A LITERATURE REVIEW ON REGION BASED ROBUST FACIAL EXPRESSION ANALYSIS

[1] A Jahnavi, [2]Shaik Taj Mahaboob, [3]M Siva Kumar

[1]PG Scholar, [2]Assistant Professor, [3]Assistant Professor (Adhoc)
[1]jahnavijanu520@gmail.com
[1]Electronics and Communication Engineering,
[1]JNTUA College of Engineering Pulivendula, Pulivendula, India

Abstract: Facial emotion recognition is becoming increasingly popular in a progressive research area and it plays a major role in Human-Computer-Interaction. In the real-world conditions, it faces many challenges, i.e., illumination changes, large pose variations and partial or full occlusions, which cause different facial areas with different sharpness and completeness. Inspired by this fact, Based on partial faces facial expression recognition is been focused. Here, we compare contribution of seven facial areas of low resolution images, including nose areas, mouth areas, eyes areas, nose to mouth areas, nose to eyes areas, mouth to eyes areas and the whole face areas. In this, we find that mouth regions contain much emotional information compared with nose areas and eyes areas through analysis on the confusion matrix and the class activation map. In Existing system it doesn't identify exact emotion of a person. In this proposed algorithm we take the larger facial areas, hybrid extraction feature and CNN classification. In this paper we present literature survey to compare the efficiencies with Other Recognition Techniques with different approaches.

Index Terms **- Facial emotion recognition, Facial areas, Class activation map, Confusion matrix.**

I. INTRODUCTION

As with the increase in the usage of artificial intelligence, human-computer dialogue systems has been explored in efficient manner, which have wild applications such as the movie booking [1], chat bots [2] and smart homes. Therefore, emotion recognition, is an essential aspect in human-computer interaction has been received a large amount of attentiveness recently. [3, 4, 5]

Facial expression recognition is a widely important research topic in the emotion recognition, which changes deeply under the ascendancy of deep learning (DL). In spite of a huge amount of attempts are made to improve the recognition performance of facial expression, many challenges still exist. In the real-world surroundings, it's strenuous to gather faces without the shade from other objects. The faces are not always in the frontal pose and the proper light conditions.

With the popularity of DL, [6] proposes to recognize facial expression based on regions of interest, which guides convolutional neural networks (CNNs) to centre on areas associated with the expression. However, it does not analyze the contribution of different facial areas for different emotions. In the meantime, we analyze the contribution of seven facial areas for the robust emotion recognition, including nose areas, mouse areas, eyes areas, nose to mouse areas, nose to eyes areas, mouth to eyes areas and the whole face areas.

In previous, we compare contribution of seven facial areas of low resolution images, including nose areas, mouse areas, eyes areas, nose to mouse areas, nose to eyes areas, mouth to eyes areas and the whole face areas are not recognized clearly and gives the less performance and accuracy. So, to overcome this, we analyze the impact of different facial areas in facial expression recognition through three experiments. First, we analyze emotion classification performance based on the whole faces, treating it as a comparison experiment. Then, we visualize activation parts of the inputs through CAM [7] technique. Finally, we compare classification accuracy and the confusion matrix of seven facial areas.

II. SYSTEM OVERVIEW

In our system, we follow the DenseNet-BC architecture [8], which has three dense blocks associated with the global average pooling (GAP) and the fully-connected layer (FC) behind. The inputs of the system are grey-scale images and the outputs are normalized emotion probabilities. To get prediction of the system, we utilize a weighted sum on the outputs of GAP, which are spatial average of the feature maps generated from the last dense block. In CAM technique, a similar idea is taken. We compute a weighted sum of the feature maps extracted from the output of the last dense block as CAM.

In this, the processing steps are first to take the seven facial area regions of the whole face as input image and applies to dense block. Here, this dense block is based on convolution and carries to each dense block and in each layer, segmentation has done. And here our image is transition into each block. To removing the noise of the image, we are taking the transition and apply to Global average pooling (GAP). Here, this pooling gives the identification of our expression. So, we are using this technique to have the transmission data in order. And here we compute a weighted sum of the feature maps extracted from the output of the last dense block as CAM.

And next fully connected layer (FC) is using for regions identification. So from the GAP, some data has been classified into parts and these parts are divided with intensity value and this divided data is mapped to FC. From this FC classifier, the classification regions are mapped to Class activation map (CAM).

Class activation mapping is adapted to visualize activation parts of different emotions. So, in this features are extracted and calculated. And after that image regions are identified and applied to CNN classification. And finally, it predicts the expression and classifies it and final output is given i.e. happy or sad etc,

III. TECHNIQUES USED FOR FACIAL EXPRESSION ANALYSIS

A. HOG for Human detection

In this paper they explained that the Histograms of Oriented Gradient (HOG) descriptor notably outperforms pervious attributes set used for human detection. In this paper they calculated the impudence of each steps of the computation on the performance, analyzing finally those fine-scale gradients, fine orientation binning, relatively coarse spatial binning, and high-quality local contrast normalization in overlapping descriptor blocks are key aspect for quality results. The new method proposed gives similar and absolute partition on the actual MIT pedestrian database, so they introduce a more arduous dataset incorporates over 1800 annotated human images with a various range of pose variations and backgrounds [9].

B. Dynamic Texture Recognition by LBP

Dynamic texture (DT) is an extension of temporal domain. Description and recognition of DTs have enthralled growing observation. A new appeal for recognizing DTs is introduced and its reduced simplification and extensions to facial image analysis were also contemplated. First, the textures are modeled with volume local binary patterns (VLBP), which are an adjunct of the LBP operator widely used in ordinary texture analysis, combining motion and appearance. To make this proposed method computationally easy and simple to enlarge, only the co-occurrences of the LBP on three orthogonal planes (LBP-TOP) are to be contemplated [10].

C. Recognizing Action Units by Facial Expression Analysis

The changes in the discrete facial features can show the human emotions and intentions. In this paper, they developed an Automatic Face Analysis (AFA) system which is used to audit facial expressions build on transient facial features (deepening of facial furrows) and permanent facial features (brows, eyes, mouth) in a almost frontal-view face image sequence. Multistate face and facial component replica has been proposed for tracking and modeling the various facial features, including eyes, lips, cheeks, brows, and furrows [11].

D. ROI Deep CNN

In this paper they introduced Facial expression recognition (FER) method which is based on region of interesting (ROI) and it shows the convolutional neutral networks (CNN) to focus on the areas connected with the expression. In test stage, they defined two recognition methods: Firstly identify the test image directly; secondly implement the decision fusion strategy on ROI areas. The model we used is fine-tuned from pre-trained deep CNN instead of training from scratch. In addition, we confer an innovative region based image augmentation method named artificial face to improve the limited database. The accomplishment of the proposed method has been vindicating on the CK+ databases [6].

E. Emotional Chatting Machine

Emotional Chatting Machine (ECM) has been proposed in this paper which can generate congruous responses not only in content (relevant and grammatical) but also in emotion (emotionally consistent). This is the primary work that discusses the emotion factor in large-scale conversation generation. ECM discourse the factor using three new methods, firstly it models the high-level abstraction of emotion expressions, secondly it represent the change of implied internal emotion states, and finally it uses precise emotion expressions with an external emotion vocabulary [4].

Table 1
A Summary of Facial Expression Analysis Techniques

Sl. No.	Title	Reference	Method/Algorithm Implemented	Advantages	Disadvantages
1	Histograms of oriented gradients for human detection	N. Dalal, and B. Triggs [9]	Detection can be described by distribution of intensity gradients	Reduces false positive rates	It is not scale and rotation invarient
2	Dynamic texture recognition using LBP with an application to facial expressions	G. Zhao, and M. Pietikainen [10]	LBP-TOP operator is computed from three orthogonal planes.	Computationally simple and robust in terms Grey-scale and rotation variation	Use of facial appearance information is limited
3	Recognizing Action Units for Facial Expression Analysis	J.F. Cohn, T. Kanade, and Y. Tian [11]	AFA is introduced to audit facial expressions	Outperforms explicit parameterization	Large databases cannot be processed
4	Improved Facial Expression Recognition Method Based on ROI DeepCNN	X. Sun, M. Lv, C. Quan, and F. Ren, [6]	Improved facial expression recognition method based on ROI	Recognition based on ROI decision fusion is superior	The distributed representation of ROI area in model memory is wasted
5	ECM: Emotional Conversation Generation with Internal and External Memory	H. Zhou, M. Huang, T. Zhang, X. Zhu, and B. Liu [4]	It models the emotional influence in large-scale conversation generation	Has more emotion accuracy	Performance is less because of loss function

IV. CONCLUSION

In this paper, we have reviewed the framework of Facial emotion recognition and presented a literature survey on the different techniques used in facial recognition system. These FER methods are measured on the basis of accuracy of detecting the emotion, computational complexity and recognition rate.

REFERENCES

[1] Y. Chen, J. Gao, X. Li, and L. Li, "End-to-End Task-Completion Neural Dialogue Systems," CoRR, 2017.

[2] A.D. Brébisson, Y. Bengio, A.P. Chandar, M. Germain, T. Kim, N.R. Ke, Z. Lin, S. Mudumba, V. Michalski, A. Nguyen, M. Pieper, J. Pineau, I. Serban, C. Sankar , S. Subramanian, J. Sotelo, D. Suhubdy, and S. Zhang, "A Deep Reinforcement Learning Chatbot," CoRR, 2017.

[3] N. Asghar, J. Hoey, X. Jiang, L. Mou, and P. Poupart, "Affective Neural Response Generation," CoRR, 2017.

[4] H. Zhou, M. Huang, T. Zhang, X. Zhu, and B. Liu, "Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory," CoRR, 2017.

[5] S. Ghosh, M. Chollet, E. Laksana, L. P. Morency, and S. Scherer, "Affect-LM: A Neural Language Model for Customizable Affective Text Generation," CoRR, 2017.

[6] X. Sun, M. Lv, C. Quan, and F. Ren, "Improved Facial Expression Recognition Method Dased on ROI Deep Convolutional Neural Network," International Conference on Affewctive Computing and Intelligent Interaction (ACII), pp. 256-261, 2017.

[7] Khosla, À. Lapedriza, A. Oliva, A. Torralba, and B. Zhou, "Learning Deep Features for Discriminative Localization," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2921-2929, 2016.

[8] Huang, Z. Liu, and K.Q. Weinberger, "Densely Connected Convolutional Networks," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261-2269, 2017.

[9] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886-893, 2005.

[10] Zhao, and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 915-928, 2007.

[11] J.F. Cohn, T. Kanade, and Y. Tian, "Recognizing Action Units for Facial Expression Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 97-115, 2001.