Original Article

Human Emotion Recognition System Using Deep Learning Technique

Sivakumar Depuru[1], Anjana Nandam[2], P.A. Ramesh[3], M. Saktivel[4], K. Amala[5], Sivanantham[6]

[1]Assistant Professor, Dept. of CSSE, Sree Vidyanikethan Engineering College, Tirupati, AP, India. E-mail: sivakumar.d@vidyanikethan.edu

[2]Assistant Professor, Dept. of CSE, Koneru Lakshmiah Educational Foundation, Vaddeswaram, AP, India, E-mail: Anjana.nandam@gmail.com

[3]Assistant Professor, Dept. of CSE, Sri Venkateswara College of Engineering, Kadapa, AP, India. E-mail: ramesh.asvarth@gmail.com

[4]Professor, Dept. of CSE, Sree Vidyanikethan Engineering College, Tirupati, India. E-mail: sakthisalem@gmail.com

[5]Assistant Professor, Dept of ECE, Sri Venkateswara College of Engineering, Tirupati, India. E-mail: amala.k@svcolleges.edu.in

[6]Assistant Professor, Dept of CSSE, Sree Vidyanikethan Engineering College, Tirupati, India. E-mail: sivanantham.s@vidyanikethan.edu

Abstract

Now a day's automatic emotion recognition system plays a vital role to recognize the human expressions. There are numerous applications available ranging from surveillance cameras to detect the emotions. Emotion recognition is an important task in emotion detection deep learning techniques are used for facial recognition. Images are used as input, and facial expressions are produced as output, such as happy, sad, disgusted, angry, fearful, surprised, and neutral. In this paper, we design a deep Convolutional Neural Network (DCNN) model. This model can be classifies seven various human facial emotions. This DCNN model is trained and tested using the FER (Facial Expression Recognition) Data set. The deep FER is analysing the methods are very difficulties. The dataset used for experimentation is FER challenge dataset available in KAGGLE repository. The implementation environment includes keras, tensorflow, and Open cv2 python packages. The results include the comparison of accuracy of emotion detection between training and testing phase. The average accuracy achieved was 86.05%.

Keywords: Convolutional Neural Network (CNN) Face Expression Recognizer (FER), Artificial Intelligence (AI), Machine Learning, Deep Learning.

DOI: 10.47750/pnr.2022.13.04.141

INTRODUCTION

The advancement and utilization of computer systems, networks, and software are increasing enormously. These systems have a crucial role in our day-to-day life and they made human survival much easier. The human facial expressions and emotion recognition system assumes plenty of significance in this era because it can apprehend the behavior, feelings, intents of humans, etc. the method of conventional have limited speed and low accuracy while the human facial expressions and emotion recognition system utilizes deep learning has proved to be the better one[13]. The convolutional neural network model is build to recognize the various human facial expressions and it may be utilized for applications like customer feedback analysis, face unlocking etc.

In this research, we present a model for finding various emotions like anger, disgust, neutral fear, happiness, sadness, and surprise utilizing facial images. Earlier research utilized deep learning technology to make models of human facial expressions based on emotions to be finding emotions [3]. Emotion is a nervous system-related mental state associated with feelings, perceptions, behavioral reactions, and a degree of gratification or displeasure [4]. Human facial emotion recognition system developed with the various machine learning techniques.

Machine learning plays a vital role in the field of computer science and it's an emerging technology. That is treated to have an effect of 95% in the coming 3 years. Deep learning is a subset of machine learning that makes use of artificial neural networks, which are algorithms inspired by the human brain. A Convolutional Neural network (CNN) is a type of deep neural network in which convolution is used as the mathematical operation. Because the dataset is made up of images, the system performs image recognition using a 2D CNN. The proposed new deep convolutional neural network has been trained not only to classify seven various human facial expressions but also with good accuracy. The model is trained by using the manually collected dataset with a mobile phone camera.

Related Work

Y. Ma and Cao. G, et al [5] proposed a CNN (Convolution Neural Network) model used to identify human expressions and emotion from the Electro cardiogram (ECG) dataset, which can then be used to categorize brain signals as well. The system provides around 83% accuracy on testing. S Suresh et al [6] introduced a sign language recognition system, which classifies six various sign languages using a Deep Neural Network (DNN). Both models with various optimizers (Adam and SGD) are compared and the Adam optimizer is used to find more accuracy with this model. ShreyModi and Mohammed Husain Bohara [8] used a CNN-based emotion recognition system that uses a feature map and max pooling layers, as well as a fully connected output layer, to detect a person's emotions. Akriti Jaiswal, A. Krishnama Raju, and et al [7] proposed a deep learning architecture depending on convolutional neural networks (CNN) for recognizing human emotions from images. This model achieves accuracies of 70.14. K. Bouaziz et al [9] demonstrated an analytics workflow that combines image recognition techniques and tools. The proposed model uses CNN architecture to classify different handwriting. Krizhevsky et al [10] used to publish a seminal paper on automatic image classification in general [8]. This research demonstrates a deep neural network with functionality similar to that of the human visual cortex. Using the CIFAR-10 dataset and a self-developed labelled array of 60,000 images divided into 10 classes, a model for categorising objects from pictures is obtained. Another significant outcome of the research is the visualisation of the filters in the network, which allows the model to be evaluated in terms of how it breaks down images. F. Zhou et al [11] proposed a deep convolutional neural network model used to recognize ships in the movement for Polarimetry Synthetic Aperture Radar (POLSAR) Images. This model utilizes a Faster Region based Convolutional Neural Network (FRCNN) method to find a ship of various sizes. NASA/JPL AIRSAR dataset used by validated the model.

Proposed Model

A neural network is a collection of algorithms that mimic the human brain and use these algorithms to find relationships between data to find solutions. Convolutional Neural Network is a type of Neural Network in which Convolutional is the mathematical operation utilized to find the difference between the data [1]-[2]. When it comes to complex problems like pattren recognition, image and video classification so on, traditional neural networks fail miserably, but Convolutional Nueral Network has had great success in these applications, yielding high accuracy.

Convolutional Neural Network contains four layers and these four layers extract the features from the input images together. After extracting the features from the input images using algorithm learns and it can be represented by each convolution filter[12].

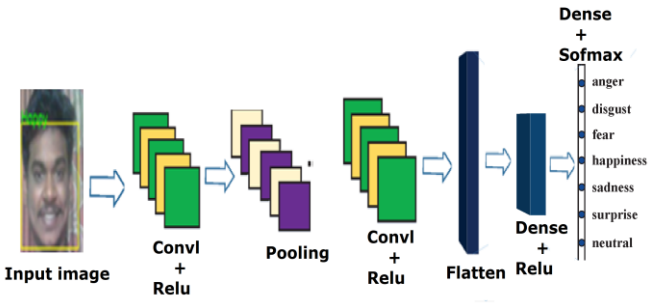

Fig:1 Human emotion recognition model using Deep convolutional Neural Network Architecture

Convolutional layers are made up of small patches that transform the total image depending on the filter values. Equation (1) is the formula for generating the feature maps, which are the output of the convolutional layer [15].

$$G[a,b]=(i*j)[a,b]= \Sigma m \, \Sigma k \, i[\, n, k] \, j[\, a - n, b - k] \text{---(eq.1)}$$

From the above equation i is represent the input image and j is represent the filter, (a,b) is the size of generated matrix result. The output of the convolutional layer is transfer to a pooling layer, which reduces its size without loss of information. The flatten layer is converts these 2-D array to 1-D, which can then be supplied to the CNN for classification [16]. This CNN employs the process of backpropogation method, which adjusts the weights based on the errors, lowering the loss function. The 'w' is updated with the help of (eq.2)

$$Wm = Wm + \Delta Wm \text{ – (eq.2)}$$

From the above equation Wm is represents the weight and ΔWm as the below equation (eq.3).

$$\Delta Wm = v * dE / dWm * xm \text{ – (eq.3)}$$

In this above equation v is represents the learning rate and E is Loss function, xm is represents the input. The proposed design is human facial emotion recognition model is depicted into figure 1. This model used to convolutional layers with dropouts after each and every convolutional layer [17] [18].

Proposed Model of Convolutional Neural Network

The input image is converted 32X32X32 into standard size of 30X30X32 is given to the initial convolutional layer and the output from the convolution layer is called the feature map (F-Map) and its further passed to an activation function called rectified layer unit(ReLu) which makes negatives to zero values and positive values unchanged. The F-Map(featrue map) is passed to the pooling layer and this layer is reducing the pool size without loss of information. The overfitting is reduced by the use of dropout layer. This mechanism again perform continues for both input convolution layer and output as well.

Finally, a 2-dimensional array is generated with some random values from the feature map then converting this 2-dimensional array to a 1-dimensional vector by use of flatten layer and passed this as the input of the neural

network is represented by the dense layer. Here two layers of the neural network are used one layer is represented as the input and another layer is represented as the output, because seven classes are to be classified. Softmax is the activation function used in the output layer. It generates the probablistic output for each and every class in this system, which can be built using KERAS deep learning library.

DATA SET

The FER dataset is available in Kaggle repository. This FER dataset used in our model is includes joyful, depressed, disgusted, angry, scared, surprised, and neutral. Each and every image has variable amount of pixel size. As shown in the bellow dataset split table 1 [19]. Each class has the same number of training samples, ensuring that they are not biased so that sample size is consider each and every train-test- validation.

Table 1: FER Dataset Split-up table

Num. of classes	**7**
Num. of training images	**28821**
Num. of validation images	**7066**
Total Num. of images	**35887**

RESULTS & DISCUSSION

The model can be implemented by using a python programming language. this model is also simulated in a jupyter notebook and combines the convolutional neural network layers to construct the model, compiling and fitting the model, Keras that runs on top of TensorFlow, is utilized as the deep learning library. Scikitlearn is a package and utilized to identify the confusion matrix that provides the model's precision, accuracy, specificity, recall, etc. The confusion matrixes, as well as other graphs such as loss and accuracy graphs, are plotted using Matplotlib and Seaborn. The FER (Facial Emotion Regonition) image dataset is used to train CNN, That uses the Adam optimizer and the loss function is categorical cross entropy. Relu and Softmax are the activation functions; Adam is the optimizers 40 epochs are used with learning rate 0.0001.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 48, 48, 64)	640
batch_normalization (BatchNo	(None, 48, 48, 64)	256
activation (Activation)	(None, 48, 48, 64)	0
max_pooling2d (MaxPooling2D)	(None, 24, 24, 64)	0
dropout (Dropout)	(None, 24, 24, 64)	0
conv2d_1 (Conv2D)	(None, 24, 24, 128)	204928
batch_normalization_1 (Batch	(None, 24, 24, 128)	512
activation_1 (Activation)	(None, 24, 24, 128)	0
max_pooling2d_1 (MaxPooling2	(None, 12, 12, 128)	0
dropout_1 (Dropout)	(None, 12, 12, 128)	0
conv2d_2 (Conv2D)	(None, 12, 12, 512)	590336
batch_normalization_2 (Batch	(None, 12, 12, 512)	2048
activation_2 (Activation)	(None, 12, 12, 512)	0
max_pooling2d_2 (MaxPooling2	(None, 6, 6, 512)	0
dropout_2 (Dropout)	(None, 6, 6, 512)	0
conv2d_3 (Conv2D)	(None, 6, 6, 512)	2359808
batch_normalization_3 (Batch	(None, 6, 6, 512)	2048
activation_3 (Activation)	(None, 6, 6, 512)	0
max_pooling2d_3 (MaxPooling2	(None, 3, 3, 512)	0

Figure 2: depicts the analysis of the proposed FER model built using Keras

Figure: 3 depict the accuracy and loss of the value

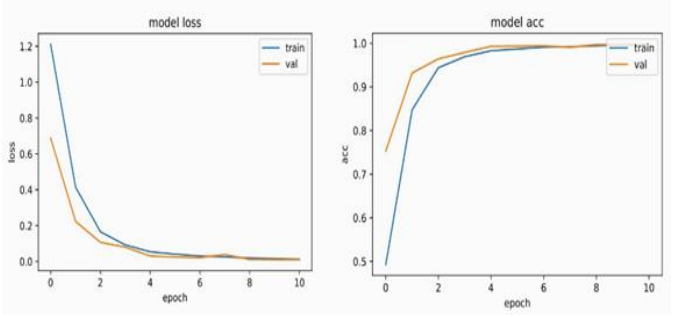

Figure: 4 depict the accuracy of training and validation (testing) data as well as model loss during the CNN training process.

The above figure 4 shows the model accuracy and loss and we can see that loss decreases with the time and accuracy

increases in both training and validation.

After executing the OpenCV model the facial expressions and emotions are detected as given below figures 5, 6, 7 indicates the 3 major emotions. Those human facial expressions and emotions are happy emotion is recognized as depicted in the following figure 5, as well as the remaining emotions are identified as the depicted in the figure 6 and 7.

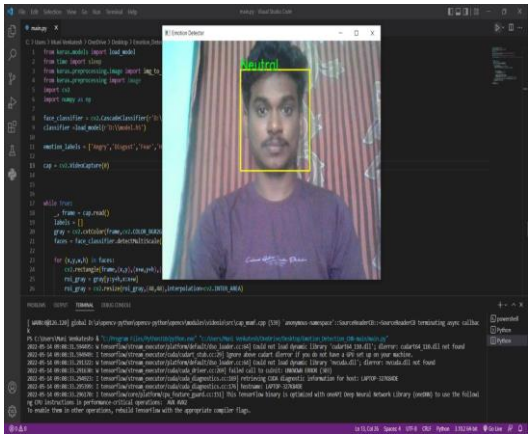

Figure: 5 Neutral emotions

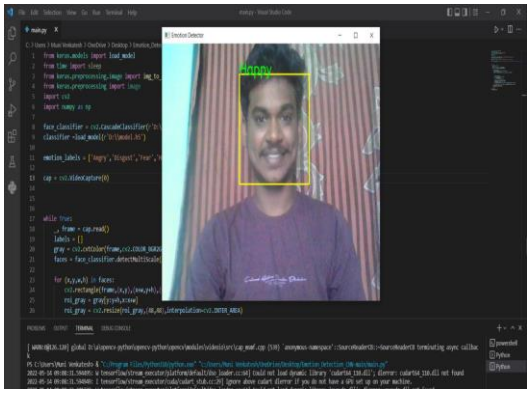

Fig: 6 Happy emotions

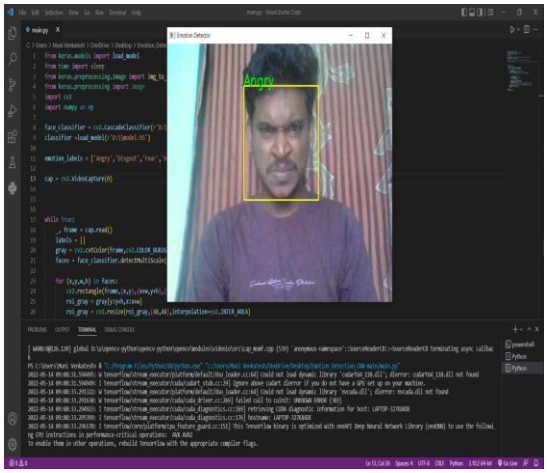

Figure: 7 angry emotions

CONCLUSION

In this paper, we propose two layers of convolutional neural network for extracting Human facial expressions and recognizing emotions. This model classifies the 7 various human facial expressions from the FER Image dataset used and 35887 images. Human emotions are considered as happy, sad, disgusted, angry, fearful, surprised, and neutral. The model has similar training and validation accuracy, indicating that it has the best fit and is generalizable to the data. The model incorporates an Adam optimizer to minimize the loss function and it has been validated to have an accuracy of 86.05%. The research work is to be expanded to detect human facial expressions and emotions changes using a video sequence, which can then be used for various real-time applications like feedback analysis and so on. Combines these systems with any other electronic devices for their effective control.

REFERENCES

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 Conference on Electronics, Electrical Engineering and Computing (INTERCON), Cusco, 2017, pp. 1-4.

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.

E. Correa, A. Jonker,M.Ozo, andR.Stolk, “Emotion recognition using deep convolutional neural networks,” Tech. Report IN4015, 2016.

Li and W. Deng, “Deep facial expression recognition: A survey,” arXiv preprint arXiv:1804.08348, 2018.

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.

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

Akriti Jaiswal, A. Krishnama Raju, Suman Deb “Facial Emotion Detection Using Deep Learning” 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020 pp. 978-1-7281-6221-8/20/$31.00 ©2020 IEEE.

Shrey modi et al. " Facial Emotion Recognition using Convolution Neural Network Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021) IEEE Xplore Part Number: CFP21K74- ART; ISBN: 978-0-7381-1327

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

A. Krizhevsky and G. Hinton. “Learning multiple layers of features from tiny images”, 2009.

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.

Siva Kumar Depuru, Dr.K.Madhavi, Autoencoder Integrated Deep Neural Network for effective analysis of malware in distributed Internet of Things (IoT) Devices, The International journal of analytical and experimental modal analysis, Page No:226-232,Volume XI, Issue XII, December/2019.

Sivanantham, S., et al. "Association Rule Mining Frequent-Pattern-Based

Intrusion Detection in Network." COMPUTER SYSTEMS SCIENCE AND ENGINEERING 44.2 (2023): 1617-1631.

M. Ramu and et al "Study on Potential AI Applications in Childhood Education" International Journal of Early Childhood Special Education (INT-JECS) ISSN: 1308-5581 Vol 14, Issue 03 2022 doi: 10.9756/int-jecse/v14i3.1215.

P. Yogendra Prasad,Dumpa Prasad,D Naga Malleswari (2022).Implementation of Machine Learning Based Google Teachable Machine in Early Childhood Education. International Journal of Early Childhood Special Education (INT-JECSE); 14(3):1308-5581.

Sakthivel M, Sivanantham S, Kamalraj R & Krishnamoorthy V (2022). An Analysis of Machine Learning Depend on Q-MIND for Defencing the Distributed Denial of Service Attack on Software Defined Network. International Journal of Early Childhood Special Education. 2022; 14(05): 3769 – 3776.

Lakshmi Haritha.M & K. Ramani. Impact of Deep Learning on Localizing and Recognizing Handwritten Text in Lecture Videos. International journal of Advanced Computer Science and Applications;12(4), 2021.

CH Prathima, R. Anusuya, M. Ram Kumar Prabhu (2022). Comprehensive Design Analysis of Digital Marketing in Agriculture Sector. International Journal of Early Childhood Special Education ; 14(5):2022.

P. Dhanalakshmi et al. (2022). Application of Machine Learning in Multi-Directional Model to Follow Solar Energy Using Photo Sensor Matrix. International journal of Photoenergy; 9:1-9.