

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 treasure 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 characteristic information 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^lumination, 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 point 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:

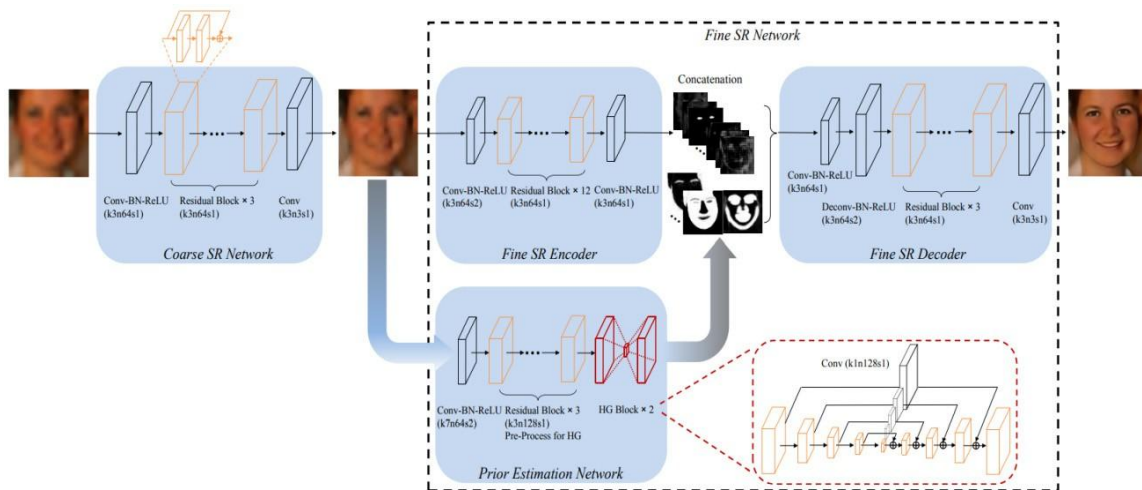


Figure 1. FSRNet Network Structure

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.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
35858		3 PrivateTe:	85 77 61 69 88 93 80 67 55 63	69 70 75 76 76 91 109 125	139 151 161 171 172 182	189 196 202 198 191 183 180 181 170 14								
35859		5 PrivateTe:	253 255 229 150 89 61 54 60 55	49 61 50 56 45 53 48 48 51 48	46 49 45 42 43 43 46 48 46 92	74 76 84 99 83 102 90 104 1								
35860		4 PrivateTe:	11 11 11 13 20 27 38 41 38 34 20	13 10 39 85 102 115 128 136	137 141 143 141 146 144 145	149 149 153 146 108 38 16 17								
35861		4 PrivateTe:	11 13 16 27 24 26 89 161 190 197	201 206 210 214 220 225 226	227 228 228 226 226 227 226	224 223 223 225 226 227								
35862		3 PrivateTe:	27 42 62 91 112 118 122 123	119 124 129 131 137 141	145 151 154 157 161 168	170 171 171 175 182 191 197	204 210 212							
35863		6 PrivateTe:	233 232 208 188 194 179 177	167 157 180 185 196 202	210 216 216 220 223 225	225 226 227 228 228 227	223 219 215 214							
35864		2 PrivateTe:	73 54 63 76 82 71 67 69 73 72	82 98 117 119 142 167	202 207 209 220 237	243 249 250 251 251 248 242	231 209 175 155 1							
35865		5 PrivateTe:	196 196 197 197 198 198 198	196 176 148 122 108 112	119 126 167 217 224	218 216 218 215 209 205	198 197 195 188 179							
35866		4 PrivateTe:	68 59 65 78 118 131 137 141	142 135 135 137 137 141	139 134 137 140 143	140 140 144 143 135 133	136 131 127 121 107							
35867		3 PrivateTe:	102 109 109 106 104 107 112	109 116 119 117 122 117	110 118 114 111 118	119 124 121 127 146	145 143 139 145 153 170							
35868		6 PrivateTe:	57 62 59 61 72 102 143	130 90 95 143 173 146	124 123 102 112 114	90 73 85 53 37 43	49 73 87 109 115 116 98 86 95 105 9							
35869		3 PrivateTe:	198 198 197 196 196 197 196	196 196 195 195 185 96	34 33 33 36 37 37 34 33	35 32 29 27 25 24	21 15 16 17 18 20 26 37 45							
35870		2 PrivateTe:	204 209 215 218 214 214	214 217 205 175 170	164 77 0 7 12 13 13	14 20 18 11 19 36 26 25 32 21	19 9 22 41 42 28 21 27 31							
35871		3 PrivateTe:	217 220 222 223 223 224 225	223 223 225 223 221 223	222 221 218 220 218	219 218 217 218	218 217 215 212 208 206 205							
35872		2 PrivateTe:	6 8 4 5 30 48 61 70 76 79	98 117 130 137 143 152	156 158 164 172 172	168 170 171 174	179 176 176 175 173 175 169 163 1							
35873		6 PrivateTe:	112 102 98 89 98 133 164 185	180 179 185 169 176	178 156 166 148 97	93 102 104 103	89 88 79 93 80 81 107 102 83 79							
35874		5 PrivateTe:	131 159 90 59 10 0 1 1 1 0	0 2 2 5 7 9 11 11 11	9 7 5 6 10 10 11 9 7 9	8 19 12 6 13 6 9 13 15 25 57 59 45 48 43 163 12								
35875		4 PrivateTe:	54 57 77 122 121 76 73 80	58 22 26 27 35 41 66	126 177 197 203 194	181 172 163 167 174	193 200 194 198 190 162 143 12							
35876		5 PrivateTe:	43 43 51 73 94 97 102 95 99	107 126 144 154 173	189 192 196 203 204	205 211 213 215 215	216 212 214 220 217 216 212 2							
35877		5 PrivateTe:	248 251 239 144 102 95 82	77 91 138 153 145 146	170 180 195 207 211	214 212 204 207 204	185 201 201 192 177 174 186							
35878		6 PrivateTe:	29 29 27 31 49 56 29 19	22 20 34 43 55 71 85	94 98 101 104 110 113	115 120 122 121 119	116 115 115 104 96 92 84 75 62 4							
35879		6 PrivateTe:	139 143 145 154 159 168 176	181 190 191 195 199	203 205 206 210	213 213 213 212 213 215	215 215 213 214 216 215							
35880		3 PrivateTe:	0 39 81 80 104 97 51 64 68	46 41 67 53 68 70 54	73 55 49 76 52 21 0	11 10 7 15 10 3 6 2 0 2 5 9 7 1 4	38 58 100 11 49 75 9							
35881		2 PrivateTe:	0 6 16 19 31 47 18 26 19 17	8 15 3 4 2 14 20 20	59 138 175 192 207	218 212 214 224 224	195 148 105 87 93 166 233 233 2							
35882		2 PrivateTe:	164 172 178 171 172 173	176 181 189 192 197	202 206 208 210	210 211 210 210 213 212	213 217 216 213 214 214 216 215							
35883		0 PrivateTe:	181 177 176 156 178 144 136	132 122 107 131 166	174 166 194 150 136 132	142 180 181 168	187 189 205 222 196 214 217							
35884		6 PrivateTe:	50 36 17 22 23 29 33 39	34 37 37 37 39 43 48 50	53 60 67 69 72 73	82 86 87 94 94 89 84 78 78	80 63 59 56 56 43 3							
35885		3 PrivateTe:	178 174 172 173 181 188 191	194 196 199 200 201	204 206 202 195	185 182 183 184 188 191	194 192 187 182 184 187							
35886		0 PrivateTe:	17 17 16 23 28 22 19 17	25 26 20 24 31 19 27	92 164 195 220 150	123 190 168 86 82 134 183	198 91 106 89 30 78 114 110 1							
35887		3 PrivateTe:	30 28 28 29 31 30 42 68 79	81 77 67 67 71 63 61 78	108 142 147 123 113 111	107 113 125 136 145 137 118	101 85 64 59 71							
35888		2 PrivateTe:	19 13 14 12 13 16 21 33	50 57 71 84 97 108 122	136 145 154 161 170	177 176 179 181 188 189	181 168 162 161 155 137 10							

Figure 2. Part of original data

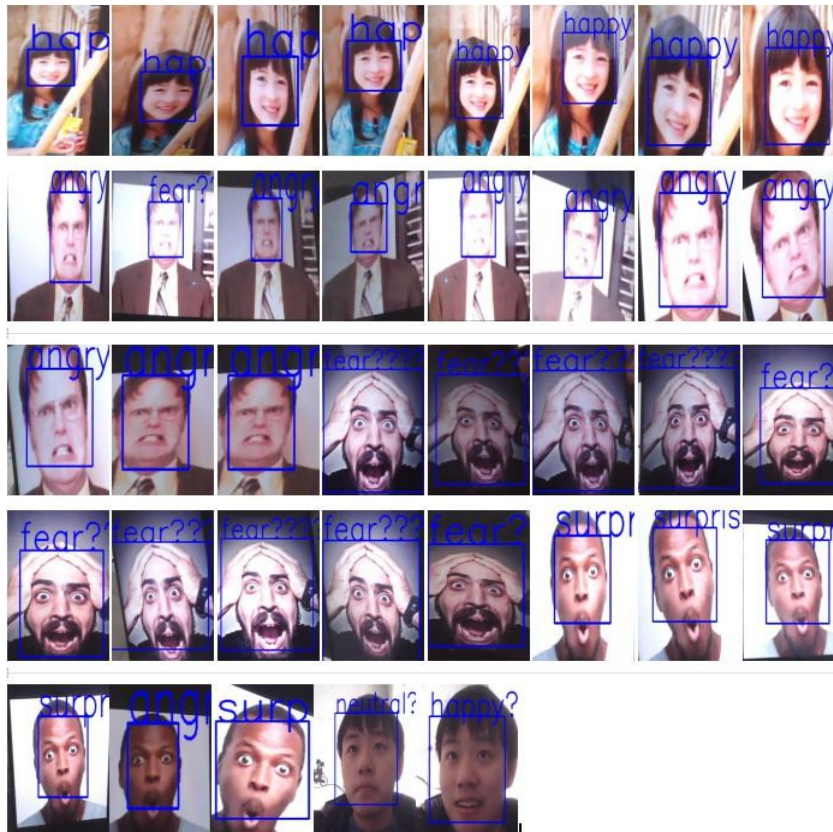


Figure 3. Partial Test Results

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 Qihong. 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.