iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IX **Month of publication:** September 2021

DOI: https://doi.org/10.22214/ijraset.2021.38100

Facial Emotions Recognition Using Deep Learning Technology

Abhinav Chaubey[1], Mr. Vijay Maheshwari[2]
[1]M.Tech. (CS), [2]Under the Guidance, Computer Science Department, "School of Engineering and Technology, Shobhit Institute of Engineering & Technology, Meerut-250110 India"

Abstract: Artificial Intelligent give us capability to detect emotions of human being. Due variation of individual expression it is difficult to find precisely. With AI we can mimics a human's capability like recognising someone with a restricted facial feature. this . paper, of mine indentify the face emotions by detecting areas of face like eyes, nose, lips, and forehead.

By implementing two repressing methods like histogram and data augmentation we propos to extract characteristics of facial emotion. Here in this paper two dimensional architecture is used. First is used for inputting greyscale of face image where as second is for accepting histograms. the final process calculate the result on the bases of KNN and SVM classifiers. The results indicates that proposed algorithm detect six fundamental facial emotions , Happiness, Anger, Fear and surprise. Précised result are expected by using trained model data set.

Keywords: SVM, KNN, FER, DNN, VGG16, HOG, HSOG.

I. INTRODUCTION

From the human face images FER (Facial Expression Recognition) try to indentify face emotions like Happy, sad, Surprise, Fear and Disgust. Just by analyzing the face images the method helps the machine to understand the human emotions. It has gained lot of attention because of its potential applications like, computer interfaces, health management, autonomous driving, detecting abnormal human behaviour and other similar tasks.

Data augmentation with Histogram are pre-processing methods, that are required for the facial images provided to make machine learn from images. Histogram equalisation is very precise , which could make the pattern of gray scale in numerous images uniformly lead to reduce the light generated obstacles. CNN method needs huge data sets to solve particular problem. FER databases which are available publicly have not sufficient images to handle the problems. For creating synthetic images from the original face image, "Simard et al. (Simard et al., 2003) suggested the DA procedure to extend the datasets."

Despite recent rapid developments, FER remains challenging due to some factors such as lighting, head deflection and some occlusions in facial regions. These impedances can affect facial recognition performance and reduce the accuracy of FER. As demonstrated in the past, handcraft features seems to be no longer appropriate for expression recognition activities with critical issues. The DNN: Deep Neural Network proves a precise answer to these problems.

Humans being can recognize emotions a with limited constrain . On account of acknowledgment of facial expressions, the utilization of full face images can be repetitive since facial expression fundamentally misshapes certain specific zones of face images.

There is one algorithm called as facial benchmark detection algorithm which is offered by Dlib helps to extricate the facial regions from the given face image. This Machine learning library offer by "by King (King, 2009)" as an open source.

That gives us 68 landmarks points on the face. Using those landmarks points, we are able to extract the regions of face like forehead, eyes, nose, and lips. Our proposed frame focuses only on these parts of the face, but to verify the efficiency of the proposed frame, we also performed experiments with the complete facial image.

This paper focuses on the double dimension architecture that processes the greyscale and the HSOG (histogram of second-order gradients) face images. HSOG a type of Histogram of oriented gradient, as indicated in "(Dalal and Triggs, 2005)". It extracts local information from the face image. DNNs are used for various channels greyscale and HSOG facial images. In one channel, a proposed VGG16 ft with original parameters acquired as in VGG16, which was trained, is created for greyscale expression. On the other channel, HSOG facial images, a proposed two-layer CNN, which refers to the development of Deep ID (Sun et al., 2015). The output of the two channels are concatenated and made an enormous feature vector.

To detect common facial gestures, SVM along with KNN with calculation of separation for classification like fear, happy, anger etc.

A. "JAFFE Database"
B. " VIDEO Database "

(Shikkenawis and Mitra, 2016), 3. CK+ Database above are utilised to test the framework to demonstrate its feasibility. This modular way is another significant commitment to current work.

First, the dynamically detect facial location from full face.

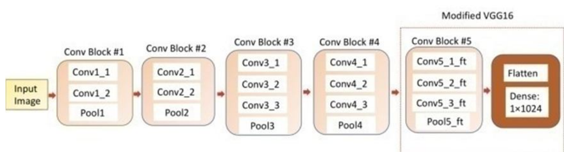
1) Second, two dimensional channels from greyscale with HSOG is extracted for FER Third, Fine tuning by trained VGG16 using DNN .
2) At last, outputs of the two channels are combined to predict a vigorous outcome. Four bench- marking datasets and a few handy facial images are utilized to assess the successfulness of our work. The rest of the article is organized as follows. The 2 section provides details of the proposed framework. The 3 section shows the results and analysis of the experiment. Section 4 Concludes the study.

II. SUGGESTED METHOD FOR FER

This section describes each of these steps in detail.

A. Histogram Equalization

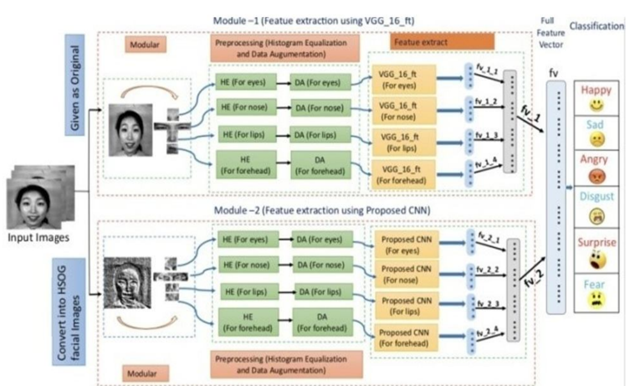
The image captured can have different lighting and shadow which can cause various bright and dark areas hence result in poor face location detection. Therefore, we tend to perform histogram equalization (HE) before recognition. HE is a simple but effective technique in image processing, execution in visual recognition and fast convergence.



III. FEATURE EXTRACTION FROM GREYSCALE

A. Facial Images

The unsatisfactory result from CNN lead to FER approach. data Expanding result in excessive contents. Hence, fine-tuning method is utilized to detect features from the input face through the deep neural network (DNN) which has improved efficiency.
Proposed method utilizes DNN for the extraction of emotion location for FER dependent on the VGG network introduced
This utilizes two versions of VGG: VGG16 and VGG-19.



Feature Extraction from HSOG (Histograms of the Second Order Gradients) Facial Images To the best of our knowledge, there is no model trained on the HSOG images. So here first compute the HSOG facial images.

Calculation of First Order Oriented Gradient Maps

It start with computing the 1st order oriented gradient map. Before initiating calculation it make sure the colour normalisation and accurate gamma value. After this it start calculating First order gradient maps.

B. Computation of Second Order Gradient (OGMs)

After Calculation of 1st OGM the result are feed to 2nd OGM calculation. For every pixel located calculated for OGM magnitude and its orientation by equation below vectors archived from different locations like lip, eye etc are connected in a long vector by VGG16 ft. then it is classified

$$Mag(x,y) = \sqrt{(\frac{\partial G(x,y)}{\partial x})^2 + (\frac{\partial G(x,y)}{\partial y})^2}$$

$$\Phi(x,y) = \arctan(\frac{\partial J_G(x,y)}{\partial y} / \frac{\partial G(x,y)}{\partial x})$$

$$\frac{\partial G(x,y)}{\partial x} = G(x+1,y) - G(x-1,y)$$

$$\frac{\partial G(x,y)}{\partial y} = G(x,y+1) - G(x,y-1)$$

Another classifier with multiple distance measurements is the K nearest neighbour (K=1,2,3) classifier. this paper used the Euclidean distance, the Chi-square distance, and the histogram intersection (HI). The loss computation for the SVM and KNN is done using Mean Square Error (MSE). To support the theoretical finish of the proposed framework, tests have been conducted on some genuine datasets, as archived in this section.

Experiments of FER have been conducted on the four FER datasets. Facial images for the most part incorporate extremely huge dimensions. Managing such broad information ends up being very hard for machines. Hence, the modular methodology is applied when just certain data areas of the face are thought of. Facial expression give signals of the individual emotional state, even without verbal correspondence. The eyes are the most open aspect of an individual face and reveal a lot about their sentiments. Not with standing the eyes, lips, forehead, nose, and so forth they are also information regions. During the expression analysis task, we saw that, in addition to the eyes, nose and lips/mouth, the forehead additionally assumes a significant role regarding expressions. Most FER strategies are presently applied to full face images. This article focuses in just on some information area of the face, as talked about. To make a correlation, we did the holistic experiments (where the full face image was utilized) as well as modular.

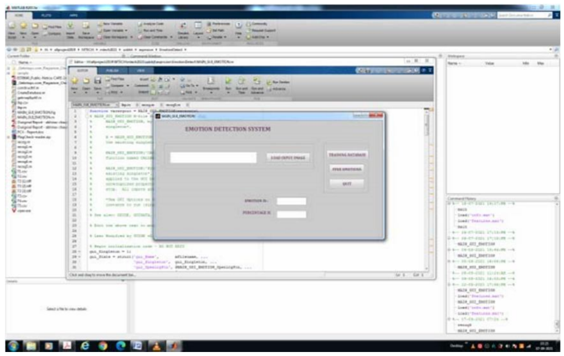
Table 2: parameter set for the proposed CNN.

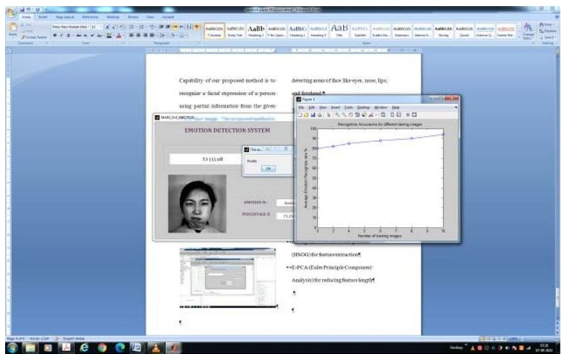
	C.1	S.1	C.2	S.2
Filters	64		256	
size	7×7	2×2	3×3	2×2
stride	1	2	1	2
pad	3	0	0	0

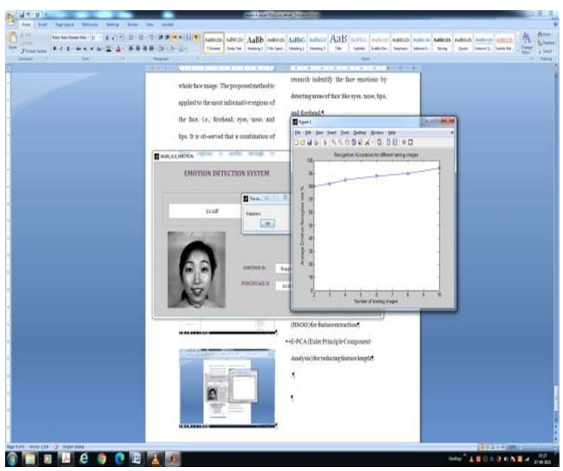
Datasets	Holistic SVM	Holistic KNN Euclidean	Holistic KNN Chi Square	Holistic KNN Histogram Intersection	Modular SVM	Modular KNN Euclidean	Modular KNN Chi Square	Modular KNN Histogram Intersection
JAFFE	90.02	87.32	81.42	80.02	95.67	92.14	88.32	86.97
VIDEO	91.55	87.43	79.89	77.12	97.77	91.54	88.12	86.91
CK+	90.15	87.71	86.41	83.54	94.78	91.24	86.79	85.64
OULU-CASIA	89.40	86.30	79.79	75.20	95.86	88.76	84.61	82.20

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IX Sep 2021- Available at www.ijraset.com

IV. COMPARISON ANALYSIS







V. CONCLUSIONS

Artificial Intelligent give us capability to detect emotions of human being. This research indentify the face emotions by detecting areas of face like eyes, nose, lips, and forehead.

The approach presented here is identifying human emotions by face few locations by implementing Deep learning with following technique.

A. Two Dimensional feature identification
B. Second order Histogram for features
C. PCA for reducing feature length

REFERENCES

[1] I. Abbasnejad, S. Sridharan, D. Nguyen, S. Denman, C. Fookes, and S. Lucey. Using synthetic data to improve facial expression analysis with 3d convolutional networks. In Proceedings of the IEEE International Conference on Com- puter Vision Workshops, pages 1609– 1618, 2017.

[2] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local bi- nary patterns: Application to face recognition. IEEE transactions on pattern analysis and machine intelligence, 28(12):2037–2041, 2006.

[3] P. AKoringa, G. Shikkenawis, S. K. Mitra, and S. K. Parulkar. Modified orthogonal neighborhood preserving projection for face recognition. In Pat- tern Recognition and Machine Intelligence, pages 225–235. Springer, 2015.

[4] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic. Netvlad: Cnn architecture for weakly supervised place recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 5297–5307, 2016.

[5] M. Baccouche, F. Mamalet, C. Wolf, C. Garcia, and A. Baskurt. Spatio- temporal convolutional sparse autoencoder for sequence classification. In BMVC, pages 1–12, 2012.

[6] S. A. Bargal, E. Barsoum, C. C. Ferrer, and C. Zhang. Emotion recognition in the wild from videos using images. In Proceedings of the 18th ACM International Conference on Multimodal Interaction, pages 433–436, 2016.

[7] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. In Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, volume 1, pages 454–461. IEEE, 2001.

[8] S. Biswas and J. Sil. An efficient expression recognition method using con- tourlet transform. In Proceedings of the 2nd International Conference on Perception and Machine Intelligence, pages 167– 174, 2015.

[9] J. Cai, Z. Meng, A. S. Khan, Z. Li, J. O'Reilly, and Y. Tong. Island loss for learning discriminative features in facial expression recognition. In 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), pages 302–309. IEEE, 2018.

[10] J. Chen, J. Konrad, and P. Ishwar. Vganbased image representation learning for privacy-preserving facial expression recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 1570–1579, 2018.

[11] L. Chen, M. Zhou, W. Su, M. Wu, J. She, and K. Hirota. Softmax regression based deep sparse autoencoder network for facial emotion recognition in humanrobot interaction. Information Sciences, 428:49–61, 2018.

[12] D. Ciregan, U. Meier, and J. Schmidhuber. Multi-column deep neural net- works for image classification. In 2012 IEEE conference on computer vision and pattern recognition, pages 3642 – 3649. IEEE, 2012.

[13] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang. Facial expression recognition from video sequences: temporal and static modeling. Computer Vision and image understanding, 91(12):160–187, 2003.

[14] C. A. Corneanu, M. O. Simón, J. F. Cohn, and S. E. Guerrero. Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications. IEEE transactions on pattern analysis and machine intelligence, 38(8):1548 – 1568, 2016.