ABSTRACT

This paper discusses the application of feature extraction of facial expressions with combination of neural network for the recognition of different facial emotions (happy, sad, angry, fear, surprised, neutral etc..). We use transfer learning on the fully connected layers of an existing convolution neural network which was retrained for human emotion classification. A varietyof datasets, as well as our own image dataset, is used to train the model .Humans are capable of producing thousands of facial actions during communication that vary in complexity, intensity, and meaning. This paper analyses the limitations with existing system emotion recognition. This project achieved 85 percent accurate results. Finally, a live video stream connected to a face detector feeds images to the neural network. The network subsequently classifies an arbi- trary number of faces per image simultaneously in real time . Purposed system depends upon human face as we know face reflects the emotions. The results demonstrate the feasibility of implementing neural networks in real time to detect human emotion .

Keywords: convolutional neural network, feature extraction,Emotion

TABLE OF CONTENT

LIST OF ABBREVIATIONS:-

ABBREVIATION	MEANING
AI	Artificial Intelligence
ANN	Artificial Neural Network
CV	Computer Vision

CHAPTER 1

INTRODUCTION

1.1 INTRO

Ever since computers were developed, scientists and engineers thought of articially intelligentsystems that that are mentally and/or physically equivalent to humans. In the past decades, theincrease of generally available computational power provided a helping hand for developing fast learning machines, whereas the internet supplied an enormous amount of data for train- ing. These two developments boosted the research on smart self-learning systems, with neuralnetworks among the most promising techniques.

1.2 HUMAN EMOTION

A key feature in human interaction is the universality of facial expressions and body lan- guage. Already in the nineteenth century, Charles Darwin published upon globally shared fa- cial expressions that play an important role in non-verbal communication . In 1971, Ekman andFriesen declared that facial behaviors are universally associated with particular emotions. Ap- parently humans, but also animals, develop similar muscular movements belonging to a certainmental state, despite their place of birth, race, education, etcetera. Hence, if properly mod- elled, this universality can be a very convenient feature in humanmachine interaction: a well trained system can understand emotions, independent of who the subject is. One should keep in mind that facial expressions are not necessarily directly translatable into emotions, nor vice versa. Facial expression is additionally a function of e.g. mental state, while emotions are also expressed via body language and voice. More elaborate emotion recognition systems should therefore also include these latter two contributions. However, this is out of the scope of this research and will remain a recommendation for future work. Readers interested in research on emotion classication via speech recognition are referred to Nicholson at all. As a nal point of attention, emotions should not be confused with mood, since mood is considered to be a long-term mental state. Accordingly, mood recognition often involves longstanding analysis ofsomeones behaviour and expressions, and will therefore be omitted in this work.

1.3 IMAGE CLASSIFICATION TECHNIQUES

The growth of available computational power on consumer computers in the beginning of the twenty-rst century gave a boost to the development of algorithms used for interpreting pic-tures. In the eld of image classication, two starting points can be distinguished. On the one hand pre-programmed feature extractors can be used to analytically break down several ele- ments in the picture in order to categorize the object shown. Directly opposed to this approach,self-learning neural networks provide a form of blackbox identication technique. In the latter concept, the system itself develops rules for object classication by training upon labelled sampledata. An extensive overview of analytical feature extractors and neural network approaches forfacial expression recognition is given by Fasel and Luettin [6]. It can be concluded that by the time of writing, at the beginning of the twenty-rst century, both approaches work approximatelyequally well. However, given the current availability of training data and computational powerit is the expectation that the performance of neural network based models can be signicantly improved by now.

1.4 DATASET

Neural networks, and deep networks in particular, are known for their need for large amounts of training data. Moreover, the choice of images used for training are responsible for a big partof the performance of the eventual model. This implies the need for a both high qualitative and quantitative dataset. For emotion recognition, several datasets are available for research, varying from a few hundred high resolution photos to tens of thousands smaller images.

The non-frontal faces and pictures with the label contemptuous are taken out of the RaFD data, since these are not represented in the FERC-2013 training set. Furthermore, with use of the Haar Feature-Based Cascaded Classier inside the OpenCV framework all data is preprocessed. For every image, only the square part containing the face is taken, rescaled, and converted to anarray with 48x48 grey-scale values.

1.5 TOOLS & PLATFORM

Minimum Requirement:-

Processor:- Pentium 4/1.2 Ghz Celeron/Duron process or

CPU Speed:- 1.4 GHz

RAM:- 512 MB

Hard Disk:- 80 GB

OS:- Windows 98/2000/XP

Mouse:-

Keyboard:-

Sound Card:- NO

Language:- Java 5.0

Server:- Apache Tomcat 4.1

CHAPTER 2

BACKGROUND

2.1 INTRO

This section focuses on the background and importance of facial expressions for sentiment analysis. It is bifurcated into four subsections. Firstly, we discuss the evolution of the time- line of facial recognition methods. Secondly, we discuss the need for Facial Detection, Dimension Reduction, and Normalization for sentiment analysis. In the third subsection, we focus on the need for feature extraction from the face image. Finally, we highlight the need for emotion classification

2.2 EVOLUTION TIMELINE

Figure 3 gives a brief overview on the evolutionary time- line of facial sentimental recognition methods given by the researchers across the globe along with the datasets. There exists various algorithms for FER such as traditional state- of-the-art algorithms and DL-based algorithms proposed by various researchers till 2020. The emotion recognition was first stated in the paper proposed by Bassili *et al.* in 1978 where authors have classified the emotions into six basic gestures such as happiness, sadness, fear, surprise, anger, and disgust. Different algorithms (traditional and DL) were used for FER by the authors. For example, Padgett *et al.*

2.3 NEED FOR FEATURE EXTRACTION

Facial Expression analysis comprises of various methods such as facial landmark identification, feature extraction, and different feature extraction databases. Facial landmarks are drawn by the facial key points which are derived from the geometry of the face [16]. Feature Extraction is done after preprocessing phase [49]. There are two methods available for feature extraction are appearance-based extraction and geometric-based extraction. The geometric-based method ex- tracts feature like edge features and corner features. Verma *et al.* [50] analyzed the performance of the feature extraction technique Gabor filter. They also tested the average gabor filter and compared both the filtering techniques to enhance the recognition rate.

The second method, which is an appearance-based method, takes care of the states of different points of the face, such as the position of the eye, shape of important points such as mouth and eyebrows using the salient point features. The majority of the traditional methods have used Local Binary Pattern (LBP) as the feature extraction technique, which is a generic-based framework for the extraction of features from the static image. It converts the most important features of the input image, as mentioned above, into a histogram.

2.4 NEED FOR EMOTION CLASSIFICATION

The third step in the FER is the Emotion Classification. There are various methods that are used for the classifica- tion of emotions after applying face detection and feature extraction algorithms. The various classification algorithms are convolutional neural network (CNN) [53], SVM, and restricted boltzmann machine (RBM). The most widely used method for classification is CNN. It is the most efficient algorithm as it can be applied directly to the input image without applying any feature extraction and face detection algorithms and still gets better accuracy over the input data [54]. The number of images in the training data set also has a huge impact on classification results. CNN faces a huge challenge in the training of limited-image dataset. So, the models which are built on a limited dataset can use the SVM algorithm for feature extraction and face detection. The emotions of a human are not static, it varies time-to-time. So, the classification of situation-based emotions is challenging.

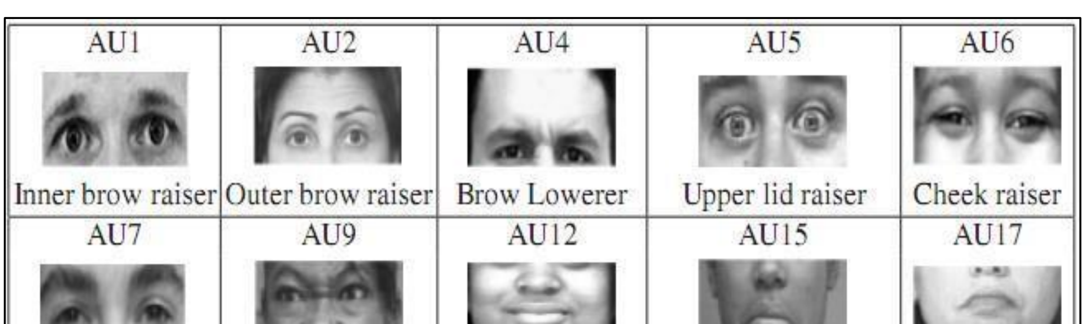
CHAPTER 3

IMAGE FEATURES

We can derive different types of features from the image and normalize it in vector form. We can employ various types of techniques to identify the emotion like calculating the ellipses formed on the face or the angles between different parts like eyes, mouth etc. Following are some of the prominent features which can be used for training machine learning algorithms:

3.1 FACTS

Facial Action Coding System is used to give a number to facial moment. Each such number is called as action unit. Combination of action units result in a facial expression. The micro changes in the muscles of the face can be defined by an action unit. For example, a smiling face can be defined in terms of action units as 6 + 12, which simply means movement of AU6 muscle and AU12 muscle results in a happy face. Here Action Unit 6 is cheek raiser and Action Unit 12 is lip corner puller. Facial action coding system based on action units is a good system to determine which facial muscles are involved in which expression. Real time face models can be generated based on them.



images, we have many feature detector algorithms in the OpenCV library such as Harris corner detector.

These feature detectors take into account many more factors such as contours, hull and convex. The Key-points are corner points or edges detected by the feature detector algorithm. The feature descriptor describes the area surrounding the key-point. The description can be anything including raw pixel intensities or co-ordinates of the surrounding area. The key-point and descriptor together form a local feature. One example of a feature descriptor is a histogram of oriented gradients. ORB (based on BRIEF), SURF, SIFT etc. are some of the feature descriptor algorithms.