

A Prediction of Emotions for Recognition of Facial Expressions Using Deep Learning

K.Sravanthi, G.Jaya Suma

Abstract—Automated facial expression recognition can greatly improve the human-machine interface. Many deep learning approaches have been applied in recent years due to their outstanding recognition accuracy after training with large amounts of data. In this research, we enhanced Convolutional Neural Network method to recognize 6 basic emotions and compared some pre processing methods to show the influences of its in CNN performance. The preprocessing methods are :resizing,mean, normalization, standard deviation, scaling and edge detection . Face detection as single pre-processing phase achieved significant result with 100 % of accuracy, compared with another pre-processing phase and raw data.

Key words: Emotion recognition; Convolutional neural network, feature extraction , facial action coding system.

I. INTRODUCTION

The process of expression recognition can be roughly divided into three parts: image acquisition and pre processing, image feature extraction and selection, classification Emotion is a kind of social facts that can be analysed through facial expressions. Basic emotions consist of anger, happy, disgust, fear, sadness, happy, and surprise . The application of automatic facial expression can be applied in wide variety of areas such as emotion and paralinguistic communication, clinical psychology, psychiatry, neurology, pain assessment, lie detection, intelligent environments, and multimodal human computer interface (HCI). There are two approaches in facial expression analysis, i.e. feature extraction and action unit detection from Facial action coding system (FACS). FACS is a framework proposed by Ekman et al. Feature extraction has two main techniques: geometric feature-based and appearance- based feature extraction. The former represents facial points to form feature vectors and to show the face from geometrical perspective. The latter is applied in the extraction of feature vectors, either in specific or holistic face area of facial image, such as in the Gabor Wavelets, or LBP application.

Deep learning is a part of machine learning approaches that can be utilized as emotion recognition and facial expression analysis. However, its performance depends on the data size. The greater the data the better the performance . The size of facial expressions datasets are still insufficient for deep learning to be implemented. Therefore, some researches apply augmentation techniques in the pre-processing step such as resize, scaling, mean, standard deviation, normalization, rotation, edge detection. This research aims to recognize emotions using deep learning and show the influences of data pre-processing in deep learning performance. Data pre-processing methods including

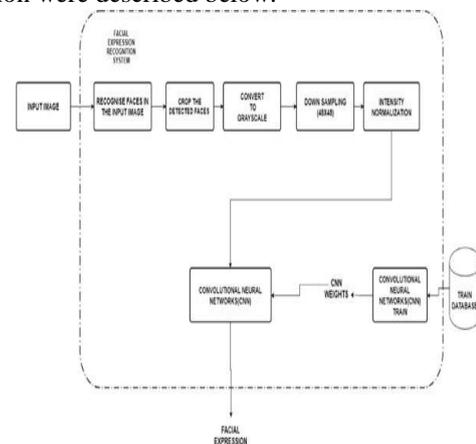
resizing, normalization, scaling , mean, standard deviation, rotation, edge detection. . We compare the accuracy of each pre-processing methods and the combination between them, then analyse it in order to see the variability of accuracies.

II. RESEARCH RELATED WORK

According to F. Defooza, Thiago olinetra- fantos research work with Title: Facial Expression Recognition with convolution neural networks which has a single solution for facial expression recognition that uses a combination of CNN and specific image pre processing steps. Cnn achive better accuracy with big data.The experiment employed to evaluate our technique were contained out using three lagely used public databases.(ck+, JAFE and BU-3DFE) Ck+ data base is fast to train and it allow for real time facial expression recognition with standard computers. According to Tarik A. Rashid research work with technique for recognizing facial expressions using different imbalanced data sets of facial expression is presented. The data is pre processed then balanced next, the significant features are used as input to a classifier model four main classifiers models are selected namely. Decission Tree (DT), Multi-layer perception (MLP),Convolution Neural Network (CNN).The convolution neural network is determined to produce the best recognition accuracy. By performing three two papers we identified CNN technique for prediction of facial expression ofter performing pre processing to the digital image and identifying the actual expression of the precession in the digital image.

III. PROPOSED METHOD

In this work, feature extraction and classification of 7 basic emotions: angry, happy, disgust, fear, sad, and surprise applied on convolution neural network (CNN). Some Preprocessing steps like resize, normalization, scaling, mean, standard deviation, rotation, edge detection. About the details explanation were described below.



Revised Version Manuscript Received on 16 September, 2019.

K.Sravanthi, Asst.Professor, Dr.Br .Ambedkar University, Etcherla, Srikakulam, Andhrapradesh, India.

Dr. G.Jaya Suma, HOD Information TechnologyJntuk,Vizianagaram, Andhrapradesh, India.

IV. DATA SETS

The experiments were performed using posed dataset JAFFE, which mean every subject was given instructions by expert to demonstra

Table 1. The JAFFE datasets

	JAFFE
Label	7 emotions
The Number	100
Participants	100 (male female)
Resolution	48/48 pixel array
Format	.tiff

te 6 emotions. Consists of 100 participants were 18-50 years . It is widely used for emotion recognition and facial expression analysis. Further more, The male and Female Facial Expression (JAFFE) has 100 images that represent 7 basic emotions from male and females .

We have taken the data set from one of the Kaggle competitions. The data we are dealing with is crowd sourced and hence vary widely. Some are cartoon images, some are front facing, some are side facing or inclined and in some rare cases the images are blacked out. The data consists of 48 x 48 pixel gray scale images of faces. The faces are more or less face centered and occupies about the same amount of space in each image. Our ultimate goal is to categorize each face based on the emotion shown in the facial expression in to one of seven categories. The file "train.csv" contains two columns, "emotion" and "pixels". The "emotion" column contains a numeric code ranging from 0 to 6, with both ends inclusive, for the emotion that is present in the image. The "pixels" ` column contains a string surrounded in quotes for each image. The contents of this string a space-separated pixel values in row major order. This string can be looked at as an array of 2304 pixel values arranged in an order. The file "test.csv" contains only the "pixels" column and your task is to predict the "emotion" column, which varies from 0-6 as categorized above.

V. PREPROCESSING

pre-processing steps such as resize, normalization, scaling, mean, standard deviation, rotation, edge detection . Image pre-processing is the study of algorithm that takes an image as input and outputs vector(features) or image(matrix) as output. We might do the following like image enhancement, feature detection, image compression. We do this kind of pre-processing to extract the hidden information or not to loose any information or to remove noise or to normalise the pixel data. About the details explanation were described below.

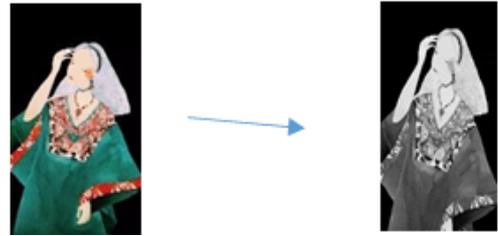
Resize:

Consider a image will be adjusted to determined width, height and converted to gray scale image.

```
Img = cv2.imread('face.jpg',0)
Img1=imutils.resize(n,width=500,heig ht=500
```

Out put:

before resizing (1920, 1080)
after resizing (888, 500)



Scaling:

While scaling an image, a new image with a higher or lower number of pixels must be generated.
Img = cv2.imread('face.jpg',0)
Res=cv2.resize(img,none,fx=0.4,fy=0.4,interpolation=cv2.INTER_CUBIC)



Output:
Before scaling(1920,1080) After scaling(768,432)

Mean:

This method is used to reduce the amount of intensity variation between one pixel and the other. It is often used to reduce noise in image.

```
Img = cv2.imread('face.jpg',0)
Print(np.mean(img,axis=0) Output:
[15.23489583 15.34635417 15.74166667.....6.0015625
5.97395833 5.82864583]
```

Standard deviation:

When we process an image with standard deviation , it will show the variance which is nearer or farther to the mean value of the given image.

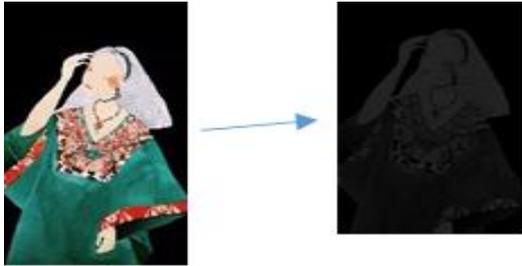


Output:
Before standard deviation(1920,1080) After standard deviation(888,500) Squeezed text(84 lines)

Normalization:

It is a process that changes the range of pixel intensity values in images with poor contrast due to glare. It is also called contrast stretching.

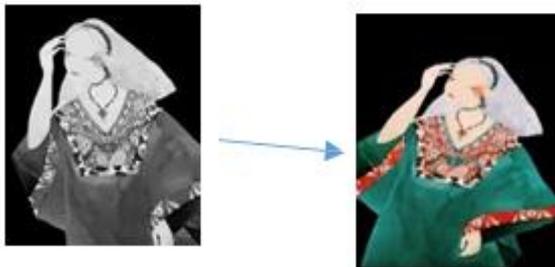
```
Img = cv2.imread('face.jpg',0)
NormalizedImg=np.zeros(800,800))
NormalizedImg=cv2.normalize(img1,
normalizedImg,0,40,cv2.NORM_MIN MAX)
```



Output:
Before normalizing(1920,1080) [[000....000]
[000....000]
[000....000]
..... [000.....000] [000.....000] [000.....000]]0
After normalizing(888,500)

Rotation:

Images can be rotated to any degree clockwise or anti clockwise. We just need to define rotation point, degree of rotation and the scaling factor.

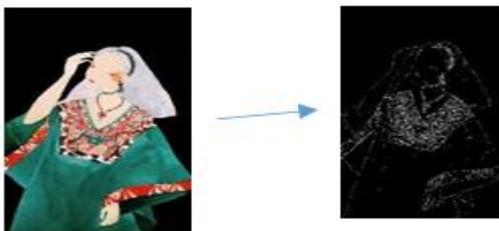


Output:
Before rotation(1920,1080) After rotation(888,500)

Edge detection:

The process of image detection involves detecting sharp edges in the image. This edge detection is essential in object detection.

```
Img = cv2.imread('face.jpg',0) Edges
=cv2.normalize(img,100,200)
Img1=imutils.resize(n,width=500,height=500)
```

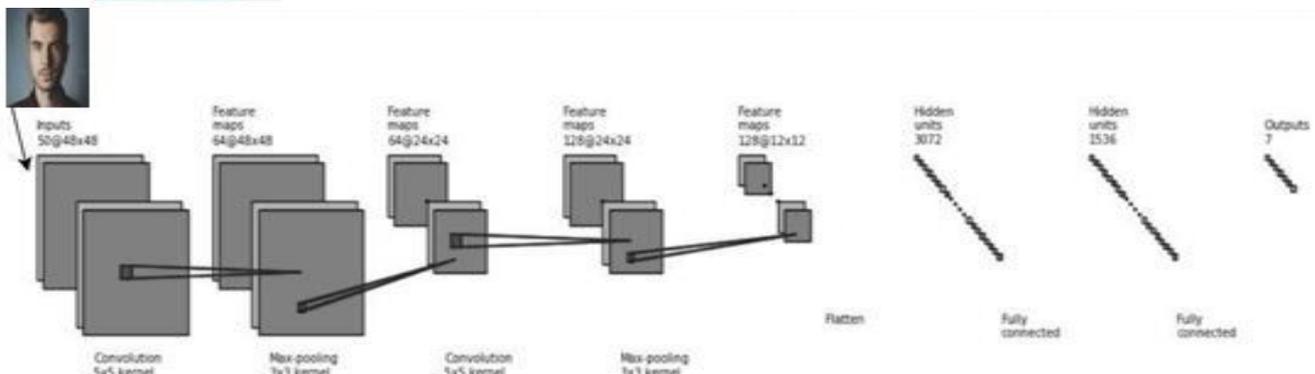


Output:
Before edge detection(1920,1080) After edge detection(888,500)

VI. CONVOLUTION NEURAL NETWORK

In deep learning a convolution neural network (CNN) is a class of deep neural methods, most commonly applied to analyzing visual imagery. CNN use a variation multilayer perceptrons designed to require minimal pre processing they are also knowing shift invariant or space invariant artificial neural networks (SIANN) based on their shared weight architecture and transcation invariance characteristics

.CNN use relatively little pre processing compared to other image classification algorithms. This means that the network learns the filters that the traditional algorithms were hand-engineered. A covolution neural network consider an input and an output layer as well as multiple hidden layers the hidden layers of a CNN typically consists of a convolution layers, RELU layer ie activation function pooling layers, fully connected layer and normalization layers. Description of the process as a convolution in neural network is by convolution mathematically it is a correlation rather than a convolution. 'Convolution layer' apply a convolution operation to the input, passing the result to the next layer. Each convolution neuron process data only for its respective field although fully connected feed forword neural networks can be used to learn feature as well as classify data. Convolution layer brings a solution to a many large input size associated with images, where each pixel is a selected variable. For this CNN brings a solution to this problem as it reduce the number of face parameters, allowing the network to be develop with few parameters by back propagation. 'pooling' convolution network may include local or global pooling layers. Polling layers reduce the dimensions of the data by combining the outputs of neuran clusters at one layer into a single neuran 'Fully connected' layers connect every neuron in one layer to every neuron in another layer. It is in principle the same as traditional multi-layer perception neural network goes through a fully connected layer to classify the image. The input area of a neuron is called its respective field. Learning in a neural network program by making adjustment the biases and weight. The vector of weight and the bias are called a filter and represents some feature of the input.

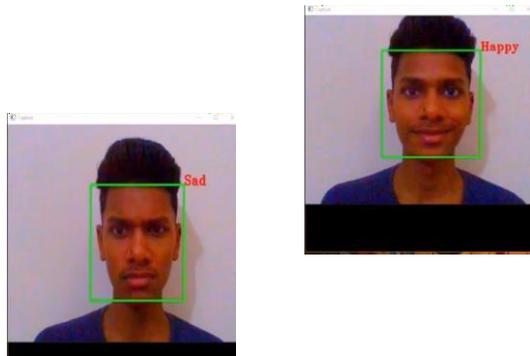


VII. EXPERIMENTAL RESULTS

Pre-processing step and CNN were performed using python libraries such Open CV and Tensor Flow on GPU NVIDIA version 375.74 from nvidia-375, processor type 16xi7- 5960X, memory 65 GB, and Ubuntu 16.04 as operating system. The scenario of this research is performed using JAFFE dataset. We applied pre-processing steps such on facial image such as resize, normalization, scaling, mean, standard deviation, rotation, edge detection.

The most discriminable emotion that can be recognized by CNN are anger, happy, and surprise. CNN gained 100% of accuracy for those emotions. Sadness is the most difficult emotion to recognize since many misclassified of sadness into anger and fear. it means, CNN is hard to distinguish sadness due sadness looks similar to anger and fear.

Identifying weather the captured image is happy or sad.



VIII. CONCLUSION

Two main contributions are presented in this paper. One contribution is that the proposed pre-processing method can assist the CNN model to gain the higher accuracy rate in the applications of facial image processing. The proposed CNN model works better on 32x32 and 64x64 resolution. It seems the capacity of model satisfies the complexity task for facial expression recognition on those resolution. We can boost the performance of CNN using data augmentation like combining data from step (b) cropping and (f) adding noises. The feature work involves exploring image synthesis techniques that may be considered as solution of augmentation data in deep learning. It aims to prevent data starvation and overfitting for small amount of data.

REFERENCES

1. Pantic M, Cowie R, D'Érrico F, Heylen D, Mehu M, Pelachaud C, et al. Social Signal Processing: The Research Agenda. In Moeslund TB, Hilton A, Krüger V, Sigal L, editors. Visual Analysis of Humans: Looking at People. London: Springer London; 2011. p. 511-538.
2. Ekman P. An argument for basic emotions. Cognition and Emotion. 1992;; p. 169-200.
3. Tian Y, Kanade T, Cohn JF. Facial Expression Recognition. In Li SZ, Jain AK, editors. Handbook of Face Recognition. London: Springer London; 2011. p. 487- 519.
4. Chen XW, Lin X. Big Data Deep Learning: Challenges and Perspectives. IEEE Access. 2014; 2: p. 514-525.
5. Gudi A, Tasli HE, den Uyl TM, Maroulis A. Deep learning based FACS Action Unit occurrence and intensity estimation. In 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG); 2015 May. p. 1- 5.
6. Khorrami P, Paine TL, Huang TS. Do Deep Neural Networks Learn Facial Action Units When Doing Expression Recognition? CoRR. 2015; abs/1510.02969.
7. Mollahosseini A, Chan D, Mahoor MH. Going deeper in facial

- expression recognition using deep neural networks. In 2016 IEEE Winter Conference on Applications of Computer Vision (WACV); 2016 March. p. 1-10.
8. Lopes AT, de Aguiar E, Souza AFD, Oliveira-Santos T. Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order. Pattern Recognition. 2017; 61: p. 610-628.
9. Shan C, Gong S, McOwan PW. Facial expression recognition based on Local Binary Patterns: A comprehensive study. Image and Vision Computing. 2009; 27: p. 803-816.
10. Liliana DY, Widyanto MR, Basaruddin T. Human emotion recognition based on active appearance model and semi- supervised fuzzy Cmeans. In 2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS); 2016 Oct. p. 439-445.
11. Jeon J, Park JC, Jo Y, Nam C, Bae KH, Hwang Y, et al. A Real-time Facial Expression Recognizer Using Deep Neural Network. In Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication; 2016; New York, NY, USA:ACM. p. 94:1--94:4.
12. Duncan D, Shine G, English C. Facial Emotion Recognition in Real Time. 2016.
13. Ouellet S. Real-time emotion recognition for gaming using deep convolutional network features. CoRR. 2014; abs/1408.3750.
14. Mayya V, Pai RM, Pai MMM. Automatic Facial Expression Recognition Using DCNN. Procedia Computer Science. 2016; 93: p. 453-461.
15. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion- specified expression. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops; 2010 June. p. 94-101.
16. Lyons M, Akamatsu S, Kamachi M, GyobaJ. Coding facial expressions with Gabor wavelets. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition; 1998 Apr. p. 200-205.
17. Aifanti N, Papachristou C, Delopoulos A. The MUG facial expression database. In 11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10; 2010 April. p. 1-4.
18. Goodfellow I, Bengio Y, Courville A. Deep Learning: MIT Press; 2016.
19. Gonzales RC, Woods RE. Digital Image Processing Third Edition New Jersey: Pearson; 2010.
20. Vidyasaraswathi HN, Hanumantharaju MC. Review of Various Histogram Based Medical Image Enhancement Techniques. In Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015); 2015; New York, NY, USA: ACM. p. 48:1--48:6.
21. Goodfellow IJ, Warde-farley D, Mirza M, Courville A, Bengio Y. Maxout networks. In In ICML; 2013.
22. Sage D. Biomedical Imaging Group. [Online].; 2002 [cited 2017 08 10. Available from: <http://bigwww.epfl.ch/sage/soft/localnormalization/>.