Transactions on Computational and Scientific Methods | Vol. 4, No. 12, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Facial Emotion Recognition Using an Enhanced Xception Network for Behavioral Analysis

Sofia Moreno
National Autonomous University of Mexico, Mexico
sofia.moreno189@unam.mx

Abstract:

The rise in mental health issues such as depression and anxiety has necessitated efficient diagnostic tools. This study integrates facial recognition with psychological analysis, utilizing deep learning techniques to achieve accurate emotion recognition. An improved Xception network was developed, featuring densely connected depthwise separable convolution modules to optimize parameter usage and reduce computational demands, making it suitable for mobile applications. The model employs the FSRNet hypernetwork for facial feature extraction and multi-scale feature learning, achieving over 95% accuracy in facial emotion recognition. This approach provides a foundation for mental health prediction and diagnosis, significantly reducing the time and cost of psychological evaluations. Future research will focus on enhancing detection speed and extending functionality across platforms for broader applications in mental health analysis.

Keywords:

Deep Learning Face Correction; Human Face Correction; Xception.

1. Introduction

With the achievement of the battle for a well-off society in an all-round way in China, the focus of people's life has shifted from food, clothing, housing and tresure of life or study has brought people mental problems, such as depression, anxiety, bisexual affective disorder, and so on. Especially since the COVID-19, there have been more than 70 million patients with depression and 90 million patients with anxiety in the world. Professor Lu Lin, academician of the CAS Member and president of the Sixth Hospital of Peking University, said: "At present, the incidence rate of mental and psychological diseases in China has reached 17%"[1]. Therefore, it is urgent to pay attention to people's mental health.

With the development of current technosed because of its contactless, efficient and convenient features[2].

The method based on deep learning mainly uses convolution neural network to extract the characteristicinformation of the face in the image. Through the use of deep learning technology, the application of CNN algorithm to achieve the effect of face recognition has become the mainstream. More and more scholars, experts and business companies are trying to use the method of deep learning to carry out various face detection and psychological prediction.

This paper combines face recognition technology with psychology, builds a face emotion recognition model through deep learning and applying the Xception algorithm. By extracting facial feature values from the image samples of facial expressions, the recognition rate of facial emotion is more than 95%. Later, it can be used for preliminary prediction of human mental health, and then for diagnosis and treatment of human mental health, greatly reducing the cost and time of psychological diagnosis.

2. Face Image Preprocessing

In the process of face recognition, the subject may have involunt[lum]ination, blurring, scale transformation and other characteristics caused by other reasons, so it is necessary to preprocess the image [3].

2.1 Face Correction

The methods of face correction include five-point alignment or 68-point alignment. This paper adopts the 68-point alignment method. Compared with the other method, this method has more face key poin out our of the detected object. Its principle is to recognize the face key points. It divides the face key points into internal key points and contour key points. The internal key points include a total of 51 key points, including eyebrows, eyes, nose and mouth, and the contour key points include 17 key points. Finally, the input image is output as a set of facial feature points through affine transformation, similar transformation and other methods to achieve the effect of facial correction[4]

2.2 FSRNet Technology of Face Reconstruction

Face reconstruction mainly relies on FSRNet technology to restore low-resolution face to high-resolution face, that is, image clarity. FSRNet network structure is divided into coarse Image Super Resolution network and fine Image Super Resolution (hereinafter referred to as SR) network. The overall structure is as follows: The first step is to build a coarse SR network to generate a coarse Image High Resolution (hereinafter referred to as HR) image. In the second step, the coarse HR image will be sent to two branches. The first branch is the fine SR encoder, which is mainly used to extract image features; The second branch is a priori information prediction network, which estimates the landmark heatmap and analytic graph. Third, the image features and prior information will be sent to a fine SR decoder to recover the HR image. The specific network structure is shown in Figure 1:

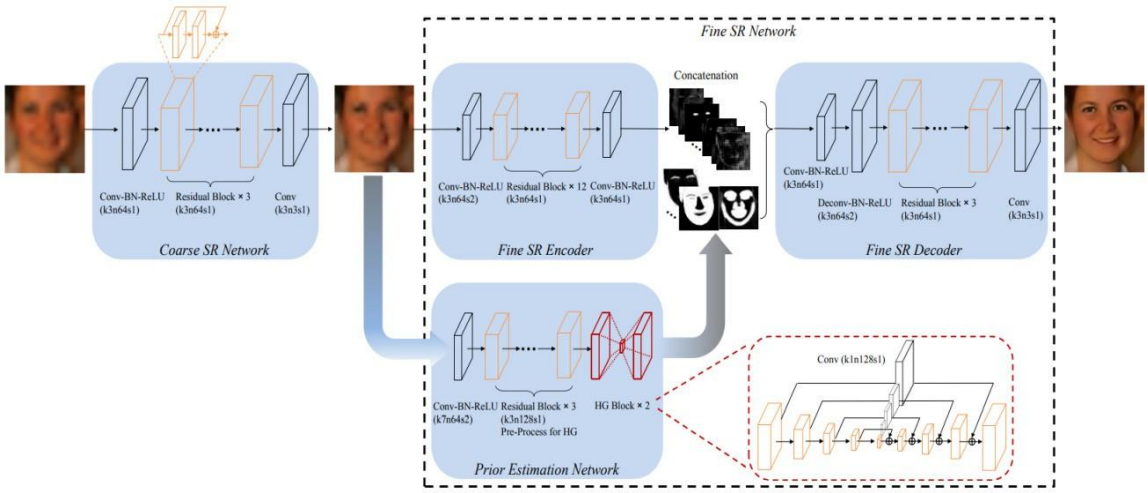


Figure 1. FSRNet Network Structure

3. Process and Pseudocode of Facial Emotion Recognition

3.1 Process of Facial Emotion Recognition

Step 1: Import mode cv2 load_ Model and other functions and libraries

Step 2: Load data and image parameters

Step 3: Start video streaming;

Step 4: extract the feature value of portrait

Step 5: import the eigenvalues into the model to get the emotional label

Step 6: Disconnect the video stream

Step 7: Export images with emotional labels

3.2 Core Code Introduction

In the process of recognition, the code for emotion recognition is written in Python language. The code flow is as follows:

Begin

Readline (Keras Callback functions, Image preprocessor ImageDataGenerator,Numpy database, sklearn database)

Use (Adam optimizer, categorical_crossentropy loss function)

Printf Model Summary

Reader (data set file) For(i=1;

i<;i++)

{i Read the first picture

Extract the image feature value;

If feature value extraction succeeded

do brings eigenvalues into the model;

printf sentiment tags

else feature value extraction succeeded

Do brings the eigenvalues into the model;

printf sentiment tags

}

Picture=emotional label + recognition picture i

Printf picture

The implementation of the image preprocessing function above is to convert the input image data into float32 format and normalize the pixel value to 0-1. If thev2 parameter is true, the pixel value will be further converted to - 1-1. The implementation of image emotion analysis, by defining a mini XCEPTION model, is used to process images with input shapes of (48, 48, 1), in which num classes are 7 (that is, 7 different categories. The dataset labels are divided into 7 categories, namely: * 0: 'angle', 1: 'dispute', 2: 'fear', 3: 'happy', 4: 'sad', 5: 'enterprise', 6: 'neutral'.

It first uses haarcascade_ frontalface_ The default.xml model detects the face in the image (this model is obtained from the network), and then uses fer2013_ mini_ XCEPTION.33-0.65.hdf5 model (this model is built by the program cnn.py) recognizes the emotion on the face, and finally draws the recognized emotion label on the image and saves it to images/predicted_ test_ Image.png file.

4. Simulation and Implementation

This paper will download FER-2013 database from a website. This database has 35887 face images in total. After applying the model, the facial emotion in the image will be recognized and simulated, as shown in Figure 2. After simulation, seven emotions, such as surprise, happiness and anger, can be identified. The recognition accuracy of the system is above 95%, as shown in Figure 3.

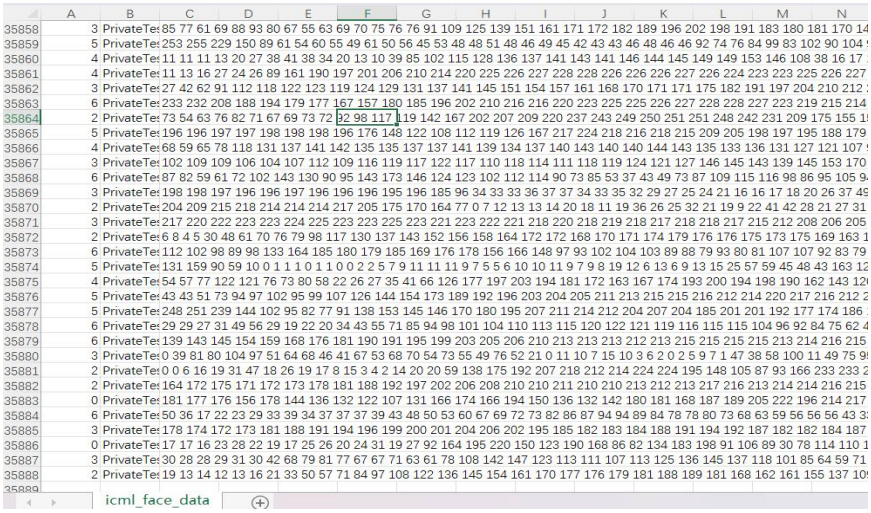


Figure 2. Part of original data

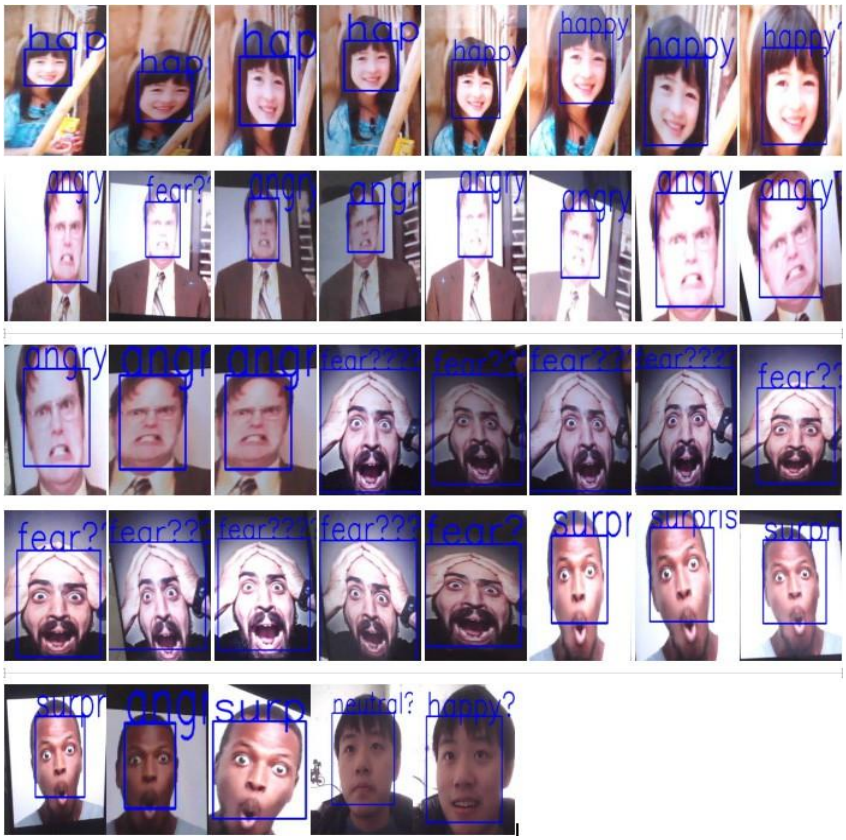


Figure 3. Partial Test Results

5. Conclusion

This paper proposes an emotion recognition method based on the improved Xception network. Through the dense connection of deeply separable convolution modules, it can reduce the amount of calculation parameters and make full use of model parameters, while taking into account the use of mobile terminals. The model uses the feature network extraction model for feature extraction and learning, which improves the ability of multi-scale features; Adopt FSRNet face hypernetwork model; At the same time, the Xception feature extraction network is applied to recognize the number of pictures, and the accuracy rate is more than 95%. In the future, the algorithm will be further improved to reduce the detection time. At the same time, the application on other platforms and psychological recognition and other functions will be studied to finally achieve mental health prediction, recognition

and diagnosis.

References

[1] Zhao Chunxiao World Mental Health Day | The incidence rate of mental illness in China has reached 17% [EB/OL] 2022[2023/2/13]. https://city.cri.cn/20221017/d53ef121-3602-5e19-8f6c-a6f2275dcaad.html.

[2] Zhang Yansheng, Nie Zhiyong, Sui Lilin. Face recognition in real complex mining environment based on deep learning [J]. Energy Technology, 2022,20 (05): 3-8.

[3] Chen Jiahao, Fu Xiaofeng, Zhang Jiaming. Face pain expression recognition based on lightweight Xception network [J]. Industrial Control Computer, 2022,35 (11): 109-110.

[4] Guo Zhongtian, Wang Ranfeng, Fu Xiang, Wei Kai, Wang Yulong. Feature extraction method of coal slime flotation foam velocity based on image feature matching [J]. Industrial and Mining Automation, 2022,48 (10): 34-39+54. DOI: 10.13272/j.issn.1671-251x.17991.

[5] Zhou Zihao, Tian Qiuhong. Gesture recognition based on improved Xception network [J]. Software Guide, 2022,21 (06): 41-48.

[6] Tan Yongjian, Tian Miao, Xu Dexin, Sheng Guanqun, Ma Kai, Qiu Qinjun, Pan Shengyong. Research on rock image classification and recognition based on Xception network [J]. Geographic and Geographic Information Science, 2022,38 (03): 17-22.