March-April 2020
ISSN: 0193-4120 Page No. 5948 - 5950

Facial Expression Recognition Using CNN

T. Pradeep Kumar[1], M. Raja Suguna[2]
[1]UG Scholar, Department of Computer Science, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai
[2]Assistant Professor, Department of Computer Science, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai
pradeepkumar.vlr98@gmail.com, suguna.raj89@gmail.com

Article Info
Volume 83
Page Number: 5948 - 5950
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 31 March 2020*

Abstract

Facial expressions is non-verbal way of communication among humans through face-face interaction. In recent decades Automatic Expression Detection in an Image or a Video plays an important role due to their drastic application such as Detection of emotions in Psychological Analysis, Cyber Forensics, Emotions Detection in Streaming or stored video for behavioral Analysis. Machine learning Algorithm other than the Neural Network requires complex Feature Extraction Phase followed by the Classification of Emotions, whereas in Classification the accuracy of the model is also less. Traditional Feature Extraction during training phase is also more time consuming. To overcome the difficulties of traditional Approach, Deep Learning Approach called Convolution Neural Network is used to Detection Facial Expression. In Convolution Neural Network, the feature extraction is done by training the Network with the large Facial Expression Image Dataset. The Accuracy of Classification also outperforms than the traditional Machine Learning Algorithms and reaches the States of Art method. Kaggle Facial recognitionFERC-2013 dataset is used.

Key Words: *Facial Expressions, Convolution Neural Network, Emotions Detection, Deep Learning Method*

1. Introduction

Facial expressions are one of the most important features to reflect the human emotional state because they convey useful information to the observer. Facial expressions convey 55% of a communicated message, which is more than the part conveyed by the combination of voice and language. Facial expressions can be divided into six basic categories, namely anger, disgust, fear, happiness, sadness, and surprise. With the development of human–computer interaction systems, such as social robots, visual-interactive games, and data-driven animation, facial expression recognition (FER) has become a popular field of study in recent years.

2. Literature Survey

Mayya et al. [1] used DCNN Features to recognize the Facial Expression. DCNN Architecture from Image net [2] compute the Facial Features of vector having 9216 dimensions. They used these higher dimensional Features for validating the SVM Classifier to recognize the facial Expression and Emotions. They have implemented their Machine Learning Model in the following datasets JAFFE and CK + with the accuracy of 8.12% and 96.02% for predicting Seven Classes of Expression. The training time for the network is high and it is end-end solutions.

Wen et al. [3] implemented a method to reduce training time of the Classifier. Trained the Ensemble of Convolutional Neural Network. They randomly generated Hundred Convolutional Neural Network through Random techniques and variety of and they used the fusion of weighted majority vote, probabilistic fusion and the majority vote. This reduces the training time but requires more computational Resources.

Zhang et al. [37] used Soft-max function to train their Convolutional Neural Network. The same Expression variations are reduced and the expressions with large variations are saved for analysis. For a single Smile or sad Emotions, 4000 images are trained by CNN respectively. The dataset for training the network should be very large do get the accurate results.

3. Facial Expressions

A facial expression is moves or places of atleast one muscle groups under the facial skin. Facial gestures are the one of the information conveyer in the non-verbal form. They are fundamental method to provide social information. Even a infant of only 36 hours old can understand the facial emotion to some extent. Even though the humans belonging to different geographical regions, race and regions the expressions produced by them using the skin movements are universal. Hence studying and developing techniques for facial Expression detection will be a very good aspect of increasing Human-Computer Interactions.

4. Proposed Method

The Proposed Method is Single Convolutional Neural Network Classifier. The Input Image should undergo Preprocessing techniques because of varying color Channels and races among the human beings. The Training dataset size is further increased by including images with horizontal flipping and any random rotations which is called Data Augmentations. The Preprocessing of Image consists of the following functions such as cropping not useful region in face, Z-score and Histogram Standardization, and down sampling standardization. The Normalized test images are send as a input to the trained CNN Model to detect the facial Expression.

1) Data Preprocessing-Normalization: It is carried out in two phases, one Normalizing all images in the dataset ,by subtracting the mean value of the image from all pixels in the image and the standard deviation is set as 3.12.the second one is Normalizing training data per pixel in which the standard deviation of all pixel for the training data set is set to 1.

2) Data augmentation: In order to increase the size of training dataset, the images are modified by the following methods and added as the new set images to the training image. The Data augmentation techniques includes mirroring the image with probable value, horizontally rotating the image by the angle of 45 degree, rescale the image to the new scale and randomly cropping the image with the previous original size.

3) Training: In Training the Convolutional Neural Network, Back Propagation algorithm is used Gaussian Distribution is used to initialize the weights of the CNN and the bias value are initialized from the Zero. The loss function used is Mini-Batch Gradient Descent Algorithm and the batch size for training the network is taken as 255.Overfitting is avoided by dropout technique in Fully Connected Layers of the CNN. The training totally takes 1200 epochs, where learning rates are decreased by 8 times in the successive epochs.

The original dataset is divided into two partitions Training set and test set. The CNN is trained by training dataset and the model is saved with the highest accuracy. The trained model is validated by the test Image dataset and the final Accuracy is estimated for Facial Expression Recognition. The following section gives the detailed description of the CNN Architecture.

5. Convolutional Neural Network

The CNN has the following layers in Sequential Model. The first layer is Convolution layer with 32 kernel of size of 5 X 5, the second layer is sub-sampling layer to reduce the image size with max pooling of 2 X 2 kernel and with the stride 2, the third layer is another Convolutional Layer with 64 kernels of size 5 X 5,and the fourth layer is sab-sampling layer of size 2 x 2 and stride 2,the activation function used is Relu in the Convolution Layer. The output from the second Sub-sampling layer is given as a input to the Output layer with the flattened vector of 1600 dimensions and the soft-max activation function with the loss functions and the other parameters are described in the previous section. Fig 1 gives the architecture of Convolutional Neural Network.

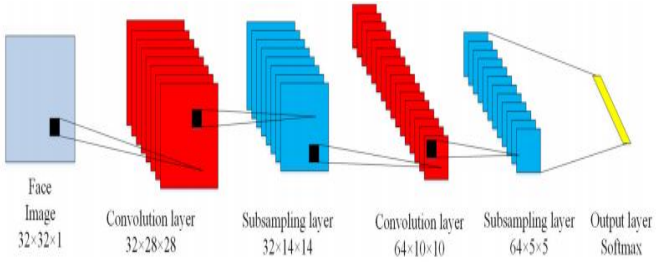

Figure 1: Architecture of Convolution Neural Network

The Model is evaluated by the metrics such as Error rate, Precision, Recall and F-score and the accuracy of the neural network is estimated from these evaluation metrics. Fig 1 shows the Accuracy of the model at the various epochs with the different face Images with Back ground and the Images without Background and Images with the Forehead.

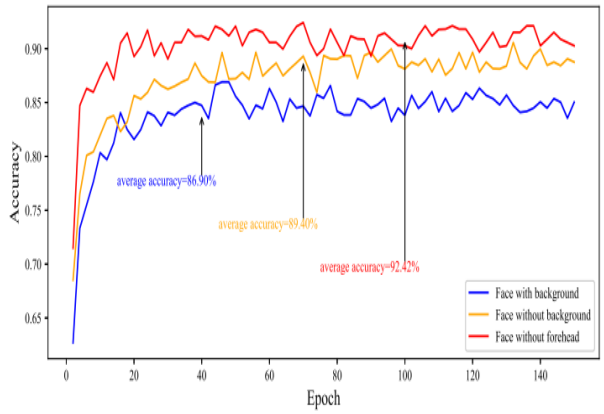

Figure 2: Accuracy of the Model

6. Conclusion

The facial expression is recognized in real-time and varying conditions. Facial Expression Reorganization test datasets has been built on varying facial expression conditions to run on the Convolutional Neural Network. The Proposed method is evaluated for the dataset built and shows the state-of-the-art output compared to the other techniques in the literature.

References

[1] Mayya, V., Pai, R.M., Pai, M.M.M.: Automatic facial expression recognition using dcnn. Proc. Comput. Sci. 93, 453–461 (2016a).https://doi.org/10.1016/j.procs.2016.07.233.

[2] Prevedello, L.M.; Erdal, B.S.; Ryu, J.L.; Little, K.J.; Demirer, M.; Qian, S.; White, R.D. Automated critical test findings identification and online notification system using artificial intelligence in imaging. Radiology 2017,285, 923–931.

[3] Wen, G., Hou, Z., Li, H., Li, D., Jiang, L., Xun, E.: Ensemble of deepneural networks with probability-based fusion for facial expression recognition. Cogn. Comput. 9(5), 597–610 (2017). https://doi.org/10.1007/s12559-017-9472-6.

[4] Zhang, K., Huang, Y., Wu, H., Wang, L.: Facial smile detection based on deep learning features. In: 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). IEEE, pp 534–538 (2015).