International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 03, 2020
ISSN: 1475-7192

PREDICTION OF FACIAL EXPRESSION USING DEEP LEARNING

[1]Dr. R.PUNITHAVATHI M.E., Ph.D ,[2]T. NAVEEN KUMAR ,[3]S.D.PONGOWTHUM,
[4]A. RAHUL ,[5]S. SANTHOSKUMAR

ABSTRACT--*Recommendation suggestion is one of the biggest problems for the modern-day applications and its developers. It deals with the content delivery and advertisement, where matching the service provider and the client who requires the service are happening. Now a days, the history and the wish list of the application user are not enough to satisfy the requirements of the user. The real time emotion of human needs to be founded out in order to provide the right content to the right user. It can be made possible by building a convolutional neural network using deep learning, where the model will be trained with several thousand images and then input is passed then output is returned.*

Keywords--*Deep Learning*, Facial Expression, Convolutional Neural Network, Real time emotion, Image resizing.

I. INTRODUCTION

Real time emotion can be obtained through facial expression of the user. Now consider a music player a decade ago, the people will download the songs in a device then the music player is used only to play music. But now a days it is not necessary to download the songs in a device. The music player itself plays through internet, even now user don't need to search and choose the song, the application itself suggest songs based on the song popularity and watch history of the user. The songs in such players are not randomly generated there is a powerful machine learning algorithm behind the scenes, which computes not a single list it prepares for each user of the application. So, what will be future of music players and the application like it which requires perfect suggestion for the users? Surely there will be a prediction of real time emotion of the users. It is possible by predicting facial expression. Through deep learning [1] by training over several thousand images it is possible to predict the emotion of human in real time.

II. RELATED WORKS

Real time facial expression in the sense the system needs to be capture the facial expression in live camera. In real time also it may look like live, but it isn't, system will take several pictures in a second, so it looks live. In camera feed, it may capture anything which may be an object, animal, etc., To predict facial expression, the first

[1] PROFESSOR AND HEAD Department Of Information Technology, M Kumarasamy College Of Engineering, Tamil Nadu, India
[2] B.Tech Student Department Of Information Technology, M Kumarasamy College Of Engineering, Tamil Nadu, India
[3] B.Tech Student Department Of Information Technology, M Kumarasamy College Of Engineering, Tamil Nadu, India
[4]B.Tech Student Department Of Information Technology, M Kumarasamy College Of Engineering, Tamil Nadu, India
[5] B.Tech Student Department Of Information Technology, M Kumarasamy College Of Engineering, Tamil Nadu, India

Received: 03 Dec 2019 | Revised: 21 Jan 2020 | Accepted: 18 Feb 2020

task is to extract face from the image. It only needs to extract human face not the other substance that appears in the image. So, Haar cascade algorithm [2] is used to extract human face from the available image. There Haar cascade classifier for frontal face is used in order to extract human face from the captured image.

There are many algorithm and datasets are already available to predict the emotion in real time. But they are poor in recognizing it when compared to level where human can understand the emotion from one to another. Through proposed system it is possible to predict the human emotion in real time in accurate.

III. DEEP LEARNING:

Deep learning is a subset of machine learning which is based on artificial neural network where the learning can be done based on supervised, unsupervised and reinforcement learning. The models in deep neural networks are classified into convolutional neural network and recurrent neural network. The neural network consists of multiple hidden layers between the input and output layer

IV. CNN:

Convolutional neural network or CNN [3] or ConvNet is a class of deep neural network which generally takes input as image which forms weights and biases with various aspects can able to differentiate between various classes

V. PROPOSED METHODOLOGY

The Proposed system combines a collection of multiple datasets. All such datasets combined then it is formed as a huge single dataset. The proposed system contains only five emotions such as happy, sad, angry, surprise and neutral. Remaining two emotions such as fear and disgust will be removed for improvement.

Pre-processing:

Pre-processing [4] only has few changes are going to happen but the impact in improving the code is going to be higher.

Received: 03 Dec 2019 | Revised: 21 Jan 2020 | Accepted: 18 Feb 2020

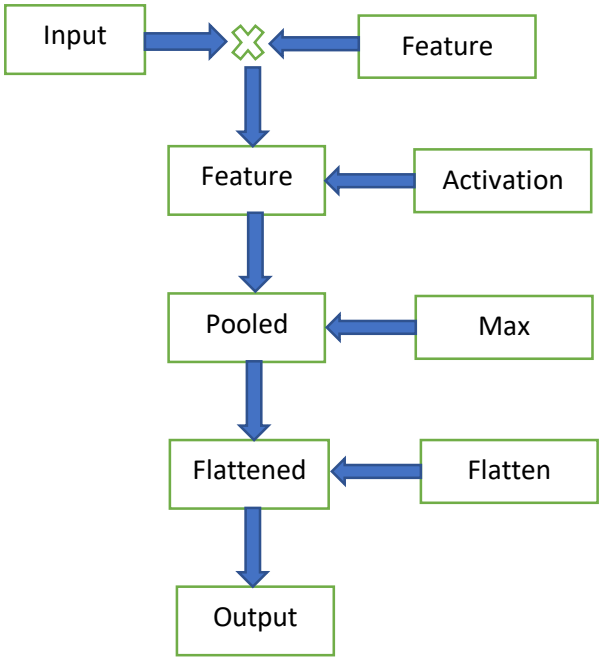

Fig. 1. SYSTEMATIC WORK FLOW

Removal of noisy images:

The collected dataset consists of several thousand images, but there may be some images that will be irrelevant to the classification but exists on the dataset. It needs to be removed.

Limiting the outcome

In collected dataset, there are totally seven kind of emotions are trained and predicted. But in the proposed system it going to be only five kinds of outcomes. In such case fear and disgust will be removed as it replicates angry and sad. And it improves the accurate prediction of the emotion.

Image resizing:

Dataset when combined may have images with different pixel values it may led to another problem because the image captured in real time from a camera [5] feed needs to be converted to the pixel value that needs to be as same as the pixel value of all images in the dataset.

Layers in CNN:

Conv2D is a Convolutional layer that creates a convolutional kernel which will be the first layer in the CNN. The role of convolutional layer is to extract features from the input image that is obtained from the real time camera feed. It extracts features by learning from pixel data of input image. Features can be extracted from various forms such as identity detection [6], edge detection, sharpening and blurring.

Activation Function in neural network is the layer which gets the input feed from the previous layer. The most common used activation functions are Rectified Linear Unit, Exponential Linear Unit, Leaky ReLu, Sigmoid, Softmax and tanh. In this system Softmax is used which gets logits data as input and returns probabilities as an

Received: 03 Dec 2019 | Revised: 21 Jan 2020 | Accepted: 18 Feb 2020

output. That probability lies between 0 to 1. It is founded by dividing exponential of the input to sum of all exponential to the input.

Batch normalization normalizes the input layer by adjusting and scaling the data impacted by activation function. Normalization is process which enhances of stability of the network by calculating the mean and variance. Batch size is directly proportional to the accuracy and reduce in loss and indirectly proportional to time consumed to train the model.

Pooling [7] is strategy to reduce the size of the data in model without changing its feature. Pooling is strategy to reduce the size of the data in model without changing its feature. There are two types of pooling they are min pooling and max pooling. In min pooling the minimal value in the desired matrix will be chosen. In max pooling the maximum value in the desired matrix will be chosen. Here as it need to deal with image and recognize pattern with the images max pooling is suitable. So max pooling of 2X2 is chosen.

Dropout layer is built to avoid overfitting in CNN. When a large dataset is taken and a CNN in meant to build, there is high of some amount of data that is irrelevant to be there in such dataset but it has a high potential to affect purpose of mislead to a wrong class or prediction.

Dense layer can also be described as fully connected layer. Dense layer [8] makes the connection between nodes from previous layer to next layer and it decides how many nodes from previous nodes need to be connected to the next layer and also it describes in which order such nodes need to be connected to one another.

Categorical cross entropy is the algorithm which is used in order to measure the loss when system needs to be prediction output with more than two classes. Instead if the system needs to predict among two values binary cross entropy should be used.

Stochastic gradient descent algorithm:

Gradient decent algorithm [9] is an optimization technique used to optimize the neural network. The primary objective of the gradient decent algorithm is not optimizing the code. It needs to be more efficient in predicting the desired output. Meanwhile the gradient decent algorithm choses suitable value for all the variants passed. But stochastic gradient descent algorithm chooses for every combination of variants passes so only takes more time, but it is efficient than the other gradient descent algorithms.

Adam optimizer:

Adam is an adaptive learning rate optimization algorithm. It is built to determine the mean and variance to choose the best. So, it is efficient than the other algorithms.

VI. EXPERIMENTNAL ANALYSIS

After creating a powerful convolutional neural network of 43 layers and model trained with 24,254 images (Angry – 4004, Happy – 7255, Neutral – 4965, sad – 4830, Surprise- 3200), with batch size 32, pixel value of 48 X 48 and trained 100 epochs the model gains the accuracy of 84.3%. Fig. 2. represents the sample output image of the facial expression. The process of constructing the model 80% of data used for training and 20% of data used for validation.

The system also recognizes the facial expression better than the expectation when it is tested on various human classification and various circumstances. This proposed approach is done by Anaconda Platform. Anaconda is an opensource software which comes with various development environments such as Jupyter notebook, Jupyter lab and Spyder, Which is a best user-friendly software to develop machine learning project in Python and R. The following Figure 2 illustrates the detected emotions from the given input face image.

Fig. 2. Resulted Emotion Detection in Face

The various packages and library files available at anaconda can be installed, updated, and maintained easily.

VII. CONCLUSION AND FUTURE WORK

The proposed system results to the successful development of facial expression detection with various facial expressions such as happy, sad, angry, surprise and neutral with the better accuracy. The results obtained are better than the expectation. The system is also fast enough and even giving better result when compared to the existing system. So, it will be a better replacement for the existing system. The result obtained from the system can be improved by using high quality datasets in increasing the count of images used. Even then the performance is directly proportional to the algorithm and the methodology chosen to compile the model. So, model build should be lite and good enough to predict the facial expression with high accuracy.

REFERENCES

1. Hepeng Zhang, Bin Huang, Guohui Tian, "Facial Expression Recognition Based On Deep Convolution Long Short-Term Memory Networks Of Double-Channel Weighted Mixture", 2019.
2. Junkai Chen, Zenghai Chen, Zheru Chi, Member, IEEE and Hong Fu, “Facial Expression Recognition in Video with Multiple Feature Fusion”, 2016.
3. Mingjing Yu, Huicheng Zheng, Zhifeng Peng, Jiayu Dong, Heran Du, “Facial expression reco4gnition based on a multi-task global-local network”, 2020.

Received: 03 Dec 2019 | Revised: 21 Jan 2020 | Accepted: 18 Feb 2020

4. Pipit Utami, Rudy Hartanto, Indah Soesanti, "A Study on Facial Expression Recognition in Assessing Teaching Skills: Datasets and Methods", 2019.
5. M. SHAMIM HOSSAIN, GHULAM MUHAMMAD, “An Emotion Recognition System for Mobile Applications”, 2017.
6. D. Yang, Abeer Alsadoon, P.W.C. Prasad, A. K. Singh, A. Elchouemi, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment", 2018.
7. Yingsheng Ye, Xingming Zhang , Yubei Lin, Haoxiang Wanga, “Facial expression recognition via region-based convolutional fusion network” 2019.
8. Yan Yan, Zizhao Zhang, Si Chen, Hanzi Wang, "Low-resolution Facial Expression Recognition: A Filter Learning Perspective", 2019.
9. Zhongzhao Xie, Yongbo Li, Xinmei Wang, Wendi Cai, Jing Rao, Zhenzhu Liu, “Convolutional Neural Networks for Facial Expression Recognition with Few Training Samples”. 2018.