Facial Expression Recognition Using Deep Convolution Neural Network With Tensorflow

Mayur Tembhurney[1], Akshay Chapekar[2], Shubham Funde[3], Prashant Shende[4]

[1] *Assistant Professor, Department of Electronics and Telecommunication Engineering, S B Jain Institute of Technology, Maharashtra, India*
[2] *Student, Department of Electronics and Telecommunication Engineering, S B Jain Institute of Technology, Maharashtra, India*
[3] *Student, Department of Electronics and Telecommunication Engineering, S B Jain Institute of Technology, Maharashtra, India*
[4] *Student, Department of Electronics and Telecommunication Engineering,S B Jain Institute of Technology, Maharashtra, India*

ABSTRACT

Facial expression recognition (FER) is an automatic system that manipulates the facial data and plays a vital role in human machine interfaces. Olden machine learning algorithms has attracted incrementing attention from researchers since the early nineties.
It approaches often requires complex feature extraction process. In this paper we reflected recent advances in deep learning to Convolutional Neural Networks (CNN). It is an prominent field which uses nowadays applications such as in robots, games and neuromarketing.
It is widely used technique uses facial expressions, eye movement and gestures which conveys the emotional status and feelings of persons. The proposed model made pivot on detecting the facial expressions of an individual from a single image. The number of parameters in our proposed networks concentrated by decreasing manner that accelerates the total performance speed and makes this well and make suitable for real time systems

Keyword: - *Convolution Neural Network (CNN) , Facial Expression, Tensorflow, Machine learning, CNN model architecture.*

1. Introduction

Facial Expressions plays an important role in interpersonal communication. The theory of facial expression comes back into Darwin's research on evolution of the species which it appeared as a shape of nonverbal communication. Automatic recognition of facial expression plays an important role in artificial intelligence and robotics and thus it is a need of the generation. Ever since computers were invented, people have build AI systems that are physically and mentally equal to any person. Companies have been experimenting with combining sophisticated algorithms with image processing techniques that have emerged in the past ten years to understand more about what an image or a video of a person's face tells us about how he/she is feeling and not just that but also showing the probabilities of mixed emotions a face could has.

1.1 Scope of the Project

In this project FER system is planned by using convolution neural network. Some application related to this include Personal identification and Access control, Videophone and Teleconferencing, Forensic application, Human-Computer Interaction, Automated Surveillance, Cosmetology and so on. The main aim of this project is to produce Automatic Facial Expression Recognition System which can take human facial images having some expressions as input and detect and classify it into seven different expression class such as : Neutral, angry, disgust, fear, happy, sadness, surprise Humans have always had the innate ability to recognize and distinguish between faces. Now computers are able to do the same. This opens up tons of applications. Face detection recognition used to improve accessibility and security, it allow payments to be processed and allow personalized healthcare and other services. It verifies if two faces are same. The use of facial recognition is huge in security, bio-metrics, entertainment, personal safety, etc.. Our testing showed it had good performance.The steps involved in facial recognition are :

1. Find face in an image
2. Analyze facial feature
3. Compare features for the 2 input faces
4. Returns True if matched or else False.

1.2 Literature Review

Two different perspectives are second hand for facial expression recognition, both of which include two different systematic approaches. Dividing the face into separate action units or keeping it as a whole for further processing appears to be the first and the primary distinction between the main approaches.face emotion recognition system contains two step process i.e. face detection (bounded face) in image represented by emotion detection on detected bounded face. The following two techniques are used for respective given tasks in face recognition system :

1. Haar feature-based cascade classifiers : It detects frontal face in an image well. It is efficient and real time in comparison to other face detector. This blog-post uses an implementation from Open-CV

2. Xception CNN Model (Mini_Xception, 2017) : We will train a classification CNN model architecture which takes bounded face (48*48 pixels) as input and predicts probabilities of 7 emotions in the output layer.

1.3 Proposed Methodology

At training time, the system get collected a training data comprising grayscale images of faces with their proper expression label and programs a set of weights for network. Thereafter, an intensity normalization is applied to the image. The normalized images are used to train the Convolutional Network. Facial Expression Recognition is a challenge for machine learning researchers because the statistical models used in machine learning require a large amount of training data to become accurate. The output of the training step is a set of weights that achieve the best result with the training data. The photos that make up the dataset predominately determine the accuracy and bias of the model. A predictive model can learn to be biased, if there is an over represented class in the dataset. For this reason, the author chose to represent each emotion evenly, even though most datasets had a surplus of neutral and happiness examples. It's output is a single number that represents one of the seven basic expressions.

Figure 3 (a): Training Phase

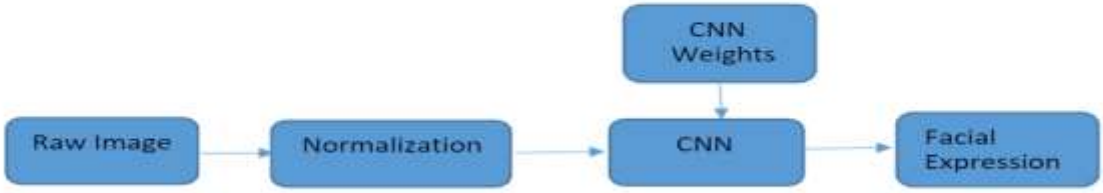


Figure 3 (b): Testing Phase

2. Data-Set

The dataset from a Kaggle Facial Expression Recognition Challenge (FER2013) is used for the training and testing. It has pre-cropped, 48-by-48-pixel grayscale images of faces each labelled with 7 emotion classes : anger, disgust, fear, happiness, sadness, surprise and neutral. Dataset contains training set of 35887 facial images with facial

expression labels. since some classes have large number of examples and some of has few. The dataset is balanced using oversampling, by increasing numbers in minority classes. The balanced dataset contains 40263 images, from which 29263 images are used for training, 6000 images are used for testing, and 5000 images are used for validation The data consists of 48×48 pixel gray scale images of faces. The task is to categorize each face based on the emotion shown in the facial expression in to one of seven categories The training set consists of 35,888 examples

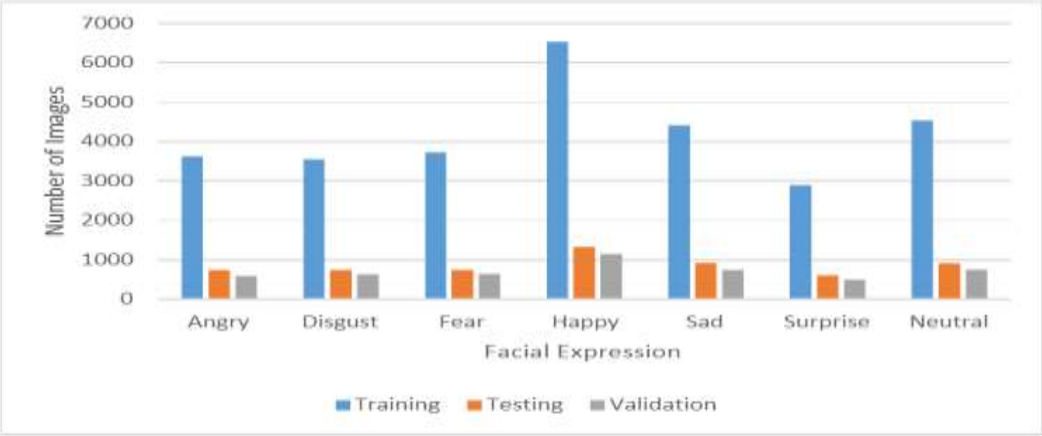


Figure 3.1: Training, Testing and Validation Data distribution

3. Loading FER Data-Set

There are two definitions in the code snippet here:

1. def load_fer2013 : It reads the csv file and convert pixel sequence of each row in image of dimension 48*48. It returns faces and emotion labels.

2. def pre process input : It is a standard way to pre-process images by scaling them between -1 to 1. Images is scaled to [0,1] by dividing it by 255. Further, subtraction by 0.5 and multiplication by 2 changes the range to [-1,1]. [-1,1] has been found a better range for neural network models in computer vision problems.

There are various techniques that can be kept in mind while building a deep neural network and is applicable in most of the computer vision problems. following are some of those methods which are used while training the CNN model :

1. **Data Augmentation** :

   More data is generated using the training set by applying transformations. It is required if the training set is not sufficient enough to learn representation. The image data is generated by transforming the actual training images by rotation, crop, shifts, shear, zoom, flip, reflection, normalization etc.

2. **Kernel regularizer** :

   It allows to apply penalties on layer parameters during optimization. Argument in convolution layer is nothing but L2 regularisation of the weights. This penalizes peaky weights and makes sure that all the inputs are considered.

3. **Batch Normalization** :

   It normalizes the activation of the previous layer at each batch, i.e. installs the transformation that stabilizes the mean activation close to 0 and the activation standard deviation close to 1. It addresses the problem of

internal covariate shift. It also used as a regularizer, some cases terminating the need for Dropout. It helps in speed up the training process.

4. **Global Average Pooling** :

   It reduces each feature map into a scalar value by taking the average over all
   elements in the feature map. The average process that forces the network to extract global features from the input image.

5. **Depth wise Separable Convolution** :

   These convolutions are composed of two different layers: depth-wise convolutions and point-wise convolutions. Depth-wise separable convolutions reduces the computation with respect to the standard convolutions by reducing the number of parameters.

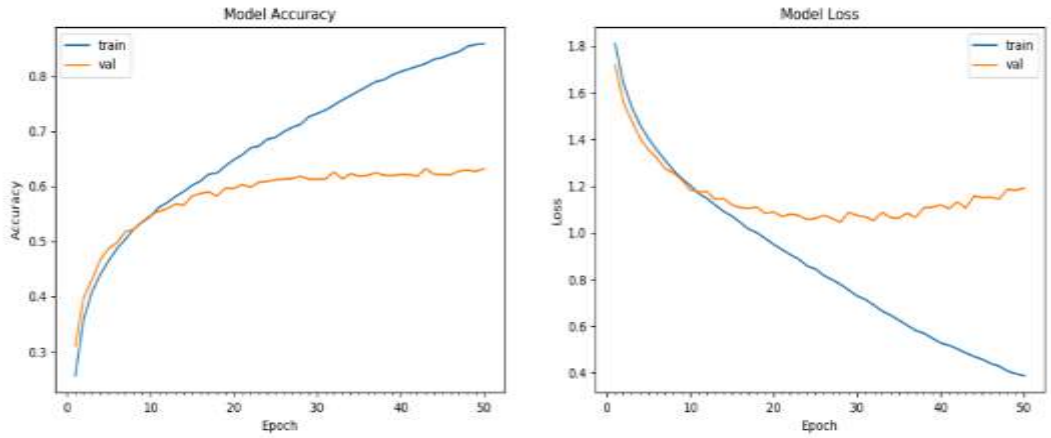
4. Algorithm

- First, we use **haar cascade** to detect faces in each frame of the webcam feed.
- The region of image containing the face is resized to **48x48** and is passed as input to the Convolutional Network.
- The network outputs a list of **softmax scores** for the seven classes.
- The emotion with maximum score is displayed on the screen.

5. Usage

The repository is currently compatible with tensorflow-2.0 and makes use of the Keras API using the tensorflow and keras library.

- First, clone the repository with git clone https://github.com/atulapra/Emotion-detection.git and enter the cloned folder: cd Emotion-detection.
- Download the FER-2013 dataset from here and unzip it inside the Tensorflow folder. This will create the folder data.
- If you want to train this model or train after making changes to the model, use python emotions.py --mode train.
- If you want to view the predictions without training again, you can download my pre-trained model (model.h5) from here and then run python emotions.py --mode display.
- The folder structure is of the form :
  - data (folder)
  - emotions.py (file)
  - haarcascade_frontalface_default.xml (file)
  - model.h5 (file)
- This implementation by default detects emotions on all faces in the webcam feed.



6. CONCLUSIONS

This paper puts a new substructure for facial expression recognition using an additional convolutional network. We believe attention is an important part for detecting facial expressions which provides neural networks with less than 10 layers to fulfil much deeper networks for emotion detection. Along with we also provided an extensive experimental analysis of our work on four FER databases and presented promising outputs also we have added a visualization techniques to highlight the salient regions of face images which are most important subjects in detecting different facial expressions.

7. REFERENCES

[1]. Shan, C., Gong, S., & Mc Owan, P. W. (2005, September). Robust facial expression recognition using local binary patterns. In Image Processing, 2005. ICIP 2005. IEEE International Conference on (Vol. 2, pp. II-370). IEEE.

[2]. Chibelushi, C. C., & Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. CV online: On-Line Compendium of Computer Vision, 9.

[3]. "Convolutional Neural Networks (LeNet) – Deep Learning 0.1 documentation". Deep Learning 0.1. LISA Lab. Retrieved 31 August 2013.

[4]. Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda (2003). "Subject independent facial expression recognition with robust face detection using a convolutional neural network" (PDF). Neural Networks. 16 (5): 555559.doi:10.1016/S0893-6080(03)00115-1. Retrieved 17 November 2013.

[5]. Le Cun, Yann. "LeNet-5, convolutional neural networks". Retrieved 16 November 2013

[6]. C. Zor, "Facial expression recognition," Master's thesis, University of Surrey, Guildford, 2008.

[7].https://appliedmachinelearning.blog/2018/11/28/demonstration-of-facial-emotion-recognition-on-real-time-video-using-cnn-python-keras/

[8]. https://github.com/atulapra/Emotion-detection.git

[9]. "Challenges in Representation Learning: A report on three machine learning contests." I Goodfellow, D Erhan, PL Carrier, A Courville, M Mirza, B Hamner, W Cukierski, Y Tang, DH Lee, Y Zhou, C Ramaiah, F Feng, R Li,X Wang, D Athanasakis, J Shawe-Taylor, M Milakov, J Park, R Ionescu, M Popescu, C Grozea, J Bergstra, J Xie, L Romaszko, B Xu, Z Chuang,and Y. Bengio. arXiv 2013.

[10].https://www.researchgate.net/publication/322103128_Facial_Expression_Recognition_System_using_Convolutional_Neural_Network