

A Review on Facial Expression Recognition using Deep Learning

Megha Chandran
P G Student
Dept.of ECE,
LBS Institute of Technology for Women,
Kerala

Dr. Naveen S
Assistant Professor
Dept.of ECE,
LBS Institute of Technology for Women,
Kerala

Abstract— The advent of artificial intelligence technology has reduced the gap of human and machine. Artificial intelligence equips man to create more near perfect humanoids. Facial expression is an important tool to communicate one's emotions non verbally. This paper puts an overview of how deep neural networks can become an efficient tool for classifying facial expressions.

INTRODUCTION

As the world is fast changing we are depending on various technologies as a part of human machine interaction. Facial expression is one of the major aspects of human emotion recognition which can be used in identifying the interpersonal relationship. Face plays an important role in social communication, and hence facial expressions are vital in judging the once attitude. Facial expressions not only exposes the sensitivity or feelings of any person but can also be used to judge his/her mental views. Facial expression recognition can be used in emotion and sentimental analysis. It is a method to recognize expressions on one's face. A wide range of techniques have been proposed to. A wide range of techniques have been proposed to detect expressions like happy, sad, fear, disgust, angry, neutral and surprise. Facial expression estimation and other gestures convey nonverbal communication cues in face to face interaction. Main blocks in facial recognition system are Image acquisition, Pre processing classification and post processing. Considering image acquisition, images can be static or image sequences. An image sequence contains potentially more information than a still image, since the former depicts the temporal characteristics of an expression.

PREPROCESSING

To learn meaningful features pre-processing is required to align and normalize the visual semantic information conveyed by the face. The steps involved in this process is detection of the face and non-face, removal of non-face. Viola Jones classifier is the state of art for this purpose which is simple and robust. As per Viola Jones, Haar like features are used. This is a widely used method in face detection. Viola Jones scores best in the sense that it requires only a small number of parameters and it classifies quickly with less rate of false positives. Rotation rectification is also needed for a facial image which is usually obtained with the help of landmark such as eyes. Before extracting the features the picture quality has to be improved such as variation in

light, effect of different backgrounds and head pose and to normalise the semantic information on the face.

FEATURE EXTRACTION

Different faces have different dimensions which include height/width of the face, colour, and width of other parts like lips, nose etc. The numerical representation of face or an element in the training set is called a feature vector. The features extracted namely are geometric features such as the shapes of the facial components such as eyes mouth etc, and locations of facial characteristics points (corners of the eyes mouth etc) or appearance features representing the texture of the facial skin in specific facial areas including wrinkles, bulges and furrows. Face detection methods can be distinguished into two groups. Holistic and analytic. Holistic means spanning the whole face, analytic means spanning subpart of the face. In geometric feature extraction facial landmark points have to be extracted and the converted into a feature vector which is the information in terms of angle, distance position [8]. When a computer is trained with numerical description of face and background the characteristics are extracted. This can be done using ADABOOST classifier. Viola Jones is a classifier used to obtain a strong classifier from the combinations of weak classifiers i.e. the integral image and cascading techniques is used here [7]. Appearance features are used to model the appearance variation of particular face and is a holistic facial analysis. Appearance feature include learned image filters from PCA, Gabor filter feature based on edge orientation histogram etc. [6] In geometric feature extraction method by segmentation process is performed to divide the face image into four regions, mouth nose and two eyes regions. Facial characteristics are recognised from the segmented regions. From these facial characteristics points scale of database. Major deep learning networks that can be used the lengths are calculated.

DEEP LEARNING BASED FER APPROACHES

Deep neural network has gained rapid advancement in human machine interface. It has shown drastic change in the field of pattern recognition. The advantages of deep neural network is that it can automatically learn the pattern and does not require a structured data as input. The successful reason behind deep learning is the availability of large for FER are Deep Belief Networks to automatically learn facial expressions and multilayer perceptron can be used to

recognize these features. SVM (Support Vector Machine) was used for classification. CNN is commonly used to extract features automatically and classify it. The best design of local to global feature learning is with convolution, pooling and layered architecture which gives good visual effect and hence used for FER. On three widely used facial expression databases (CK+, Oulu-CASIA, and MMI), Lisa Graziani Dinfo, Stefano Melacci In [6] used CNN to study the effects of the different face parts in the recognition process, based on the appearance and shape representations for the face by localizing face area with a set of landmark points and landmark detector [6]. The detector uses the classic Histogram of Oriented Gradients (HOG) features combined with a linear classifier, an image pyramid, and a sliding window detection scheme. From these landmark points appearance based and shape based representations were obtained and finally applied to CNN for classification process. Muhtahir, O. Oloyede, Gerhard P. Hancke and Herman C. Myburgh, suggested a new approach for improving FER using image enhancement technique by applying a transformation with an evaluation function [1]. Hybrid features are used for feature extraction from that comes from PHOG, EHD and LBP. Local binary pattern is a texture feature having the advantage of easy calculation and small data size. These hybrid features then applied to CNN so that it transforms the input features in a layer by layer to final class. According to Ji-Hae Kim, Byung-Gyu Kim, Partha Pratim Roy, Da-Mi Jeong [2], dimension reduction method is used through a fusion of binary feature extraction. Dimensionality reduction is achieved by using PCA (Principal Component Analysis) and LDA and fusion techniques applied are LBP and LDA (Linear Discriminant Analysis). A. T. Lopes, E. de Aguiar, A. F. de Souza, and T. Oliveira-Santos [9] experimented with three learning stages with CNN as a classifier. They also attempted to reduce rotation problem and then fed to CNN which is good at learning transformation invariant function. In this algorithm they applied a cropping technique to remove the unnecessary elements around the face and emotions are classified into six to seven using CNN. Tong Zhang, Wenming, Zhen Cui, Yuan Zong, Jingwei Yan, Keyu Yan [5] proposed a deep neural network driven feature learning method which was applied to multiview facial expression recognition. In this method a scale invariant feature transform (SIFT) features corresponding to a set of landmark points are extracted from each facial image. This SIFT vectors is the data which when given to DNN and it learns a set of optimal features and classifies the facial expressions across the different facial views. The data base used were BU-DFE and multi-PIE. Another fusion technique for automatic facial expression recognition system using DNN proposed by Anima Majumder, Laxmidhar Behera, Venkatesh K Subramanian [3] consists of SOM (self organizing map) based classifier. The geometric and appearance features are then fused using autoencoders to get a better representation of facial attributes. The technique was implemented with a database of MMI and CK+. Auto encoder is an unsupervised learning algorithm

which applies back propagation algorithm to approximate the features vector.

FACIAL EXPRESSION DATA SET

Every machine learning algorithms take a dataset as input and learns from the data given. The algorithm goes through the dataset and identifies the pattern in it. Facial expression may be interpreted in different ways so an ideal dataset should contain large sample images with face attribute labels. Most commonly used datasets for FER are CK(+), MMI, JAFFE, TFD, FER2013, AFEW, Multi-PIE, BU-3DFE.

DEEP LEARNING NETWORKS FOR FER

By considering relatively small dataset to avoid overfitting certain networks can be selected for different task. The networks can be fine-tuned at the end layers using back propagation algorithm for efficient results. Few of them are Alexnet, VGG, VGG(face), Googlenet

CONCLUSION

This review focuses on various techniques for implementing a deep neural network based facial recognition system. This also briefs the networks and data bases used for FER system.

REFERENCES

- [1] Muhtahir, O. Oloyede, Gerhard P. Hancke and Herman C. Myburgh "Improving Face Recognition Systems Using a New Image Enhancement Technique, Hybrid Features and the Convolutional Neural Network", IEEE Access, December, 2018.
- [2] Ji-Hae Kim, Byung-Gyu Kim, Partha Pratim Roy, Da-Mi Jeong "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure, IEEE Access April, 2019.
- [3] Anima Majumder, Laxmidhar Behera, Venkatesh K Subramanian "Automatic Facial Expression Recognition system Using Deep Network Based Data Fusion" IEEE transactions on Cybernetics, Vol. 48, No. 1, January 2018
- [4] Si Miao, Haoyu Xu, Zhenqi Han, Yongxin Zhu "Recognising facial Expressions using a shallow Convolutional Neural Network", IEEE Access, June 27, 2019.
- [5] Tong Zhang, Wenming, Zhen Cui, Yuan Zong, Jingwei Yan, Keyu Yan "A Deep neural Network-driven Feature Learning Method For Multiview Facial Expression Recognition, IEEE transactions on Multimedia, Vol. 8, No. 12, December 2016
- [6] Lisa Graziani DINFO, University of Florence, Italy lisa.graziani@unifi.it Stefano Melacci, Marco Gori "Coherence Constraints in Facial Expression Recognition" cs.CV, October 2018.
- [7] Tanmoy Paul, Unmmul Afia Shammi, Mosabbear Uddin A F Med, Ra Shehur Rahmn, Syoji Atiqur Rahman AhAD "A Study on Face Detection Using Viola Jones Algorithm for Various Backgrounds Angles and Distances" Biomedical soft Computing and
- [8] Aliaa A A Youssif "Automatic Facial Expression Recognition system based on geometric and appearance features" Computer and information Science, Vol. 4, No. 2, March 2011
- [9] A. T. Lopes, E. de Aguiar, A. F. de Souza, and T. Oliveira-Santos, "Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order," Pattern Recognit., vol. 61, pp. 610–628, Jan. 2017.
- [10] Y. Liu et al., "Facial expression recognition with PCA and LBP features extracting from active facial patches," in Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR), Angkor Wat, Cambodia Jan. 2016, pp. 368–373.