Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 93 (2016) 453 – 461

Procedia Computer Science

6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India

Automatic Facial Expression Recognition Using DCNN

Veena Mayya, Radhika M. Pai*, Manohara Pai M. M.

Department of Information & Communication Technology, Manipal Institute of Technology, Manipal University, Karnataka-576104, India

Abstract

Face depicts a wide range of information about identity, age, sex, race as well as emotional and mental state. Facial expressions play crucial role in social interactions and commonly used in the behavioral interpretation of emotions. Automatic facial expression recognition is one of the interesting and challenging problem in computer vision due to its potential applications such as Human Computer Interaction(HCI), behavioral science, video games etc.

In this paper, a novel method for automatically recognizing facial expressions using Deep Convolutional Neural Network(DCNN) features is proposed. The proposed model focuses on recognizing the facial expressions of an individual from a single image. The feature extraction time is significantly reduced due to the usage of general purpose graphic processing unit (GPGPU). From an evaluation on two publicly available facial expression datasets, we have found that using DCNN features, we can achieve the state-of-the-art recognition rate.

Peer-review under responsibility of the Organizing Committee of ICACC 2016

Keywords: Facial expression recognition;Computer Vision;Machine Learning;Confusion Matrix;Support vector machine; JAFFE; CK+; Deep Convolutional Neural Network;

1. Introduction

Facial expression is an important part of nonverbal communication. Human expression recognition is influenced by certain context. When a subject is being investigated, the investigator might be diverted by the subject's voice tone or argument and may forget to keep track of the facial expressions. Automatic facial expression recognition systems are exempt to such contextual interference. Such systems can be beneficial in many fields, like gaming applications, criminal interrogations, psychiatry, animations etc. State-of-art approaches attempt to recognize six basic facial expressions such as anger, disgust happiness, sadness, surprise and fear.

Facial expression recognition techniques are based on either appearance features or geometry features [1]. Geometric features are extracted from the shape of the face and its components such as the eyebrows, the mouth, the nose etc. Appearance features are extracted using the texture of the face caused by expression, such as furrows, wrinkles etc. In 1970s Paul Ekman and Wallace V. Friesen, developed Facial Action Coding System (FACS)[2] which is the most

* Corresponding author. Tel.: +91-08202925362;
E-mail address: radhika.pai@manipal.edu

Peer-review under responsibility of the Organizing Committee of ICACC 2016
doi:10.1016/j.procs.2016.07.233

widely used method for describing and measuring facial behaviors. FACS is a system designed for human observers to describe changes in facial expression in terms of observable facial muscle actions known as facial action units or AUs. FACS is demonstrated to be a powerful means for detecting and measuring facial expressions and is recently used for feature extraction in combination with other techniques such as Dynamic Bayesian Network (DBN)[3] and Local Binary Pattern (LBP)[4]. Histograms of oriented gradients(HOG)[5], Scale Invariant Feature Transform (SIFT)[6], Local Binary Pattern (LBP)[7] are few state-of-art techniques for extracting facial features. Most of the above techniques use handcrafted features for facial expression recognition, and therefore require particular efforts both in terms of computation cost and programming effort.

In recent years, deep learning using convolution neural networks(CNNs) for feature extraction of image data is becoming more popular. Their popularity stems from their ability to extract good representations from image data. DCNN's computation intensive tasks can run on GPU, which results in high performance at very low power consumption. They have also yielded high performance for some of challenges such as the CNN based model proposed by Kim et al.[8]. CNN is extensively used for facial feature extraction for determining age[9], gender[10] etc.

2. Related Work

Several methods have been reported in the literature to automatically recognize facial expressions. Lucey et al.[11] manually labeled 68 facial points in key frames and used a gradient descent Active Appearance Model (AAM) to fit these points in the remaining frames. It may not be possible to obtain accurate key points in many practical situations. A study on using LBP for facial expression recognition is proposed by Shan et al.[12]. Here expression recognition is accomplished using support vector machine (SVM) classifiers with boosted-LBP features. In[12], authors manually labeled eye positions, which is not feasible in many practical cases. Computer Expression Recognition Toolbox (CERT) is proposed by Littlewort et al.[13]. CERT convolves through the registered face image with Gabor filters to extract the facial features and uses SVM and multivariate logistic regression (MLR) classifiers to recognize facial expressions. Lyons and Akamatsu[14] proposed a system for coding facial expressions with 2D Gabor wavelets for feature extraction, having clustering for classification. These methods also require particular efforts, both in terms of computation cost and programming effort.

Deep Convolution Neural Network (DCNN) framework is widely used for extraction of features from the images. DCNN uses several layers leading accurate feature learning. Here the prelearned features are used as filters and these filters convolves through the input image and produces the features which in turn are used by other layers of the network as discussed by Krizhevsky et al.[15]. Techniques based on convolutional neural networks have been proposed for facial expression recognition such as the model proposed by Kahou et al.[16]. But they extensively train the model with other facial dataset. Sébastien[17] used Deep Convolutional Activation Feature for Generic Visual Recognition(DeCAF)[18] for facial feature extraction that does not require extensive training, but DeCAF is too slow to use it for training even the small image dataset as it does not support GPU.

3. Proposed Method

In this work, automatic facial expression recognition using DCNN features is investigated. Two publicly available datasets CK+[11] and JAFFE[20] are used to carry out the experiment. Pre-processing step involves face detection for the above two datasets. The frontal faces are detected and cropped using OpenCV[21]. Then facial features are extracted using the DCNN framework. Algorithm 1 illustrates the steps for recognizing facial expressions. Subsection 3.1 describes the facial feature extraction using Convolutional Architecture for Fast Feature Embedding(Caffe) framework[19].

3.1. Feature Extraction using Caffe

Feature extraction is performed using Caffe on Graphics Processing Unit (GPU). The convolution neural network architecture, which is used for ImageNet [15] object detection is used to extract facial features. ImageNet uses eight learned layers which includes five convolutional and three fully connected layers for object detection. In this work, features are extracted using only first five layers. These layers are combination of convolution, Rectified Linear