INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 3)
Available online at: www.ijariit.com

Emotion recognition from facial expression

Vinay Kumara R. S.
vinayrs.60413@gmail.com
BNM Institute of Technology, Bangalore, Karnataka

Shraddha P. W.
shraddhaleonax04@gmail.com
BNM Institute of Technology, Bangalore, Karnataka

Ketan Kulkarni
vasudev32095@gmail.com
BNM Institute of Technology, Bangalore, Karnataka

Goutham Y.
gouthamy219@gmail.com
BNM Institute of Technology, Bangalore, Karnataka

Manohar G. M.
manohargm32@gmail.com
BNM Institute of Technology, Bangalore, Karnataka

ABSTRACT

The aim of image (photo) based emotion recognition system emotion is to detect and classify the emotion from the human face from an image. The automatic system mainly has two components that is extraction of facial feature and classification of facial feature. To classify the feature precisely effective feature classification methods should be applied. Many recent systems use additional information from classification which reduces accuracy as well as increase execution time. In this system Double Local Binary Pattern (DLBP) is used for feature extraction which is a variant of Local Binary Pattern (LBP). Since DLBP has small dimension size which reduces the detection time. To handle local illumination problem Logarithmic Laplace-Domain is proposed (LoL-Domain). Finally, Convolution Neural Network (CNN) is used for an obtained feature for classification.

Keywords— *Double Local Binary Pattern, Convolution Neural Network, Emotion recognition, Feature extraction*

1. INTRODUCTION

Facial Emotion Recognition (FER) mainly predicts the emotion from facial expression. (That is neutral, sadness, happiness, surprise, disgust, fear, and angry) from the human face. It can be noted that 60% of communication is contributed by facial expression, 32% by voice, and 8% by language. Since emotion on the face plays a major role in communication and it is major exposed part of the body, it qualifies the computer vision system's usage (usual cameras) for analysing the face from the image for emotion recognizing. Mainly there are two factors that impact the quality of the emotion detection system that is illumination on the face and pose of the face in the image.

As it is obvious that recognition of an emotional state of a person is done immediately by humans, but it is very much necessary to train the computer to detect the emotion of a person at real time. This made the path for researchers to carry out the work on the methods that best fit the system for training purpose. Mainly, recognition of emotion is used in mood-based video players, human-abnormal behaviour detection, and computer-human intelligent interactions like instructor, helper and tutor. Recently, immersive works have been brought out towards the facial extraction of features from the face with the goal of maximizing variability between the class and minimizing variability within the class.

Employing deep learning techniques for effective learning of representing the feature had pioneered a variety of vision task for computers. Some deep neural networks had made a close move with the performance on the task of recognition of human emotions. An enormous amount of well-known problems in the area of computer vision recently have been benefited from the upswing of feature representation and classifiers in deep learning. Human beings day-to-day possess and express emotions while interacting with each other. Emotions are often bounced back on the face, gestured of hand, body, and in the voice, to convey their likings and feelings. Coming to medical applications, the task of recognizing a individual's incapability to express certain facial expressions may help in diagnosing psychological disorders at early stages.

Before going for feature extraction image should be pre-processed such as face detection, face clipping. For detecting the face in the image Viola-Jones algorithm may be used efficiently. Once the face is detected then the background is cropped, further process is carried out.

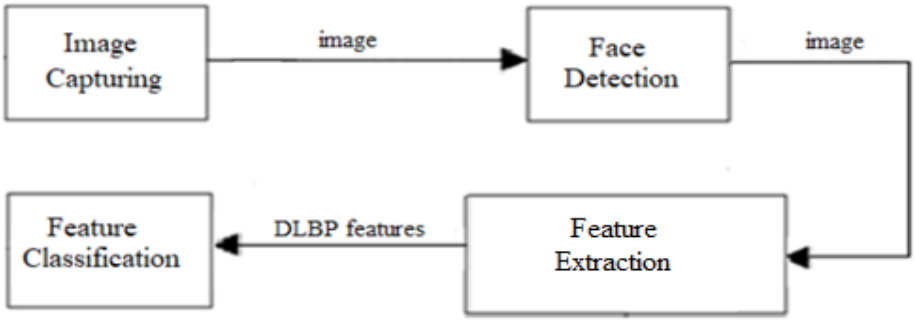


Fig. 1: Overview of the proposed system

Since this system works on a real-time basis image capturing module is been introduced for capturing an image of the person on a real-time basis. Here Double Local Binary Pattern (DLBP) is used for extracting the feature, which is the enhanced version of the Local Binary Pattern (LBP), prior to that local illumination problem should be resolved so as to get accurate results.

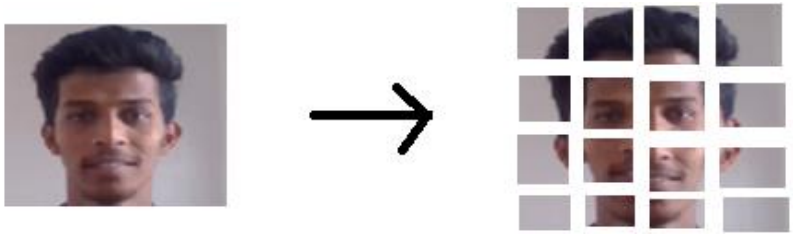
To solve local illumination variation problem Logarithmic Laplace (LoL)-Domain is used and Convolution Neural Network (CNN) is used for classifying the feature for the first time with the DLBP and LoL-Domain methods.

2. RELATED WORK

In real life human will understand the emotion of the person interacting with him/her but coming to the computer system it is difficult to detect emotion from the face of the person unless and until training is carried out. This makes the training process essential. As nowadays there are various techniques for training the system but the proposed system is efficient because of using both efficient methods DLBP and CNN for feature extraction and feature classification. Our system aims at the automatic detection of emotion. Below are the algorithms for feature extraction and feature extraction.

2.1 DLBP Algorithm:

Step 1: An image is divided into N blocks called as Texture Unit (TU) according to the size and resolution of the image.



Here each TU will be a group of nine pixels

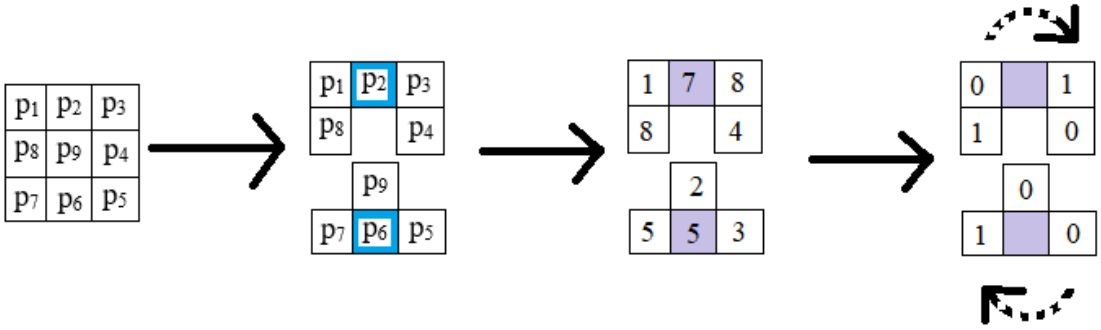
Step 2: DLBP operation is done for each TU

$$DLBP = (DLBP^0, DLBP^1)$$

$$DLBP^0 = \sum_{j=1}^{4} f(p_i - p_c) * 2^j$$

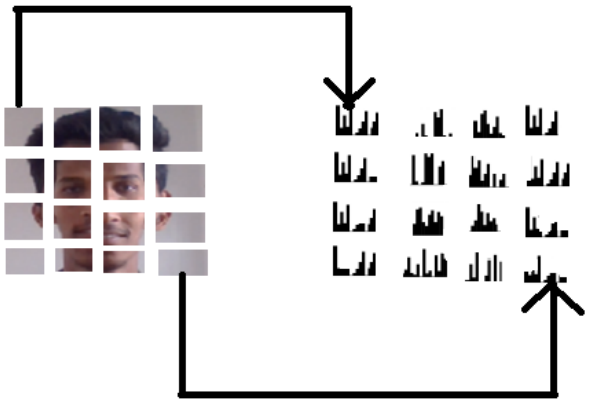
$$DLBP^1 = \sum_{j=1}^{3} f(p_i - p_c) * 2^j$$

$$f(x) = \begin{cases} 1, & \text{if } x >= 0 \\ 2, & \text{if } x < 0 \end{cases}$$

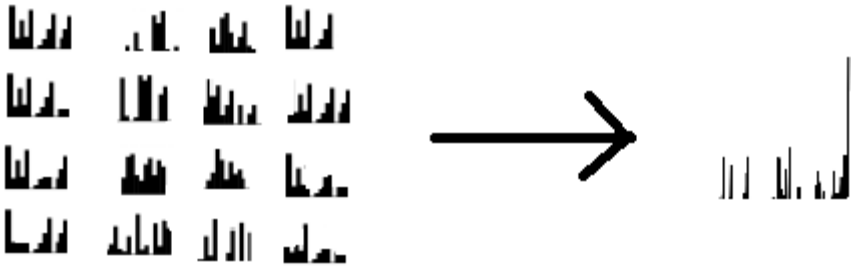


Here '0' represents top pixels and '1' represents bottom pixels.

Step 3: Build Histogram for each TU.



Step 4: Obtained histograms from all TU should combine and make as only one histogram.



After all these steps at last DLBP image is obtained from which features are obtained? Once DLBP image is obtained CNN method is applied for feature classification

Below is the algorithm for CNN method: CNN algorithm mainly has two processes: convolution and sampling.

(a) Convolution process: here trainable filter Fx is used, next deconvolution of the input image is carried out (the first stage is the input image, the input of the after convolution is the feature image of each layer, namely Feature Map), then bias bx is added in order to get convolution layer Cx.

(b) A sampling process: through pooling steps n pixels from each neighborhood, become a pixel, and next bias bx + 1 is added to weight wx+1 which is obtained by scalar weighting, lastly, a narrow n times feature map Sx + 1 is produced by the function called activation function.

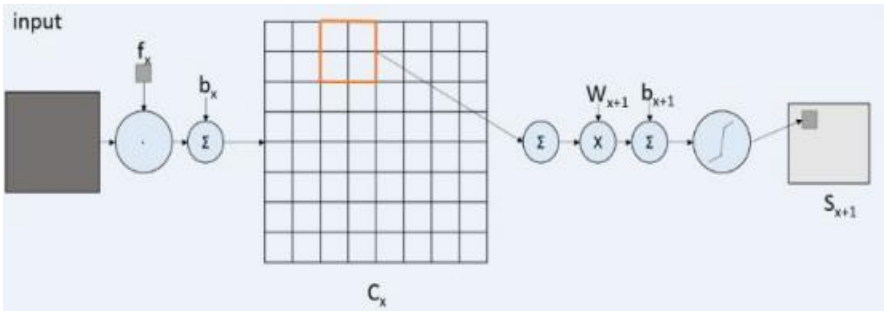


Fig. 2: Main process of CNN

In CNN local receptive field, sharing of weights, sub sampling by time or space are key technologies for extracting feature and reducing the size of the training parameters. The benefits of CNN algorithm are that explicit feature extraction is avoided, and the learning process is carried out implicitly from the training data. Benefits of sub adopting sampling structure by time or space can achieve some amount of robustness, scale and deformation displacement.

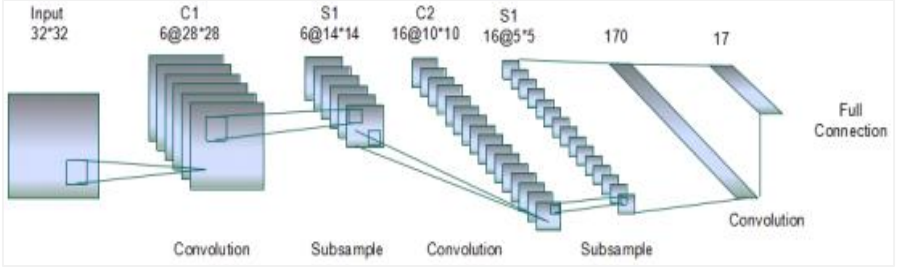


Fig. 3: Architecture of CNN in training faces

CNN algorithm needs to be improved in design of architecture that is need to be trained with a large dataset in the practical application, in order to obtain the most suitable model based on the application. Considering any size of image gray image as the

input, the size of the image is turned to 32 × 32 in the pre-processing stage. Since 7 layered convolution model has more efficiency so this is used, 7 layes of CNN for feature classification are an input layer, convolution layer C1, sub sampling layer S1, convolution layer C2, sampling layer S2, hidden layer H and output layer F.

4. EXPERIMENTAL RESULTS

Below figure shows, the results of conversion from the image captured in real time to DLBP figures. 4 (a) shows the original image. b) Shows the image converted from RGB to gray image. c) Shows the gray image converted to DLBP image that is features).

(a) (b) (c)

Fig. 4: (a) shows the original image (b) Shows the image converted from RGB to gray image (c) Shows the gray image converted to DLBP image

(a) (b) (c) (d) (e)

Fig. 5: (a) Happy (b) Anger (c) Neutral (d) Disgust (e) Surprise expressions as the output

5. CONCLUSION

Facial Emotion Recognition is accomplished using Convolutional Neural Network (CNN). Kaggle dataset is used for classification of emotion in static images as well as for detection of the emotion in the real-time-based video as well as image.

Mainly our system has two key components viz. extraction of facial feature and classification of the extracted. Here DLBP algorithm is used for obtaining the feature of the image (that is face) with high accuracy. DLBP is an effective alternative of the LBP algorithm and in the task of extraction of feature, LBP-based methods have been for a long time. Generally, all methods only deal with global illumination problem which results in poor performance coming to local illumination problem which is not dealt with in any existing methods. To overcome the local illumination problem, LoL domain is proposed in this paper in order to improve the performance as well as accuracy. Finally, CNN is used for classification of the obtained feature so as to get the appropriate emotion.

Future work might be dealt with overcoming of pose variation in efficient manner compared to the existing methods. Since poor pose variation may result in improper classification of emotion.

6. REFERENCES

[1] Wenming Zheng, Yuan Zong, Xiaoyan Zhou and Minghai Xin," Cross-Domain Color Facial Expression Recognition Using Transductive Transfer Subspace Learning", IEEE Trans, vol 9,no1, pg 21-37, Jan-Mar2018.

[2] Yuanyuan Ding, Qin Zhao, Baoqing Li, Xiaobing Yuan," Facial Expression Recognition From Image Sequence Based on LBP and Taylor Expansion", IEEE Access, vol 5, pg 19409-19419, Oct 2017.

[3] P.Ekman," Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique", Psychol.bull, vol no.2, pp.268-287, Mar-2014.

[4] Biao Yang 1, Jinmeng CAO1, RongRong NI2, Yuyu Zhang," Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images", IEEE Access, Vol 6, Feb-2018.

[5] Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, Remigiusz J. Rak," Emotion recognition using facial expressions", International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland.

[6] Y. Tong, R. Chen, and Y. Cheng, ``Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle," Optik-Int.J. Light Electron Opt., vol. 125, no. 16, pp. 4186_4189, 2014.

[7] S.-C. Hsu, C.-T. Hsu, and C.-L. Huang, ``Video-based facial expression recognition using Hough forest," J. Inf. Sci. Eng., vol. 32, no. 2, pp. 477_494, Mar. 2016.

[8] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, ``Dropout: A simple way to prevent neural networks from over_tting," J. Mach. Learn. Res., vol. 15, no. 1, pp. 1929_1958, 2014.

[9] M. R. Mohammadi, E. Fatemizadeh, and M. H. Mahoor, ``PCA-based dictionary building for accurate facial expression recognition via sparse representation," J. Vis. Commun. Image Represent. vol. 25, no. 5, pp. 1082_1092, 2014.

[10] A. R. Rivera, J. R. Castillo, and O. O. Chae, ``Local directional number pattern for face analysis: Face and expression recognition," IEEE Trans. Image Process., vol. 22, no. 5, pp. 1740_1752, May 2013.

[11] Lundqvist D., Flykt A., Öhman A., The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet, no. 1998.

[12] Microsoft Kinect, https://msdn.microsoft.com/en-us/library/jj131033.aspx.

[13] K. Seshadri and M. Savvides, "Towards a unified framework for the pose, expression, and occlusion tolerant automatic facial alignment," IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 10, pp. 2110–2122, Oct. 2016.

[14] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE Trans.Pattern Anal. Mach. Intell., vol. 29, no. 6, pp. 915–928, Jun. 2007.

[15] Z. Li and J. R. Eastman, "Commitment and typicality measures for the self-organizing map," Int. J. Remote Sens., vol. 31, no. 16, pp. 4265–4280, 2010.

[16] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Lake Placid, NY, USA, 2016, pp. 1–10.

