons.wikimedia.org/w/index.php?curid=45679374

METHOD 2: Image feature extraction method, these are the following steps:

1. Image preprocessing

The image preprocessing step comprises operations like image scaling, image brightness, and contrast adjustment and other image enhancement operations.

2. Feature extraction

The set of key parameters that best describe the particular set of facial expressions needs to be extracted from the image such that the parameters can be used to discriminate between expressions is called the feature vector of the image that helps in feature extraction technique.

3. Two D Gabor function

4. Principal component analysis

PCA is a technique used to lower the dimensionality of a feature space that takes a set of data points and constructs a lower-dimensional linear subspace.

5. Classifier

- Assemble the training data.
- Create the network object.
- Train the network.
- Simulate the network response to new inputs.

→ Image Processing → Feature Extraction → classification → Happy Sad Surprise Anger Disgust Fear Neutral

4. Experiments and Conclusion

Visualization of principal components via eigenfaces, and facial expression recognition as mentioned above was implemented in Python. Along with the Python programming language, Numpy library was used.
The Nearest Neighbour classifier provided 66.47 % accuracy.
The confusion matrix for seven facial expression classes is shown below:

Predicted Actual	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
Angry	0.65	0.01	0.06	0.11	0.07	0.07	0.04
Disgust	0.04	0.67	0.05	0.1	0.05	0.04	0.05
Fear	0.06	0.01	0.65	0.1	0.06	0.07	0.05
Happy	0.06	0.01	0.06	0.67	0.08	0.07	0.05
Sad	0.06	0.01	0.06	0.12	0.65	0.07	0.05
Surprise	0.06	0.001	0.06	0.11	0.08	0.65	0.03
Neutral	0.04	0.01	0.05	0.07	0.05	0.04	0.74

Classification report:
The precision, recall and F1-score of each facial expression class is shown below:

	precision	*recall*	*f1-score*	*support*
angry	0.64	0.65	0.65	4953
disgust	0.62	0.67	0.64	547
fear	0.65	0.65	0.65	5121
happy	0.68	0.67	0.68	8989
neutral	0.65	0.65	0.65	6198
sad	0.66	0.65	0.65	6077
surprise	0.70	0.74	0.72	4002
accuracy			0.66	35887
macro avg	0.66	0.67	0.66	35887
weighted avg	0.66	0.66	0.66	35887
accuracy:	*0.6644746008303842*			

The overall precision and recall are 0.66 and 0.67 respectively. The model performs really well on classifying positive emotions resulting in relatively high precision scores for happy and surprised. Happy has a precision of 0.68 and recall of 0.67 which could be explained by having the most examples (6500) in the training set. Interestingly, surprise has the highest precision and recall as 0.70 and 0.74 having the least examples in the training set. There must be very strong signals in the surprise expressions.

Model performance seems similar across all emotions on average. The emotion sad has a precision of only 0.66 and recall 0.65. The model frequently misclassified angry, fear and neutral as sad. In addition, it is most confusing when predicting sad and neutral faces because these two emotions are probably the least expressive (excluding crying faces).

The overall F1-score is also 0.66 . F1-score is highest for surprise. Happy and surprise have higher F1-score as 0.68 and 0.72 respectively compared to other emotions. Disgust has the least F1-score as 0.64 and sad, anger and neutral also have similar F1-score.