International Journal of Research Publication and Reviews, Vol 2, no 12, pp 808-814, December 2021

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Facial Emotion Recognition Using Deep Learning

Akhila J[1], Prof Rajitha P.R[2], Dr.T.mahalekshmi[3]

[1]Final year Student, Sree Narayana Instiute of Technology, Kollam, Kerala,India
[2]Assistant Professor, Sree Narayana Instiute of Technology, Kollam, Kerala,India
[3]Principal, Sree Narayana Instiute of Technology, Kollam, Kerala,India

ABSTRACT

The fast development of computerized reasoning has contributed a great deal to the innovation world. As the customary calculations neglected to meet the human necessities progressively, Machine learning and profound learning calculations have acquired extraordinary accomplishment in various applications, for example, characterization frameworks, proposal frameworks, design acknowledgment and so forth Feeling assumes an essential part in deciding the considerations, conduct and sensation of a human. A feeling acknowledgment framework can be worked by using the advantages of profound learning and various applications, for example, input investigation, face opening and so forth can be carried out with great precision. The principle focal point of this work is to make a Deep Convolutional Neural Network (DCNN) model that groups 5 distinct human facial feelings. The model is prepared, tried and approved utilizing the physically gathered picture dataset.

Keywords—Facial Emotion Recognition, Deep Convolutional Neural Network, Classification, Adam.

I. INTRODUCTION

The turn of events and utilization of PC frameworks, programming and organizations are developing colossally. These frameworks play a significant part in our regular daily existence and they make human existence a lot more straightforward. Facial feeling acknowledgment framework expects a ton of significance in this time since it can catch the human conduct, sentiments, aims and so on The customary techniques have restricted speed and have less exactness while facial feeling acknowledgment framework utilizing profound learning has ended up being the better one. This framework plans to fabricate a profound convolutional neural organization model that perceives 5 distinct human facial feelings and this can be utilized for applications, for example, client criticism examination, face unlocking and so on.

In the field of software engineering, AI is one of the arising advancements that is considered to have an effect of 90% in the following 4 years. Profound learning, a subset of AI utilizes fake neural organization, which is a calculation roused from the human mind. Convolutional Neural Network (CNN) is a class of profound neural organization that utilizes convolution as the numerical activity. As the dataset comprises of pictures, the framework utilizes a 2D CNN for the acknowledgment task. The proposed profound convolutional neural organization is prepared not exclusively to arrange 5 unique human facial feelings, yet in addition to yield a decent precision. The model is prepared utilizing the dataset which is gathered physically utilizing a cell phone camera.

II.PROPOSED SYSTEM

This section describes CNN and the architecture of the DCNN model of the proposed system.

A. Convolutional Neural Network

Neural network is a bunch of calculations that emulate the human cerebrum and it tracks down a connection between the information to get arrangements utilizing these calculations. CNN is a sort of Neural Network where the numerical activity used to observe the connection between the information is Convolution. Customary neural organization bombs when coming to complex issues like picture order, video arrangement, design acknowledgment, and so forth however CNN has made incredible progress in these applications, yielding great precision.

CNN comprises of essentially 4 Layers, convolutional, pooling, dropout and completely associated layers. These layers together concentrate the elements from the information. The calculation gains from the element, where the elements of interest are addressed by every convolution channel. The convolutional layer comprises of little fixes, which changes the whole picture dependent on the channel esteems. Condition is the recipe to make include maps, for example the result from the convolutional layer.

$$G[m,n] = (f * h)[m,n] = \sum_i \sum_k h[j,k] f[m-j, n-k]$$

Where *f* is the input image, h is the filter, (ǎ̄ ȯ) is the size of the resulting matrix generated.

The result from the convolution layer is given to a pooling layer where its size gets diminished with no deficiency of data. These 2-dimensional exhibits are changed over to a solitary dimensional vector utilizing the level layer with the goal that it very well may be taken care of to the neural organization for arrangement. The neural organization utilizes the back-spread calculation where the blunders are back engendered to change the loads, consequently decreasing the mistake (misfortune) work. The weight and ़ ᴾ□ is given by the delta rule as in the equation

$$,\Delta W_i = \mathrm{n}\frac{dE}{dW_i} x_i$$

Where n is the learning rate e is the error function and x_i is the input.

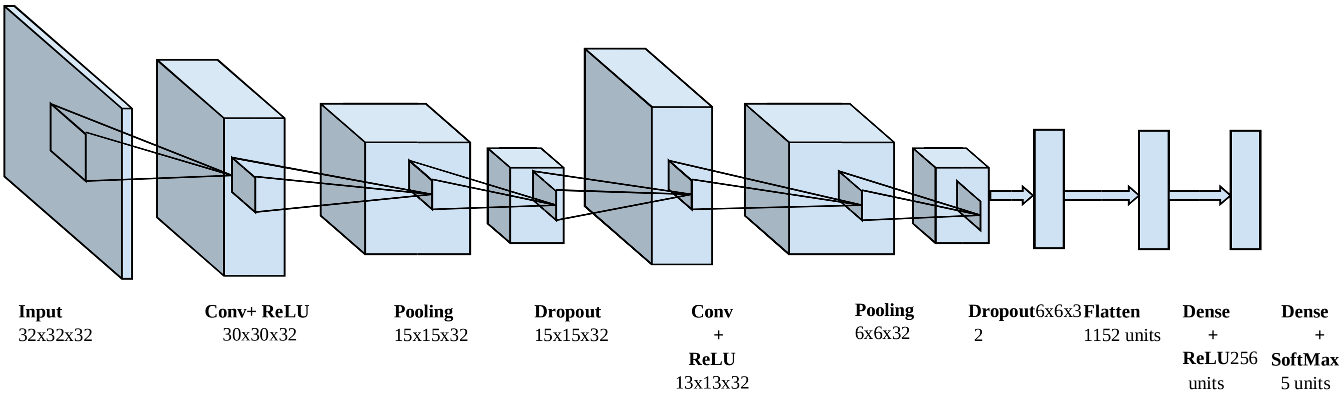

Figure 1: Architecture of the proposed facial emotion recognition model using DCNN

B.Proposed CNN Model

The design for the proposed facial emotion recagnition model is portrayed in Figure 1. The model uses two convolution layers with dropouts later every convolution layer. The info picture is resized to 32 x 32 and is given to the primary convolution layer. The result from the convolution layer, called highlight map, is gone through an initiation work. The enactment work utilized here is ReLU (Rectified Linear Unit) that makes the negative qualities zero while the positive qualities continue as before. This element map is given to the pooling layer of pool size 2 x 2 to lessen the size without losing any data. Dropout layer is utilized to lessen the over fitting. This cycle again proceeds for the following convolution layer too. At long last, a 2dimensional cluster is made with some component esteems. Smooth layer is utilized to change these 2-dimensional exhibits over to a solitary dimensional vector in order to give it as the contribution of the neural organization, addressed by the thick layers. Here a two-layer neural organization is utilized, one is input and the other is yield. The result layer has 5 units, since 5 classes should be ordered. The enactment work utilized in the result layer is SoftMax, which creates the probabilistic result for each class. Figure 2 portrays a depiction of the model synopsis of the proposed framework which is fabricated utilizing the Keras DL Library.

III SCOPE

This venture work manages perceiving the looks of individuals. As there exist a few techniques it isn't practical to run these calculations in genuine climate, hence, a continuous framework is need to create. Likewise, existing works perceive 2 or 3 looks, subsequently there is a need to carry out a framework that can ready toperceive various feelings and foreseewhether or not the individual is anxious.

IV DATASET

The dataset for the proposed model incorporates 5 unique facial feelings viz. irate, glad, impartial, dismal and shocked. These are gathered physically utilizing a 48 MP camera. Each picture has a pixel size of 1920 x 2560. The dataset split is displayed in Table I. Each class comprises of the very number of preparing tests with the goal that they are not one-sided. The train-test-approval split is in the proportion 8:1:1.

TABLE I. DATASET SPLIT RESULTS

Number of classes	5
Number of training images	2040

Number of validation images	255
Number of testing images	255

V METHODOLOGY

Utilizing python as the programming language, the model is carried out. The whole model is reenacted in the Jupiter Notebook. For building the model, adding the convolution layers, assembling and fitting the model, Keras, which runs on top of TensorFlow, is utilized as the profound learning library. Scikitlearn is the bundle utilized for tracking down the disarray lattice that gives the exactness, accuracy, affectability, particularity, review, and so on of the model. For plotting the disarray network and different diagrams like exactness and misfortune, matplotlib and seaborn are utilized.

VI. LITERATURESURVEY

1. Deep Convolutional Neural Networks For Facial Emotion Recognition

For facial feeling acknowledgment, this paper proposes a two-layer convolution network model. The model uses the between them. The main convolution layer gets the picture dataset to group five distinct facial feelings. The model utilizes two convolution layers, with dropouts input picture, which is resized to 32 x 32 pixels. The model has tantamount preparing and approval precision, showing that it has the best fit and is generalizable to the information. The initiation work utilized here is ReLU (Rectified Linear Unit), which diminishes negative qualities to nothing while at the same time keeping up with positive qualities. This element map is applied to a pooling layer with a pool size of 2 x 2 to decrease the size without forfeiting any detail. The model diminishes the misfortune work utilizing an Adam streamlining agent, and it has been checked to have an exactness of 78.04 percent.

2. Emotion Recognition using Deep Neural Network with Vectorized Facial Feature

In this paper, proposed a DNN model which utilizes vectorized facial highlights as information. Vectorized facial component for look will be presented. The vectorized facial component can be utilized to fabricate a DNN (Deep Neural Network) for feeling acknowledgment. The proposed facial element model cannot just reflect looks effectively, it can likewise be utilized for DNN with high proficiency. To test the effectiveness of such technique, a DNN is prepared to perceive some general articulations. Contrasted and other PC vision fueled framework, vectorized facial highlights can accomplish comparative precision as other AI calculations (CNN). However, it decreases the information just as the time needed for preparing. Such benefits can essentially speed up building applications including feeling acknowledgment. The model can anticipate various feelings with an exactness of 84.33%.

3. Sign Language Recognition System Using Deep Neural Network

This paper proposes a basic 2 layer convolutional neural network (CNN) to classify sign language image datasets. The classifier was found to perform with varyinglightingand noisy image datasets. This model has classified 6 different sign languages using two differentoptimizers, SGD and Adam with an accuracy of 99.12% and 99.51% respectively. More accuracy is obtained when using the Adam optimizer. Future work of this Sign LanguageRecognition System can be extended to improve the performance by tuning the hyper parameters and implement a sign language recognition system from video sequence using CNN LSTM. This sign language recognition system can also be made to control certain devicessuch as home robot. The two CNN models developed have different type of optimizers, the Stochastic Gradient Descent (SGD) and Adam.

4. Speech Emotion Recognition Using Deep Neural Network and Extreme Learning Machine

In this paper we propose to use profound neural organizations (DNNs) to remove undeniable level highlights from crude information and show that they are successful for discourse feeling acknowledgment. gauge feeling states for every discourse section in an expression, build an expression level element from fragment level assessments, and afterward utilize an ELM to perceive the feelings for the expression First produce a feeling state likelihood conveyance for every discourse portion utilizing DNNs. Then, at that point, build expression level elements from fragment level likelihood conveyances. These expression level elements are then taken care of into an outrageous learning machine (ELM), a unique basic and productive single-stowed away layer neural organization, to distinguish expression level feelings. The exploratory outcomes exhibit that the proposed approach successfully takes in enthusiastic data from low-level elements and prompts 20% relative precision improvement contrasted with the best in class draws near.

5. Video-Based Emotion Recognition using CNN-RNN and C3D Hybrid Networks

In this paper, present a video-based feeling acknowledgment framework. The center module of this framework is a half breed network that consolidates repetitive neural organization (RNN) and 3D convolutional networks (C3D) in a late-combination style. Additionally add a sound classifier into the framework. RNN and C3D encode appearance and movement data in various ways. In particular, RNN takes appearance highlights extricated by convolutional neural organization (CNN) over individual video outlines as information and encodes movement later, while C3D models appearance and movement of video at the same time. Joined with a sound module, our framework accomplished an acknowledgment precision of 59.02% without utilizing any extra feeling named video cuts in preparing set. Use CNN-LSTM and C3D organizations to all the while model video appearances and movements. Particularly, we observed that the mix of the two sorts of organizations can give great outcomes, which showed the adequacy of the technique.

VII RESULTS

CNN is trained with the emotion image dataset, utilizing Adam as the optimizer and the categorical cross-entropy as the loss function. The model parameters are shown in Table II.

TABLE II. MODEL PARAMETERS

Model Parameters	**Values**
Total images	2550
Activation	ReLU and SoftMax
Learning rate	0.01
Epochs	11
Optimizer	Adam
Loss function	Categorical Cross-entropy

Adam is an optimization algorithm that can be used instead of the classical stochastic gradient descent algorithm to update the network weights with individual learning rate for each of the weights. For each weight of the neural network it uses first and second moment estimations of gradient to adapt the learning rate. The nth moment of the random variable is provided in equation,

$$m_n = E[X^n]$$

Where m is the moment and X is the random variable. The first moment is given by the mean and the second moment is given by the un-centered variance. Adam uses exponentially moving averages to estimate the moments. Moving averages of gradient and squared gradient are given by (1) and (2) respectively.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)g_t \quad (1)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)g_t^2 \quad (2)$$

Categorical cross-entropy, the loss (error or cost) function used for optimizing classification models, is given by,

$$L(y, y') = -\sum_{j=0}^{M} \sum_{i=0}^{N} (y_{ij} * \log(y'_{ij}))$$

Where Σ is the predicted value. This function will compare the distribution of the predicted values with the distribution of the actual values.

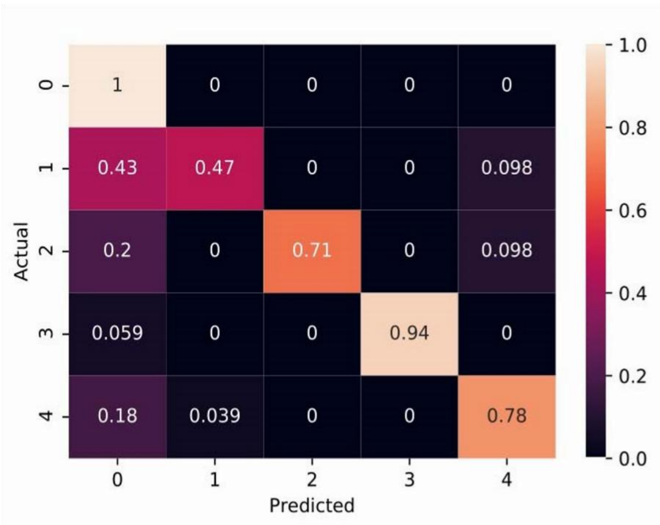

Figure 2: Confusion Matrix

Figure 2 depicts the normalized confusion matrix for the test samples using the proposed DCNN model.The specificity (recall) i.e. the coverage of positive samples shows that most of them are predicted as positive itself except class 1 (happy). Class 0 (angry) and Class 3 (neutral) are having good prediction results.

Figure 3 (a) and 3 (b) depicts the model accuracy and training loss of the model respectively for the entire epochs. From the plots, it can be observed that the model is not over fitting. The classification results of the model based on precision, sensitivity, specificity, F1-score and accuracy are provided in Table.

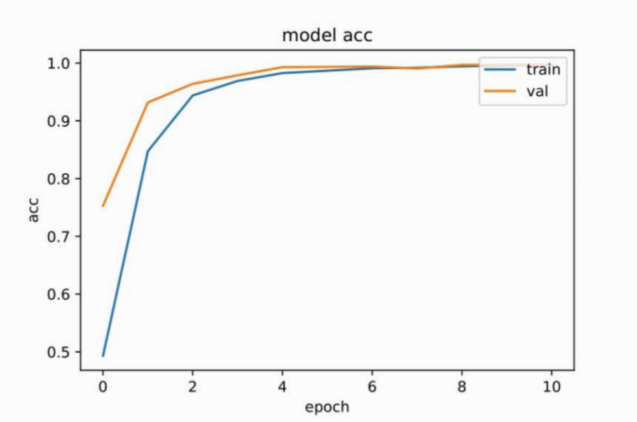

Figure 3 (a): Accuracy of training and testing data

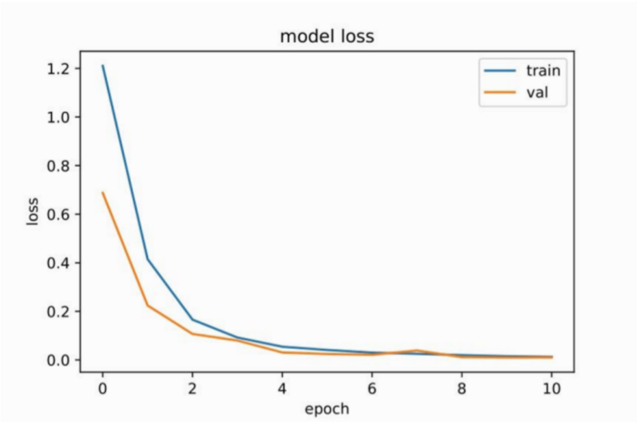

Figure 3 (b): Model loss during the training process of CNN

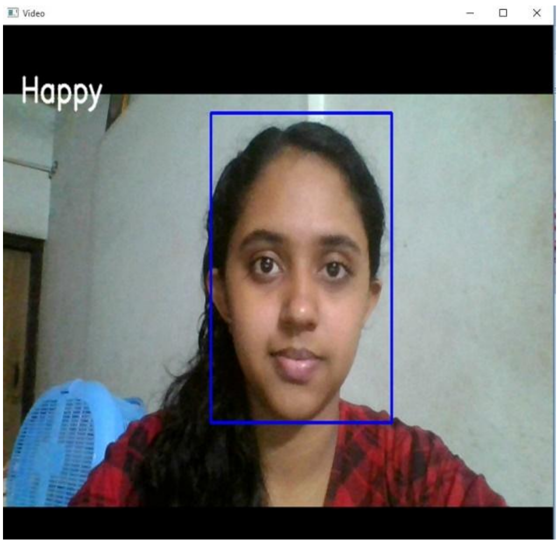

Fig: Analyse Staff

Fig : Staff Not Stressed

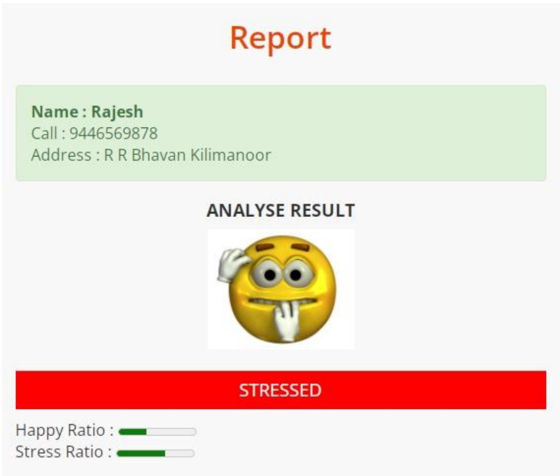

Fig : Staff Stressed

Table Iii. Classification Results Of The Model

Classes	Precision	Sensitivity(Recall)	Specificity	F1 Score	Accuracy(in %)
0-angry	0.537	1.000	0.784	0.699	82.75
1-happy	0.923	0.471	0.990	0.623	88.63
2-neutral	1.000	0.706	1.000	0.828	94.12
3- sad	1.000	0.941	1.000	0.969	98.82
4- surprise	0.800	0.784	0.951	0.918	91.76

VIII CONCLUSION

This paper proposes a 4-layer convolution network model for facial feeling acknowledgment. The model orders 7 unique facial feelings from the picture dataset-irate, disturbed, unfortunate, glad, nonpartisan, miserable, and shocked. The model has equivalent precision which pass on that the model is throwing a tantrum and is summed up to the information. The model uses an Adam enhancer to decrease the misfortune work. The work is stretched out to discover the progressions in feeling utilizing a video grouping which thusly can be utilized for various ongoing applications like input examination, and so forth The framework likewise foresee the whether or not the individual is focused on dependent on the feeling distinguished. Likewise, imagines the pressure proportion and bliss proportion.

REFERENCES

[1] G. Cao, Y. Ma, X. Meng, Y. Gao and M. Meng, "Emotion Recognition Based On CNN," 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 8627-8630.doi: 10.23919/ChiCC.2019.8866540

[2] G. Yang, J. S Saumell, J Sannie, " Emotion Recognition using Deep Neural Network with Vectorized Facial Feature" 2018 IEEE International Conference on Electro/Information Technology (EIT)

[3] Lu Lingling liu, "Human Face Expression Recognition Based on Deep Learning-Deep Convolutional Neural Network", 2019 InternationalConference on Smart Grid and Electrical Automation (ICSGEA)

[4] S. Suresh, H. T. P Mithun and M. H. Supriya, "Sign Language Recognition System Using Deep Neural Network," 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 614-618. doi: 10.1109/ICACCS.2019.8728411

[5] Lu Feng Zhou, Weiwei Fan, Qiangqiang Sheng and Mingliang Tao, "Ship Detection Based On Deep Convolutional Neural Networks For PolsarImages "International Geoscience and Remote Sensing Symposium 2018

[6] K. Bouaziz, T Ramakrishnan, S. Raghavan, K. Grove, A.A.Omari, C Lakshminarayan, " Character Recognition by Deep Learning: An Enterprise solution.", 2018 IEEE Conference on Big Data.

[7] C. J. L. Flores, A. E. G. Cutipa and R. L. Enciso, "Application of convolutional neural networks for static hand gestures recognition under different invariant features," 2017 IEEE XXIV International Conferenceon Electronics, Electrical Engineering and Computing (INTERCON), Cusco, 2017, pp. 1-4.

[8] I. Rocco, R. Arandjelovic and J. Sivic, "Convolutional Neural Network Architecture for Geometric Matching," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 39-48.

[9] J. Shijie, W. Ping, J. Peiyi and H. Siping, "Research on data augmentation for image classification based on convolution neural networks," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 4165-4170.

[10] A. Krizhevsky, I. Sutskever, and G. E. A Convolutional Neural Network Hand Tracker", in proceedings: Neural Information Processing Systems Foundation. 1995Hinton, "Imagenet classification with deep convolutional neural networks," in Proc. Annual Conference on Neural Information Processing Systems (NIPS), 2012, pp. 1106–1114.

[11] Diederik P. Kingma, Jimmy Ba, "Adam: A Method for Stochastic Optimization," 33rd International Conference for Learning Representations, San Diego, 2015