

Facial Emotion Recognition of Students using Deep Convolutional Neural Network

K.Nithiyasree^a, A.Nisha^b, S.Shankar^c, N.AkshayKumar^d and T.Kavitha^e

^{abcd}Students, ^eAssistant Professor(SS)

Department of Computer Science and Engineering

PeriyarManiammai Institute of Science & Technology, Vallam,Thanjavur 613403, Tamil Nadu, India

Article History Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Understanding the emotions of the students in a classroom lecture will improve the Teaching – Learning process. It is difficult to find out the students' interest in a conventional classroom lecture. Hence, we are providing a solution by analyzing the facial emotions of the students in a classroom with the help of a deep learning network. Deep learning is an advancement of machine learning technique which gives more accurate results than the machine learning algorithms. The proposed system will give an accurate prediction with the help of deep convolutional neural networks using VGG-16 architecture. This computer vision model will analyze every individual student's emotion from the live video taken from the video camera fixed in the classroom and provide the overall emotion of the class based on the highest probability of students' emotion. Later, our proposed VGG-16 architecture will be compared with models such as Alexnet and Resnet architecture to ensure its better accuracy.

Keywords: Deep Learning, Emotion Database, Face Emotion Detection, Facial Features, VGG16-Model.

1. Introduction

Facial expression for emotion detection has always been an easy task for humans, but achieving the same task with a computer algorithm is quite challenging. With the recent advancement in computer vision and machine learning, it is possible to detect emotions. Facial and emotional expressions are the most significant nonverbal ways for expressing internal emotions and intentions. Education, there is already a growing awareness on the importance of social-emotional competencies in terms of improving the student learning process, academic performance mindset and behavior. It is stated that emotion has a significant impact on thinking and reflective learning. Thus, with the use of effective teaching approaches based on the students' experiences in the academic study, teachers should create positive emotions inside the classroom to achieve a better learning outcome. Teachers should learn the physiological changes through reading the faces of their students for them to assess their level of understanding. Moreover, a teacher should understand that cognitive appraisal is greatly associated with emotion and it effectively motivates the students' behavior. The freedom of learning is highly dependent on emotions, thus making the educators play a significant role in controlling over and promoting an emotionally engaging environment for the students to acquire knowledge.

The objective of this paper is to detect the emotion of students using VGG16 architecture to be able to identify the overall interest of students about the lecture based on the facial expression of students. The system will predict the facial expressions such as surprise, fear, disgust, sad, happy, angry and neutral from the student face and identify the overall interest of the students based on the highest probability of the emotion. As a result of this the efficiency of Teaching – Learning methodology is improved by analyzing the student's emotion from their faces which will help teachers to adjust their strategy and their instructional materials. Further, the VGG16 architecture is compared with ResNet and AlexNet architecture to ensure the VGG16 better accuracy.

2. Related Works

ImaneLasri et.al. discussed the emotions to analyse the facial features for classification of face images. The dataset used is KDEF (Karolinska Directed Emotional Face). There are three approaches used in this paper: (i) For recognizing the area of facial skin, EBM is used. (ii) For the extraction of facial elements such as eyes, lips, and nose, Viola-Jones algorithm is used. (iii) For the extraction of anthropometric features such as distance between the eyes, eyebrow) extraction, FAST algorithm used. For classification the KVM, SVM, Bagged tree classifiers are used for testing in order to find the best classifier among the three. The average classification accuracy 57.7% for 6 different emotions (joy, surprise, sadness, anger, fear and disgust) is achieved for the Bagged Trees algorithm, the reason for this accuracy is the emotion fear is difficult to recognize and average classification accuracy 95.9% for 2 emotions (joy, surprise).

Rong Fu et.al. discussed the emotion of students to be able to identify the most appropriate teaching pedagogies. The dataset used is Filipino facial expression database The proposed prototype model has three major steps: a. Face detection, b. Emotion detection, c. Aggregation of emotion. The advantage of this model is it provides solutions pertaining to what appropriate teaching pedagogy must be applied. The model gained 80.11%

accuracy using the SVM classifier. The accuracy for each emotion is “Anger”78.43%, “Disgust” 100%, “Neutral” emotion with 47.06%, “Sad” 58.82% and “Surprised 78.47%.

Yang Tao et.al. discussed about the facial emotion such as neutral, happy, anger, surprise using the combination of a VGG16 CNN and a Bi-LSTM RNN with an attention mechanism. The dataset used is The IST-EURECOM Light Field Face Database (LFFD) . Proposed a novel light field based facial emotion recognition method with combination of three major elements: a. deep VGG-Face network, b. Bi-LSTM RNN , c. An attention mechanism.

D. Yang et.al. proposed a model to predict the emotion of the student’s expression based on mini-Xception optimization in the convolution kernel of the third convolution layer is changed to 5×5, which enables the network to obtain more features and further improve the classification efficiency of facial expressions.

Arpita Gupta et.al. proposed a model to classify seven emotions. They proposed a deep learning framework which includes CNN, ResNet and attention block. They proposed a model with three different networks based on CNN named A, B and C. The network A is based on Krizhevsky and Hinton and consists of three convolution layers, two fully connected layers, max pooling and dropout. The network B is based on Alexnet consisting of three convolution layers, three connected layers, normalization and dropout. The network C is based on Gudi consisting of one convolution layer, normalization, fully connected, dropout and max pooling. The proposed model has outperformed the existing CNN based networks by achieving the higher accuracy of 85.76 in training phase and 64.40 in testing phase. The proposed model has shown better results on the FER dataset and could be used in real-time applications.

Mehmet Akif OZDEMIR et.al. proposed a Convolutional Neural Network (CNN) based architecture for facial expression recognition called LeNet architecture (Real Time Emotion Recognition). In this paper they Merged 3 datasets (JAFFE, KDEF and their custom dataset). They got 0.0887 as training loss, training accuracy as 96.43%; validation loss as 0.2725 and validation accuracy as 91.81%. As per confusion matrix, the proposed LeNet model was more exact at expectation of surprised, fear, neutral emotion states and less precise at forecast of sad emotion state.

RohitPathar et.al. discussed the categorization of a facial image into one of the seven emotions by building a multi class classifier. In this paper they used convolutional neural networks (CNNs) for training over gray scale images obtained from FER2013 dataset. The first to eight convolution layer is increasing 32 3x3 filters, which increases 32 filters for each layer. FC layer concludes the class scores. The dataset contains around 35,887 well-structured 48x48 pixel grayscale images. As a result of comparing the accuracy at different depths they achieved 89.98% as maximum accuracy. There are comparatively less number of images for particular emotion like disgust in FER2013 dataset that results in average performance of the model in recognizing the disgust emotion.

The author JikeGe et.al. made a study on recent works of facial recognition FER via deep learning. The automatic FER task goes through different steps like: data processing, proposed model architecture and finally emotion recognition. They used FER dataset. According to these studies, researchers create two novel models of CNN which achieve average 65.23% and 65.77% of accuracy. The particularity of these models is that they do not contain fully connected layers dropout, and the same filter size remains in the network.

Oussama El Hammoui et.al. discussed the Deep CNN as a Stacked Convolutional AutoEncoder (SCAE) in a greedy layer-wise unsupervised fashion for Emotion recognition using facial expression images. They developed a SCAE with reduced number of deep learning layers for emotion recognition and compared the result with a CNN for emotion recognition. They used batch normalization to speed up the training process and improved the classification performance. The CNN with Batch Normalization (BN) and the SCAE emotion recognizers were trained and tested using the KDEF dataset. When trained from scratch for 500 epochs and with a random weight Initialization, the deep CNN model with BN produces an accuracy rate of 100% on the training set and a peak performance of 91% on the testing.

The author Min Hu et.al. discussed facial expression recognition in e-learning systems and facial expression recognition using CNN. Their system consists of three main steps: preprocessing, features extraction and classification. They tested the proposed system with students aged between 8 and 12 years old, in an educational game. The model is trained and tested on a combination of images from the CK+ and the KDEF databases. The test was successful and the system was able to detect faces and classify emotions with an accuracy of 97.53% on test data and an accuracy of 97.18% on the JAFFE dataset.

3. PROPOSED SYSTEM

We proposed a system which will predict the facial emotions (happy, sad, angry, surprise, neutral) of a student face, using the VGG16 model. For this the VGG16 model has to be trained with a FER2013 dataset through which the model can learn how to recognize the emotions. Then, we have to load the images of students of the classroom with their name for training to get precise recognition. The model will create an array for the known face encoding. Now, the model is trained.

Now, the live video will be fed into a trained model which will be converted into frames and then the emotion labels will be generated to the person in the video. Then, the system will provide the solution based on the highest probability of students' emotion, for this we use matplotlib to represent the result in the form of a pie chart. Finally, our proposed system compared with other existing systems to ensure its better accuracy.

4. BLOCK DIAGRAM

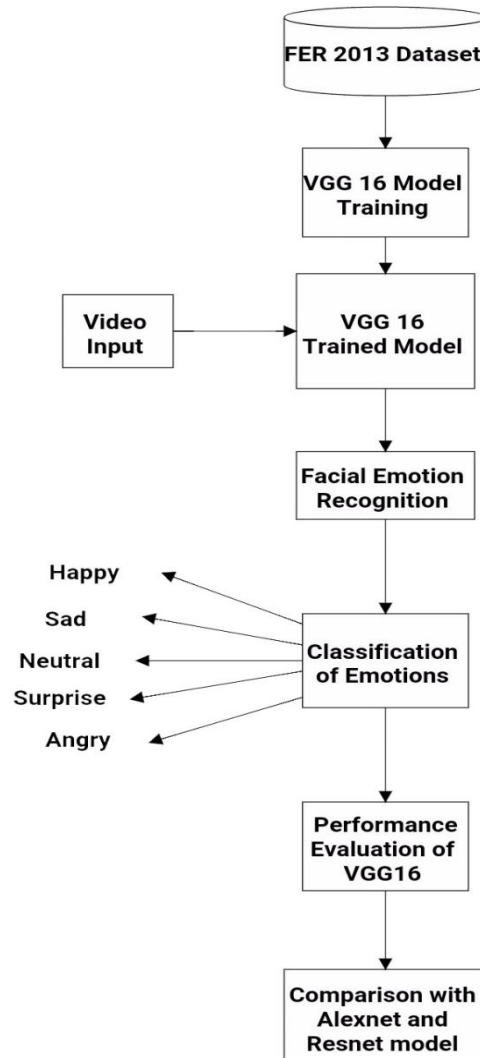


Figure 1: Facial Emotion Recognition System

5. EXPERIMENTAL RESULTS

a. Dataset

The FER-2013 dataset consists of 28,000 labelled images in the training set, 3,500 labelled images in the development set, and 3,500 images in the test set. Each image in FER-2013 is labelled as one of seven emotions: happy, sad, angry, afraid, surprise, disgust, and neutral.

b. Model Implementation

The live classroom video fed into the model. The video is read from frame to frame. The video frame will label the student with his emotion and then overall emotion of the classroom will be generated as a pie chart. Figure 1 and 2 is the output of implementing the model.



Figure 2: Students Emotion Recognition

Emotions Recorded

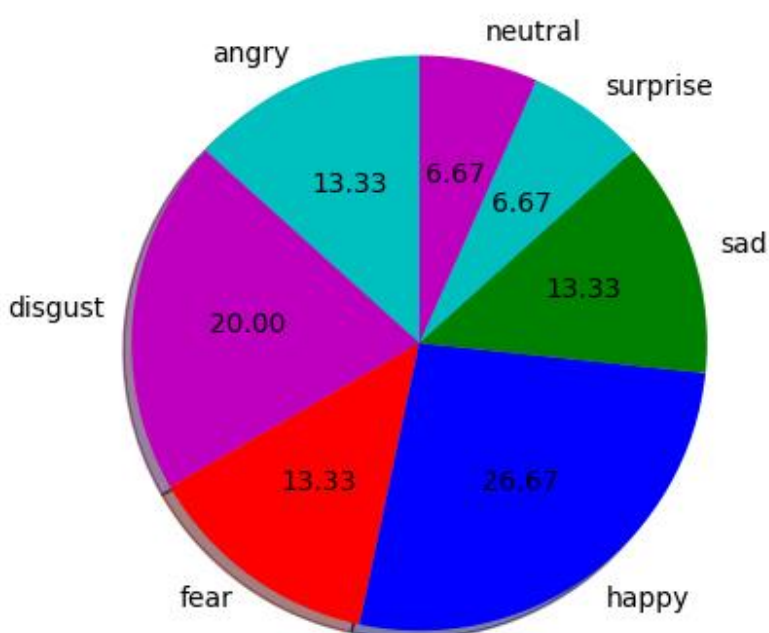


Figure 3: Classroom Emotion Pie Chart

Further, we compared our VGG-16 model with the AlexNet and ResNet model, the accuracy results shown in Table 1, in which VGG-16 achieved better accuracy than the other two models.

Table 1. Architecture accuracy comparison

S.NO	Architecture	Accuracy
1.	VGG16	89%
2.	AlexNet	87%
3.	ResNet	71%

6. CONCLUSION

In this paper, the deep CNN based face emotion detection system is implemented. The VGG16 model is trained using the FER2013 dataset. Then, the live video of the classroom is given as input to the trained VGG16 model for the detection of the face emotion (happy, sad, neutral, surprise, angry) of the students. Based on the obtained emotion of the students the overall students emotion is evaluated which will be diagrammatically represented. This information will help the teacher to improve their teaching methodology. Later, our proposed model is compared with model such as Alexnet and Resnet in which we got accuracy as VGG16 (89%), ResNet

(87%) and AlexNet (71%) architecture for classifying the different emotions of students in which the VGG16 architecture outperforms well among the three architectures.

References

1. ImaneLasri ,AnouarRiadSolh, Mourad El Belkacemi .(2019). Facial Emotion Recognition of Students using
2. Convolutional Neural Network .IEEE .
3. Rong Fu, Tongtong Wu, ZuyingLuo,AndFuqingDuan,XuejunQiao, Ping Guo College of Artificial Intelligence,
4. Beijing Normal University, Beijing.(2019). Learning Behavior Analysis in Classroom Based on Deep Learning. China.10th International Conference on Intelligent Control and Information Processing Marrakesh, Morocco.
5. Yang Tao*,Yuanzi He ,Wei Zhang. (2020). An Application of Face Recognition Technology in University
6. Classroom Teaching . Nanchang Institute of Science and Technology Nanchang, China 2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI).
7. D. Yang, A. Alsadoon, P. W. C. Prasad, A. K. Singh, and A. Elchouemi. (2018). An Emotion Recognition Model
8. Based on Facial Recognition in Virtual Learning Environment. Procedia Computer Science.
9. ArpitaGupta^a, SubrahmanyamArunachalam^a, RamadossBalakrishnan^{a,*}. (2019). Deep self-attention network for
10. facial emotion recognition. ,Third International Conference on Computing and Network Communications (CoCoNet'19) ,ELSEVIER.
11. Mehmet Akif OZDEMIR, Berkay ELAGOZ, Aysegul ALAYBEYOGLU, Reza SADIGHZADEH³and Aydin
12. AKAN .(2019). Real Time Emotion Recognition from Facial Expressions Using CNN Architecture. Department of Biomedical Engineering, Department of Computer Engineering, Business Administration,IEEE.
13. RohitPathar ,AbhishekAdivarekar, AbhishekAdivarekar, ,AnushreeDeshmukh. (2019).Human Emotion
14. Recognition using Convolutional Neural Network in Real Time”, Information Technology Rajiv Gandhi Institute of Technology, Mumbai University Mumbai, India.
15. An Liu, JikeGe, Dong Chen, Guorong Chen. (2018). An Online Classroom Atmosphere Assessment System for
16. Evaluating Teaching Quality. School of intelligent technology and engineering Chongqing University of Science and Technology,IEEE.
17. Oussama El Hammoumi, FatimaezzahraBenmarrakchi, NihalOuherrou, Jamal El Kafi, Ali El Hore.
18. (2020).Emotion Recognition in E-learning Systems. Dept. of Computer Science ChouaibDoukkaliUniversity,ElJadida, Morocco, IEEE.
19. Min Hu, Haowen Wang, Xiaohua Wang, Juan Yang, Ronggui Wang.(2018). Video facial emotion recognition
20. based on local enhanced motion history image and CNN-CTSLSTM networks. ELSEVIER.