RECOGNITION OF FACIAL EXPRESSION USING CNN AND VISUALISATION OF IT USING PLOTLY PYTHON

- **Ramachandran R**

CONTENTS

ABSTRACT

The aim of Real-time Data Visualization of Facial Expression Data using Keras and Plotly is to detect and classify human facial expressions from image sequence this is also used to cope with emotional health problems caused due to our negativity in our day to day lives. This software uses **biometric markers to detect emotions in human faces**. More precisely, this technology is a sentiment analysis tool and is able to automatically detect the six basic or universal expressions: happiness, sadness, anger, surprise, fear, and disgust. Emotional health plays very important role to improve people's quality of lives, especially for the elderly. This kind of improvement is an important progress in this artificial intelligence era. Facials expressions plays a key role in our daily communications. Many deep learning approaches have been applied in recent years due to their outstanding recognition accuracy after training with large amounts of data. This can be done basically in here steps: Firstly, Locating faces in the scene, in an image or video footage. Secondly, extracting information about facial features from detected faces and finally analyzing the movement of facial features or changes in the appearance of facial features and classifying this information into expression-interpretative categories such as facial muscle activations like smile or frown; emotion categories happiness or anger.

PROBLEM STATEMENT

Given a photo of a person, recognize different types of expression of the person by using deep learning techniques and visualize the types of expression by using different visualization tools.

INTRODUCTION

The first step towards the automatic recognition of facial expressions was taken in 1978 by Suwa et al. Suwa and his colleagues presented a system for analyzing facial expressions from a sequence of images (movie frames) by using twenty tracking points. Facial expression recognition has become an increasingly active research topic in the field of computer vision, as it plays an important role in many applications such as human-computer interaction and health care. Facial expression is a major way of human emotional communication. It is reported that facial expression constitutes 55% of the effect of a communication message while language and voice constitute 7% and 38%, respectively. Automatic face expression recognition systems find applications in several interesting areas. With the recent advances in robotics, especially humanoid robots, the urgency in the requirement of a robust expression recognition system is evident. This software find uses in a host of other domains like Telecommunications, Behavioral Science, Video Games, Animations, Psychiatry, Automobile Safety, Affect sensitive music juke boxes and televisions, Educational Software, etc. Practical real-time applications have also been demonstrated. Bartlett et al. have successfully used their face expression recognition system to develop an animated character that mirrors the expressions of the user. Another interesting application has been demonstrated by Anderson and McOwen, called the 'EmotiChat. As expression recognition systems become more real-time and robust, we will be seeing many other innovative applications and uses.

PYTHON MODULES

- Plotly
- Keras
- TensorFlow
- Pandas

LITERATURE SURVEY

We employ two standard facial expression databases for the simulation, which are both widely acknowledged by academia. JAFFE contains 213 images of 10 Japanese women, while CK+ covers the expression images of all races of people and has 328 pictures totally. Before the recognition, some pre-processing work need to be done firstly. In our image pre-processing procedure, we run a two-step process to reduce the interference in the original images, which are Face Detection and Histogram Equalization. The first step of image pre-processing is face detection. In the face detection part, detection results are based on the Haarlike feature in OpenCV, which is one of the most classic features for face detection. It was originally proposed by Papageorgiou et al. [2] and also known as the rectangular feature. Haarlike feature templates are divided into three categories, namely edge features, linear features and center surround features. [1]

A depth camera is used to capture various depth images from depth videos of the objects that generate RGB and depth information at the same time. The depth sensor video information or data demonstrate the scope of every pixel in the science as a gray level power or intensity. Figs. 2(a) and (b) represent a happy RGB and depth image respectively. [3] The depth images indicate the bright pixel values for near and dark ones for far distant face parts. Fig. 3 shows a sequence of grey and depth faces from surprise and disgust expressions respectively. [2]

The facial expression in a video shows as a successive process from the onset state, the apex state to the offset state. The peak frame is defined as the frame in the image sequence with a maximum expression state in the apex. Usually, a given image sequence contains a lot of frames, and it is impossible to consider every frame to recognize the expression in the video considering two aspects. The first one is that extracting the face feature of every frame is time-consuming which is not suitable for real-time application. Besides, an expression frame which is not in the maximum state contains too much ``noise" information which is often caused by various emotions. [2] Thus, the recognition result based on such a frame is not reliable. In our automatic facial expression recognition frame, a referential frame in the natural expression is pre-captured. Then, we propose the DLBP algorithm to extract the feature of every frame in the image sequence and the referential frame, respectively, to detect the peak frame. The proposed DLBP is a variant of LBP with only 24-demension which can effectively reduce the peak frames detection time in the real-time applications. [3]

Benefiting from the power of BRNN to store and access to the long-range contextual information, we propose a Part based Hierarchical Bidirectional Recurrent Neural Network (PHRNN) for facial expression recognition. The temporal information is the variations of the facial critical areas implied in sequential frames, which can be well mapped to facial landmarks. According to facial physical structure of a human face, we divide facial landmarks into four parts, *i.e.*, eyebrows, eyes, nose and mouth. [2] All of facial expressions can be performed by these parts. For example, happiness causes the corners of the lips up, disgust causes the eyebrows and eyes shrink, while surprise can be decomposed to eyes larger with mouth open widely. In order to learn powerful features from the facial critical areas, the four parts of landmarks are fed into four BRNN subnets, respectively. [4]

The concept of Convolutional Neural Network (CNN) was presented by Yann LeCun et al. in [7] in the 1980s, where a neural network architecture was composed of two kinds of basic layers, respectively called convolutional layers (C layers) and subsampling layers (S layers). However, many years after that, there was still not a major breakthrough of CNN. One of the main reasons was that CNN could not get ideal results on large size images. But it was changed when Hinton and his students used a deeper Convolutional Neural Network to reach the optimal results in the world on ImageNet in 2012. Since then, more attention has been paid on CNN based image recognition.[5]

Facial expression recognition is the most natural way to express human emotions. In the last few decades, there has been a lot of research on facial expression recognition from videos [1], [2]. A typical expression recognition system con- sists of face image acquisition, feature extraction, training, and recognition. Most of the facial image features are very sensitive to noise and illumination variations. Hence, fea- tures tolerating noise and illumination changes can strongly contribute to generating a robust expression recognition system.[6]

Extended Cohn-Kanade (CK+) [2] has been widely utilized to research facial expression recognition for years. In each expression category for a person, there are about 15 images in a sequence, and the expression intensity changes from low to high. The first one or two images are regarded as the neutral expression while the last one or two images are selected as the expressions in full effect. Consequently, there are 630 images from CK+. [7] For experiments, a depth video-based facial expression database was recorded. The database consisted of frontal face-based RGB and depth videos of six different expressions (i.e., surprise, sad, happy, disgust, anger, and neutral). In the videos, a head motion was assumed to be very small and hence neglected. For each expression, there were 40 videos. [2] After obtaining the database, leave-one-out cross-validation or forty-fold cross-validation was considered on the videos of each facial expression. For each fold, 39 videos were applied for training and the remaining one for testing. Finally, results of these forty folds were combined to represent the final recognition rate. [8]

The Local Directional Strength Pattern (LDSP) assigns an eight-bit binary code to each pixel of a depth face. This pattern is calculated by considering positions with highest edge strengths in bright and dark regions for a pixel. [3] For a pixel in the image, the eight directional edge response values are calculated by Kirsch masks. In typical LDP, absolute values of edge responses are taken. Then, directions representing top absolute strengths are set to 1 and rest of them 0. Unlike LDP, the strengths are kept with their signs. Then, the binary positions with highest and lowest strength values are considered. LDSP patterns represent robust features for salient pixels in an image, especially edge pixels. [9] The Radboud Faces Database (RaFD) [3] is a high quality database of faces, which contains pictures of 8 emotional expressions, including Caucasian males and females, Cau- casian children, both boys and girls, and Moroccan Dutch males. Head poses vary from left side to right, and each pose is shot with three eye gazing directions. Compared to CK+, RaFD is more challenging to the recognition model.[10]

OBJECTIVE

The objective of this project is to develop an algorithm to recognise the expression of human face accurately and to visualize the expression into six universal emotions using different visualization techniques. In recent years, many papers have been published that use deep learning for facial emotion recognition. These papers used freely available datasets with state of art models achieving an accuracy of 0.66. With this in mind, a number of different models both new and old will be experimented with to arrive at a final model with comparable results.

PROPOSED SYSTEM

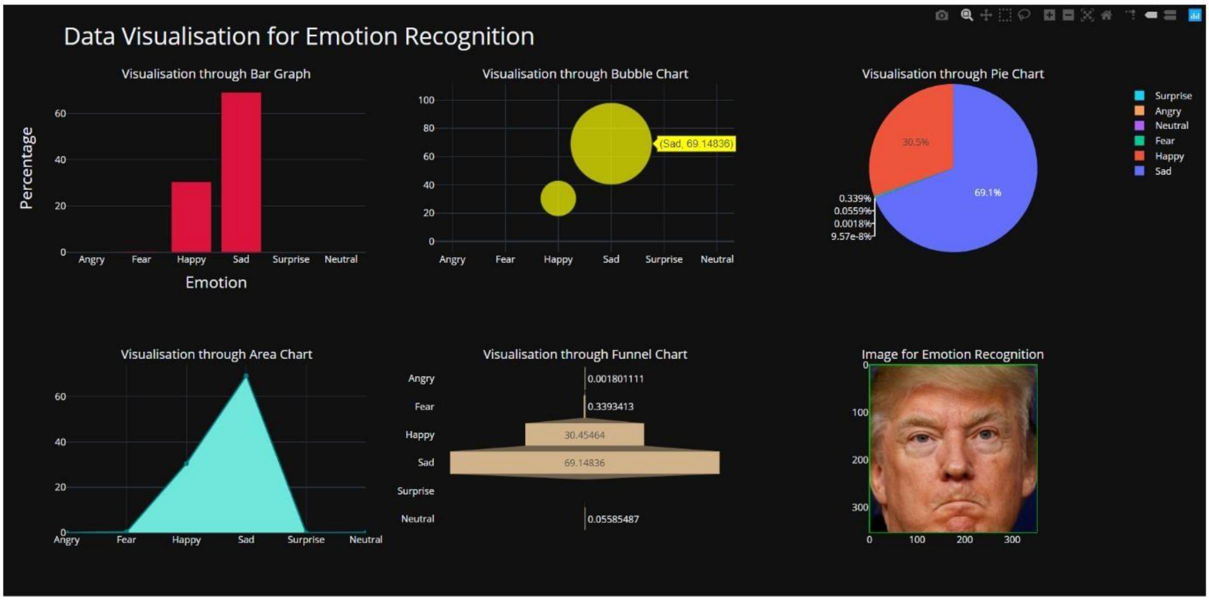
Using deep learning techniques such as Backpropagation Algorithm, we have improved the existing model such that it can predict the Facial Expression Data more efficiently.

CODE

```
from keras.preprocessing import image
from keras.preprocessing.image import ImageDataGenerator

import numpy as np
import tensorflow
from keras.models import load_model
import plotly.graph_objects as go
import plotly.offline as plo
from plotly import subplots
from skimage import data

model = load_model('model25.h5')
def emotion_analysis(emotions):
    objects = ('Angry', 'Fear', 'Happy', 'Sad', 'Surprise', 'Neutral')
    y_pos = np.arange(len(objects))
    fig1 = go.Bar(x=objects, y=emotions*100, marker={'color': 'crimson'}, show
legend=False, name="")

    fig2 = go.Pie(labels=objects, values=emotions*100, name="")

    fig3 = go.Funnel(y=objects,x=emotions*100,name="",marker={'color' : 'tan'}
, showlegend=False)                #funnel

    fig4 = go.Scatter(x=objects, y=emotions*100, name="",marker={'color' : 'te
al'}, showlegend=False, fill= 'tonexty', fillcolor='rgb(111, 231, 219)')

    fig5 = go.Scatter(x=objects, y=emotions*100, name="",mode= 'markers', mark
er={'color' : 'yellow', 'size' : emotions*150}, showlegend=False)

    img = data.imread("photo.jpg")
    fig6 = go.Image(z=img)

    figure = subplots.make_subplots(
    rows=2,
    cols=3,
    specs=[[{"type": "bar"}, {"type": "scatter"}, {"type": "pie"}],
           [{"type": "scatter"}, {"type": "funnel"}, {"type": "image"}] ],
    subplot_titles= ("Visualisation through Bar Graph","Visualisation through
Bubble Chart","Visualisation through Pie Chart",
```

```
                    "Visualisation through Area Chart","Visualisation through
Funnel Chart","Image for Emotion Recognition")
    )

    figure.add_trace(fig1, 1, 1)
    figure.add_trace(fig5, 1, 2)
    figure.add_trace(fig2, 1, 3)
    figure.add_trace(fig4, 2, 1)
    figure.add_trace(fig3, 2, 2)
    figure.add_trace(fig6, 2, 3)
    figure.update_layout(
        {
            "autosize" : True,
            "title" :{"text" : "Data Visualisation for Emotion Recognition", "
font" : {"size" : 30}},
            "xaxis_title" : {"text" : "Emotion", "font" : {"size" : 20}},
            "yaxis_title" : {"text" : "Percentage", "font" : {"size" : 20}},
            "template" : "plotly_dark"
        }
    )
    figure.show()

file = 'photo.jpg'
true_image = image.load_img(file)
img = image.load_img(file, color_mode = "grayscale", target_size=(48, 48))

x = image.img_to_array(img)
x = np.expand_dims(x, axis = 0)

x /= 255

custom = model.predict(x)
emotion_analysis(custom[0])
```

OUTPUTS



CONCLUSION

Face expression recognition systems have improved a lot over the past decade. The focus has definitely shifted from posed expression recognition to spontaneous expression recognition. The next decade will be interesting since robust spontaneous expression recognizers will be developed and deployed in real-time systems and used in building emotion sensitive HCI interfaces. This is going to have an impact on our day to day life by enhancing the way we interact with computers or in general, our surrounding living and work spaces. Having improved techniques to cope with expression variation, in the future it may be investigated in more depth about the face classification problem and optimal fusion of colour and depth information. Further study can be laid down in the direction of allele of gene matching to the geometric factors of the facial expressions. The genetic property evolution framework for facial expressional system can be studied to suit the requirement of different security models such as criminal detection, governmental confidential security breaches etc.

REFERENCES

[1] Shan, Ke, Junqi Guo, Wenwan You, Di Lu, and Rongfang Bie. "Automatic facial expression recognition based on a deep convolutional-neural-network structure." In 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA), pp. 123-128. IEEE, 2017.

[2] Uddin, Md Zia, Mohammad Mehedi Hassan, Ahmad Almogren, Atif Alamri, Majed Alrubaian, and Giancarlo Fortino. "Facial expression recognition utilizing local direction-based robust features and deep belief network." IEEE Access 5 (2017): 4525- 4536.

[3] Ding, Yuanyuan, Qin Zhao, Baoqing Li, and Xiaobing Yuan. "Facial expression recognition from image sequence based on LBP and Taylor expansion." IEEE Access 5 (2017): 19409-19419.

[4] Zhang, Kaihao, Yongzhen Huang, Yong Du, and Liang Wang. "Facial expression recognition based on deep evolutional spatial-temporal networks." IEEE Transactions on Image Processing 26, no. 9 (2017): 4193- 4203.

[5] Yang, Biao, Jinmeng Cao, Rongrong Ni, and Yuyu Zhang. "Facial expression recognition using weighted mixture deep neural network based on double-channel facial images." IEEE Access 6 (2017): 4630-4640.

[6] Uddin, Md Zia, Weria Khaksar, and Jim Torresen. "Facial expression recognition using salient features and convolutional neural network." IEEE Access 5 (2017): 26146- 26161.

[7] Wu, Bing-Fei, and Chun-Hsien Lin. "Adaptive feature mapping for customizing deep learning based facial expression recognition model." IEEE access 6 (2018): 12451- 12461.

[8] Lopes, André Teixeira, Edilson de Aguiar, Alberto F. De Souza, and Thiago Oliveira-Santos. "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order." *Pattern Recognition* 61 (2017): 610-628

[9] Mollahosseini, Ali, David Chan, and Mohammad H. Mahoor. "Going deeper in facial expression recognition using deep neural networks." In *2016 IEEE Winter conference on applications of computer vision (WACV)*, pp. 1-10. IEEE, 2016.

[10] Kim, Bo-Kyeong, Jihyeon Roh, Suh-Yeon Dong, and Soo-Young Lee. "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition." *Journal on Multimodal User Interfaces* 10, no. 2 (2016): 173-189.