Emotion Detection for Live Video Face Expression – Standard Learning Techniques

Reeja S.R., PhD
Vellore Institute of Technology
Amaravati, Andhra Pradesh

Ishfaq Yaseen
Prince Sattam bin Abdul Aziz University
Alkharaj, Saudi Arabia

ABSTRACT
Facial Emotion detection is taken into consideration very full-size for Human Computer Interaction and that they play vital position in regular people life. In latest years Convolutional Neural Network (CNNs) has end up very famous amongst researchers for picture primarily based totally analysis, due to the fact CNNs have generated satisfactory results. However, CNNs desires a number of records to train. This trouble has been addressed via way of means of numerous researchers who'veeducated CNNs with tens of thousands and thousands of pictures, this education understanding also can be utilized in a unique challenge that's called Transfer Learning. We assume the Facial Emotion Recognition version emotion detection to be beneficial in lots of programs including predictive mastering of students, lie detectors, etc.The proposed technique received 100% accuracy in emotion detection and 99.52% accuracy achieved for the proposed metho.

Keywords
Emotion Detection, Face Expression

1. INTRODUCTION
Facial Emotions detection has been a simple task for humans, however obtaining the feelings from the face with an algorithmic program is difficult. It is potential to notice that, Detecting emotion from picture using machine learning and computer vision is easy. The only way to perceive the feelings is by emotion detection that permits to grasp the mental or emotion state of the person. The feeling is detected by transfer learning. Transfer learning is a Deep learning methodology wherever a selected trained model developed for a task is reused for a few events of another task.

Emotion recognition are often employed in several applications through this feeling recognition technique lies are often detected throughout crime interrogation, Can notice gamers expertise through their emotions and create it higher per gamers want. It are often used for educational functions moreover, will understand students' interest in a course and create it higher so students will understand well, throughout interviews of recent candidates, it are often used for medical purpose moreover that's recognizing emotions of syndrome patients which might facilitate. Doctors to grasp during which state they're and what treatment they need also play a significant role in choice of advertisements on social media in step with current emotion of the user.

The main objective of this project is to implement a Emotions detection for face expressions using Transfer Learning (VGG16) with the use of a Convolutional Neural Network. There are many other approaches regarding Emotion Detection that have already been stated by date. However, the research work concentrate solely on face detection while leaving the back screen alone, resulting in a slew of useless as well as confusing elements which muddle the Whole training phase.

2. LITERATURE SURVEY
The author [1] discovered that effect of visual irregularity as an option of extracting features in a previous research using the Facial Expression recognition Dataset. According to our research, right disparity is superior to left imbalance. Facial discovery has a problem regarding facial expression. The answer for face pose attractiveness diversity is presented to [2]. He employed a particular aspect descriptor-based three dimensional posture consistent approach. Convolutional neural network (CNN) are used to overcome a number of challenges that arise during dealing upon that database, such as inappropriate cosmetic posture or emotion.

Currently, experts had represented a tremendous breakthrough in visual feature identification, that intends to promote study in the area of visual cues by improving neuroscientific and cognition technologies. Furthermore, advances in machine learning and artificial intelligence methods have made cognition rather more precise, simple, and services to a wider public. As a consequence, Visual Cognition as an inter and intra area of computer vision is quickly expanding. Biological interface, psychological assessments, intoxicated vehicle identification, and, most importantly, truth serums are all key uses.
The paper [3] shows the summaries of Alex Net and VGG CNN technique image data is mentioned in Fig.1.

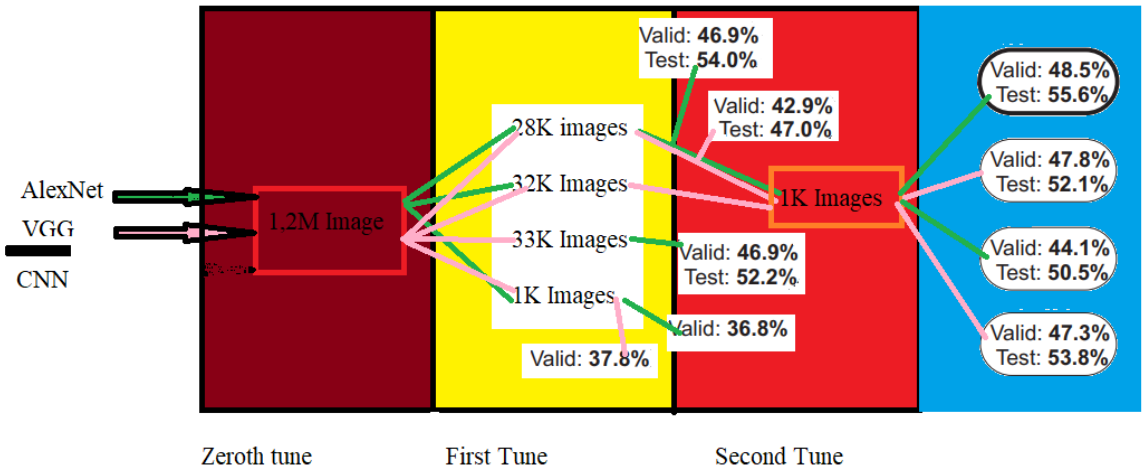

Fig.1.Alex Net and VGG CNN Technique image data

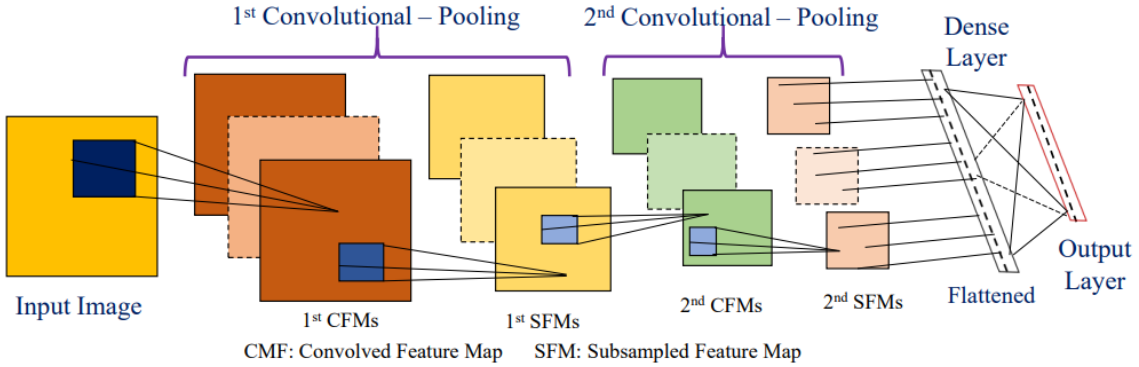

Fig.2. Layer of Pooling 2nd Convolution

The author [4] shows the layers of pooling with 2nd convolution is shown in fig.2. The accuracy obtained for this method is 96.5%.

In paper [5], the author used child and youngsters' pictures for FER. The goal of this work is to decide the number of image feature required points to get substantial expression categorization outcomes. In a critical appraisal setting, [6] used GMM-DNN method generates outcomes that are comparable toward that are achieved by others natural evaluation. The author used noisy and normal environment. The noise can be removed by various methods, the author of [9],[10],[11],[12],[13] and [14] are mentioned the details clearly about the approaches. The author [7] used Visualization of an unique Psychological Imagination Picture. facial image identification using a genuine synthetic facial landmark feature approach to detect the picture feeling [8] is another type of approach.

3. TRANSFER LEARNING METHOD

The following block diagram shows in fig.3, the system architecture of our work. The suggested Classification structure is shown in Fig 3 utilizing VGG-16, a very well pre trained DCNN algorithm. The VGG-16 algorithm is used to categorize thousands of picture components using the Matching database. The pre-trained model is fine-tuned using sensitivity variables before being updated for event detection by reconfiguring the deep levels. So last deep levels of the pre-trained network are substituted only with newest deep layer(s) to recognise a visual picture into either of 7 feeling categories when establishing the design (afraid, angry, disgusted, sad, happy, surprised, and neutral).

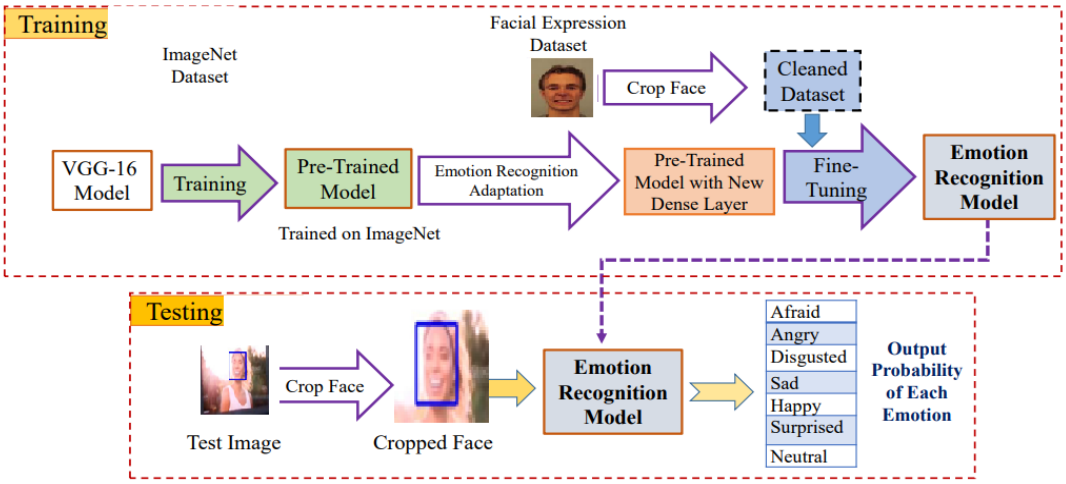

Fig.3. Proposed model

A deep learning model in a neural network is a regular, fully connected, linear layer that accepts specified scale as input and provides a scalar of the size requires. As a result, the output unit had only 7 neurons. A deep learning method in a neural network is a normal, densely integrated, sequential level which accepts certain scale as input and provides a vector of the desired dimension. As a result, neurons are present in a hidden layers. The design with multilayer foundation of the pretrained model and the convolution layers is finetuned. To train in finetuning, a purified emotive sample generated using pretreatment (that is scaling, cutting and other operations) has been used. In the instance of evaluation, a chopped picture is sent into the software's intake and the choice was made

value of maximum likelihood of sensation response. Some alternative deep convolutional neural structure, such as ResNet, DenseNet, or Fusion, can be used in replacement of VGG-sixteen. As a result, the suggested figure's scale determined by the scale of the pretrained structure and the design of the new deep network.

4. EXPERIMENTAL ANALYSIS

Transfer learning is a concept that entails transferring the knowledge of different pre-trained systems to real resources. Instead of developing our the own neural net, we'll employ widely available post versions, transfer all information to them, and eventually receive the selections for our photographs. Processing element provides choices for photographs by default. It performs inversion on each component of the images and, as a result, produces a 'n' arrays which is nothing more than the images' learned choices.

4.1 Data Processing

The data collection is called fer2013, and it's a transparent information which was publicly revealed for a Kaggle challenge. It has bitmaps picture of the facial that are 48 X 48 pixels in height and width. The information is divided into the following categories: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral. The File format has two columns: emotional, which is a numbering system between 0 to 6, & pixel, which is a phrase surrounding in marks for every visual. It takes in one numpy array and convert it to a 48*48 2D array. It will also perform the rescale part so that every element in the array is a number between 0 and 1.

4.2Extracting key frames

Facial reactions are detected using Visual Reactions recognition using a live video camera stream. In this scenario, the variation among each film is estimated whenever the source feed to the Visual Emotion Classification in a multimedia screen. Once the terms of inter variance is zero, the films are optimally steady. After that, most of the films were stabilized, and a Canny Edge Detector (ED) was used to estimate the accumulated average including all white pixels. Upon adding up additions of all steady films, the film with the highest summation is chosen because it has the most of the functionalities as far as boundaries are concerned. This film has been chosen as a FER feed.

4.3 Experimental Set-up

The culminating picture alternatives which we have an inclination to receive somewhere at deep convolution channel's peak are referred to as constriction possibilities. Those constriction possibilities are the pictures' learned alternatives, which are subsequently fed towards the MLP, which functions as an upper edge. This MLP then decreases the cost operation and changes the levels in the MLP as well as the CNN residuals. To even get constraint alternatives, use the VGG-sixteen pre-trained neural net. The VGG-sixteen net has 16 units, 13 of which would be basic convolution layers. This neural network is very well trained using the FER collection, which comprises a large number of images. This VGG-sixteen trained network will conveniently available in Keras. As a result, this VGG-sixteen layer has been post and preloaded with payloads. VGG- sixteen is one of the benchmarks utilized in this work**.**

5. RESULTS

We have tested the Face emotions detection for face expressions on both live web cam and still images also. Each expression was tested individually is shown in fig.4. The accuracy for the emotion happy is 100 percentage, Neutral is 98.1%, sad is 99.7%, angry is76.9%, surprise is 99.9% , disgust is 88.8%.

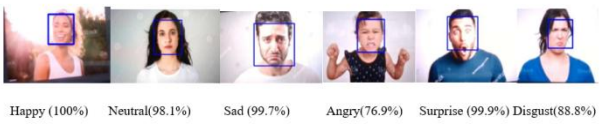

Fig.4. Various emotion Accuracy

The accuracy comparison for various methods is shown in table.1. The proposed technique received 99.52% Accuracy. The accuracy for various emotions are shown in Fig.5.

Table.1. Accuracy Comparison

Method	Accuracy
DWT with 2D-LDA and SVM Classification	95.70
Gabor, Haar and SVM,KNN and LDA classification	89.50
Std CNN and fully connected layer classification	91.67
Hybrid deep learning with CNN and RNN	94.91
Proposed Method	99.52

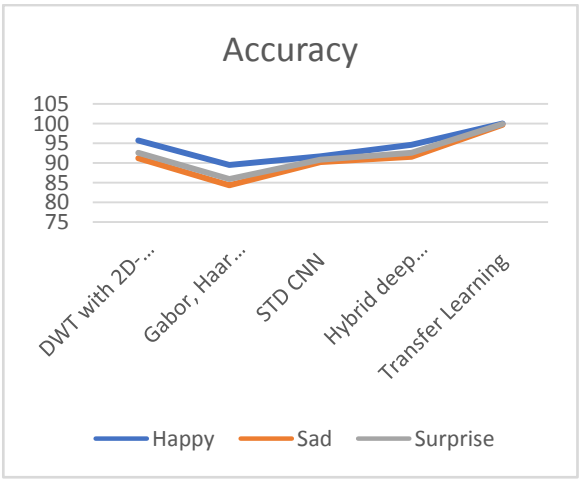

Fig.5. Accuracy for various emotion

6. CONCLUSIONS

We proposed fine tune pooling with transfer learning technique for real video facial images. The result narrates six different visual data collection is feed to various trained model with various views. All the classification shows the six features of visuals with very good accuracy. Most of the emotions, we achieved 100% accuracy. This method can be useful in hospital for monitoring the patients. The data collections are the real visual video expressions. By comparing with all the existing method, our proposed method shows 99.52 percentage of accuracy.

7. REFERENCES

[1] Sajid, Muhammad, et al. "Facial asymmetry-based anthropometric differences between gender and ethnicity." Symmetry 10.7 (2018): 232.

[2] Ratyal, Naeem, Imtiaz Ahmad Taj, Muhammad Sajid, Anzar Mahmood, Sohail Razzaq, Saadat Hanif Dar, Nouman Ali, Muhammad Usman, Mirza Jabbar Aziz Baig, and Usman Mussadiq. "Deeply learned pose invariant image analysis with applications in 3D face recognition." Mathematical Problems in Engineering 2019 (2019).

[3] Ng, Hong-Wei, Viet Dung Nguyen, Vassilios Vonikakis, and Stefan Winkler. "Deep learning for emotion recognition on small datasets using transfer learning." In Proceedings of the 2015 ACM on international conference on multimodal interaction, pp. 443-449. 2015.

[4] Akhand, M. A. H., Shuvendu Roy, Nazmul Siddique, Md Abdus Samad Kamal, and Tetsuya Shimamura. "Facial Emotion Recognition Using Transfer Learning in the Deep CNN." Electronics 10, no. 9 (2021): 1036.

[5] Rao, A., Ajri, S., Guragol, A., Suresh, R., & Tripathi, S. (2020). Emotion Recognition from Facial Expressions in Children and Adults Using Deep Neural Network. In Intelligent Systems, Technologies and Applications (pp. 43-51). Springer, Singapore.

[6] Shahin, Ismail, Ali Bou Nassif, and Shibani Hamsa. "Emotion recognition using hybrid Gaussian mixture model and deep neural network." IEEE access 7 (2019): 26777-26787.

[7] Mal HP, Swarnalatha P (2017) Facial expression detection using facial expression model. In: 2017 International conference on energy, communication, data analytics and soft computing (ICECDS). IEEE, pp 1259–1262

[8] Danisman T, Bilasco M, Ihaddadene N, Djeraba C (2010) Automatic facial feature detection for facial expression recognition. In: Proceedings of the International conference on computer vision theory and applications, pp 407–412

[9] S. R. Reeja and N. P. Kavya, "Real time video denoising," 2012 IEEE International Conference on Engineering Education: Innovative Practices and Future Trends (AICERA), 2012, pp. 1-5, doi: 10.1109/AICERA.2012.6306745.

[10] Reeja, S.R., Kavya, N.P.: Noise reduction in video sequences: the state of art and the technique for motion detection. Int. J. Comput. Appl. 58(8), 31–36 (2012)

[11] Reeja, S. R., and N. P. Kavya. "A System for Movement Detecting Congestion.", INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY 13 (4), 4302-4307

[12] Reeja, S. R., & Kavya, N. P. Hypothesis Testing for Motion Detection and Noise Estimation of an Image Frame, IFRSA's International Journal Of Computing 4 (2), 599-603

[13] Kavya, R. S., & Dr, N. P. (2014). An Approach for Noise Removal from a Sequence of Video. International Journal of Scientific & Engineering Research, 5(4), 1266-1270.

[14] S.R. Reeja, Rino Cherian, Kiran Waghmare, Jothimani, Chapter 7 - EEG signal-based human emotion detection using an artificial neural network, Editor(s): Hemanth D. Jude, Handbook of Decision Support Systems for Neurological Disorders, Academic Press, 2021, Pages 107-124, ISBN 9780128222713, https://doi.org/10.1016/B978-0-12-822271-3.00007-4.