An Emotion Intelligence Auto Music Play for Stress Burst Based on Nomogram Model for Quarantine and Self-IsolationPeople Using Deep Learning in Pandemic

Dr. I. Manimozhi [1] , Dr. T. K Sateesh[2]
[1]Associate Professor, [2]Professor & Principal
Computer Science & Engineering
East Point College of Engineering & Technology
Banglore, India,

Abstract
Emotions are one of the normal habits to talk sentiments and people\'s sentiments can be used in delight and Human Machine Interface (HMI) fields. Information from facial expressions is distributed in specific areas of the face and each of them has distinctive data so that mouth and eyes consist of greater facts than cheek and brow. To develop an emotion recognition system for the cognitive components of automatic thoughts and internal dialogue on the appraisal of stressors through their facial expression. To develop an intelligent device the usage of CNN that may without problems understand the facial expression of the consumer photograph recognition primarily based on deep getting to know models. To segregate and play music automatically based on their emotional intelligence feeling when they are in Quarantine /Isolation /Lonely situations. No program or application can predict the user’s emotion and play music accordingly except the manual selection playlist.

Keywords: stress burst CNN model, Feature extraction, Prediction, Emotions

I. Introduction

Feelings are a significant property of people and are fundamental for powerful collaborations among the general public. People correspondence [1]can be either verbal of nonverbal, which it has been shown the vast majority of them allude to nonverbal correspondence. In nonverbal correspondence, feeling assumes powerful part since it passes on people feeling about the subject, and in the brain science research it is demonstrated that looks is more compelling than verbally expressed word in discussion.

Programmed feeling demeanor acknowledgment incorporates three stages: face picture acquisition,[2] highlight extraction, and facial feeling appearance acknowledgment. In the ideal removed provisions, inside class varieties of articulation ought to be least while between-class varieties ought to be greatest. On the off chance that the extricated highlights are not suite for task close by and need more data, even the best classifier might be ineffective to have a best execution.

Component extraction for feeling acknowledgment can be separated into two methodologies: Mathematical element-based techniques and appearance-based strategies. In the main techniques, area and state of parts of the face like the eyes, mouth, eyebrows, and nose are thought of, while in the subsequent strategies, specific districts or entire of face are thought of. In light of separating articulations [3] highlight space is a troublesome issue, so articulation acknowledgment is as yet a difficult assignment for PCs. A few issues might be because of that, extricated highlights from two countenances with equivalent demeanor might be unique, while separated elements from one face with two appearances might be equivalent, or some appearance, for example, "dread" and "pitiful" are basically the same.

hese days, with the types of progress in the space of advancement diverse music players [4] are sent with features like exchanging the media, speedy sending it, streaming playback [5]with multicast streams. Yet these arrangements satisfy the crucial necessities of the customer, yet one requirements to actually surf for the tune from a huge plan of tunes, according to the current circumstance and attitude. This is a drawn-out task [6] that has serious room for improvement and resistance. The guideline objective of this work is to encourage an adroit structure that can without a lot foster a shrewd framework that can without much of a stretch perceive the feeling through look [7] and appropriately play a music track dependent on that specific articulation/feeling perceived.

II LITERATURE REVIEW

Vibha. V. Salunke and Dr. C. G. Patil [1] proposed to plan a misleadingly savvy framework equipped for feeling acknowledgment through looks of obscure individuals. The association in this paper consists of three convolutional layers each followed by max pooling and ReLU. The association is ready on FER2013 dataset [8] and took a stab at RaFD dataset subsequently giving a wide extent of planning pictures to the association, with the objective that it can vanquish the crucial issue of affirmation of dark appearances. This paper basically centers around neural organization based misleadingly shrewd frameworks equipped for determining the feeling of an individual through photos of their face.

Amine Trabelsi and Claude Frasson [2]explores the feasibility[9] of furnishing PCs with the capacity to anticipate, in a setting of a human PC connection, the likely client's feeling and its force for a given feeling inspiring circumstance. An Information Part which contains every one of the information and data expected to perceive a feeling and an Expectation Part made out of the KNN and ANN calculations.

Sanghyuk Kim, Gwon Hwan An and Suk-Ju Kang propose Haar-like components, and the area of interest is reset to reduce the variable of appearance changes. The FER association isolates histogram of organized points (Crowd) features from each facial locale, and a while later, a help-vector machine is performed to bunch the last look. In the exploratory results, the structure unequivocally apparent the vibe of somebody specifically, and the proposed system had the F1 score of 0.8759.Jianzhu Guo and Zhen Lei First, the quantity of public information bases in this field is restricted. Second, current accessible public datasets have few classifications which might cover simply a little part of all conceivable compound feelings.

Third, the marks furnished with Feeling Net dataset are identified with consequently recognized Activity Units (AU),[10 which are utilized for compound feeling investigation. Albeit the AUs can be changed over to intensify feeling classification, the outcomes probably won't be precise because of mistakes presented by the AU acknowledgment module. Chao Qi and Min Li demands Looks Acknowledgment Dependent on Insight and Planned Parallel Examples. As of late, numerous strategies for programmed look acknowledgment have been proposed, for example, head part investigation (PCA), direct basis examination (LDA) versatile bundle diagram coordinating (EBGM), autonomous part examination (ICA),[11] two-dimensional head part investigation (2D-PCA), artificial neural organizations, implanted secret Markov models (EHMM), Gabor wavelets, and others.

III PROPOSED METHOD

Pre-handling, highlight extraction, and forecast are the three stages of the undertaking as displayed in Figure 1. To make the model more strong, the dataset is increased. Changes in shear, zoom, and x or y pivot directions are all important for the increase.

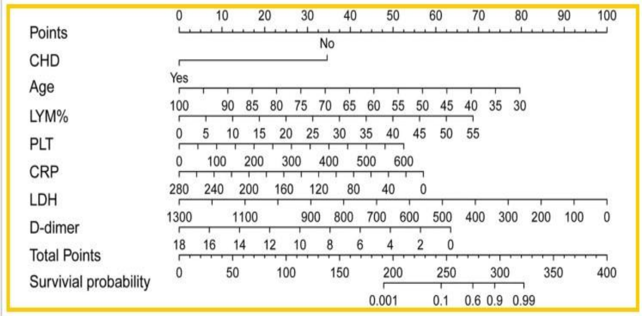

Fig1 Nomogram model

The gender is delegated look in the expectation stage utilizing a CNN model [12] with input from a picture improvement or the picture can be straightforwardly recovered from a Cameras . Element extraction will be examined in more detail in a later meeting.

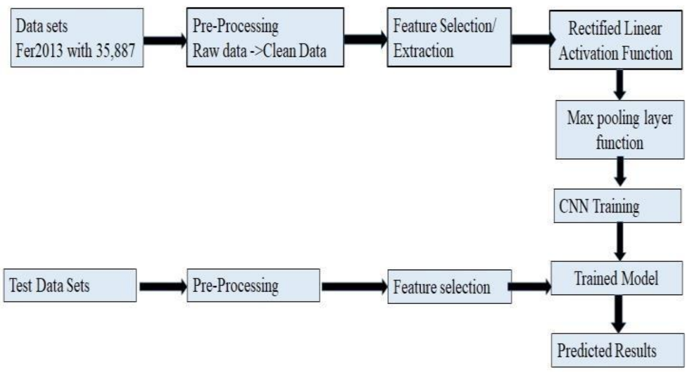

Fig 2 System Architecture

A. *Data sets*

The model was developed using a variety of male and female Facial pictures. The model will classify the input image when it has been trained. In each training and test set, there are both male and female facial expressions . The flow diagram of our system is shown in Figure 2 and steps involved in the CNN layer [13]are given below.

The steps involved-

- Define the problem and assemble a dataset.
- Augment the data and prepare it for training which prevents the data from over fitting, removes duplicates and performs normalization process.
- Choose and train the model.
- To gauge the model's objective performance, evaluate it using some metric or a set of metrics.
- Update the parameters of the model through Adam optimization.
- Tune model parameters for improved performance with more number of epochs.
- Make predictions using further data (test set).

B. Preprocessing of Data

It is carried out via using an image-enhancement software program. Image statistics augmentation is a way for falsely expanding the scale of an instruction dataset via adjusting the pics inside the dataset.

The Keras deep learning neural network toolkit's ImageDataGenerator class allows you to fit models with picture data augmentation. Image Data Generator class permits you to fit models with picture information expansion. The Image Data Generator class in Keras simplifies it to add information to the pictures. Normalization, turn, shifts, flips, brilliance changes, and more are among the expansion decisions accessible. At every age, the Image Data Generator class ensures that the model gets new varieties of the pictures. At each epoch, the ImageDataGenerator class guarantees that the model receives new variations of the images.

C. Feature Extraction and Classification

For clustering, the clustered data has been taken. In this technique, initially the agent of swarm is distributed in search space. The dataset that is being utilized in this undertaking is Fer2013. Fer2013 is an open-supply dataset this is first made for a continuous challenge through Pierre-Luc Transporter and Aaron Courville, then, at that point shared freely for a Kaggle competition, in a matter of seconds before ICML 2013 [14]. This dataset consists of 35,887 pictures. A preparation set of 35887 photographs and a test set of 460 pictures,, each with a width and height of 64 × 64 pixels. A Fully Connected Neural Network is formed for the dataset before the CNN model is trained.

The least complex technique for building a model in Keras is successive. It permits you to assemble a model layer by layer. The \ 'add ( )\' strategy is utilized to add layers to the model. The elements of the approaching picture are first smoothed to 1D. The pixel esteems in the picture are standardized. Thick layers are utilized to build the model design. Adam improvement is a stochastic angle drop strategy dependent on the versatile first-and second-request second assessment. Adam consolidates the best elements of the AdaGrad and RMS Prop strategies to make an advancement method for uproarious issues with meager inclinations. Adam is easy to set up, and the default arrangement boundaries function admirably for most of issues. The model is trained on data created batch-by-batch by fit generator. For efficiency, the generator runs in parallel with the model. The model is then trained to make predictions. The dimensions of the incoming image are first flattened to 1D. The pixel values in the image are normalized. Dense layers are used to construct model architecture. The model is then trained to make predictions.

The CNN layers are: Convo2D, maxPooling2D, flatten and dropout. The most frequent type of convolution layer is the 2D convolution layer (conv2D layer)[15] .A filter or kernel in a conv2D layer has a height and breadth. Because filters or kernels are often smaller than the input image, we transfer them across the entire image.

FLOWCHART

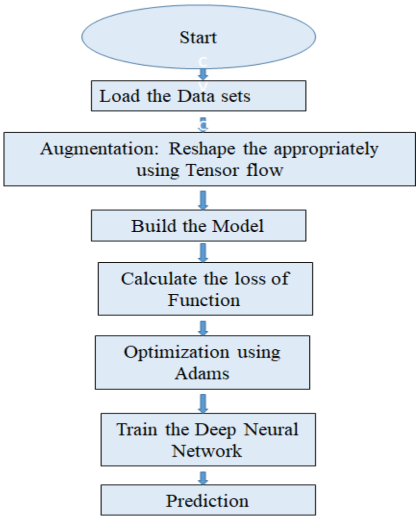

Fig 3. Proposed Process flow

Conv2D Filters are applied to all three channels in an image (Red, Green, and Blue). Each channel's filters may also be different. The final convoluted image is created by adding the individual convolutions for each channel. A feature map is the result of a convolution operation on a filter's output.

For 2D spatial information, the max-pooling manner is utilized. It takes the largest price for each channel of the enter over an input window to downsample the enter alongside its spatial dimensions (top and breadth) (of a length decided by using pool size). Consider the subsequent equation.

Classifier.Upload(MaxPooling2D(pool_size=(2,2)))   (1)

Pool_size is an integer or a tuple of numbers that represents the window length over which the maximum should be taken. If only one integer is furnished, the window duration for both dimensions could be the same.

The input is flattened by means of a knocking down layer, which has no impact at the batch size.

During the schooling period, the Dropout layer prevents overfitting by using randomly setting input gadgets to zero with a rated frequency at each step. Inputs that are not set to 0 are scaled up via 1/ (1 - charge) to maintain the whole quantity. The result is shown in Figure 7.

IV IMPLEMENTATIONS

The dataset this is being used in this challenge is Fer2013. Fer2013 is an open-supply dataset it really is first, created for an ongoing challenge through Pierre-Luc Carrier and Aaron Courville, then shared publicly for a Kaggle opposition, hastily earlier than ICML 2013. This dataset consists of 35.887 grayscale, 48x48 sized face pics with various feelings, all categorized.

Emotion labels in the dataset:

- 0: 4593 photos - Angry
- 1: 547 photos – Disgust
- 2: 5121 pictures – Fear
- 3: 8989 pix – Happy

4: 6077 photos – Sad
5: 4002 pix – Surprise
6: 6198 photos – Neutral

The above is the variety of attributes in the dataset. There are 35887 instances gift in the dataset.

Convolutional neural network version.

ArtificialNeural Networks are used in different forms of duties like photos,audio, word s. Different styles ofNeural Networks are used for distinct purposes,for illustration for pro gnosticating the gathering of words we use Intermittent Neural Networks redundant exac tly an LSTM, also for photo type we use Convolution Neural networks. In this blog, we're going to make a abecedarian constructing block forCNN.Complication layers encom pass a fixed learnable pollutants (a patch insidethe below snap).Every clear eschewal has small range and zenith and an identical depth as that of entering volume (3if the enter sub caste is snap enter).

For case, if we need to run complication ona picture with dimension 34x34x3.The possible length of filters may be axax3, wherein ‘a’ may be three, 5, 7, and many others but small compared to picture length. During in advance skip, we slide each clear out throughout the whole enter amount step by step in which every step is known as stride (that could have price 2 or three or perhaps four for immoderate dimensional snap shots) and

Cipher the fleck product betweenthe weights of pollutants and patch from enter quantum. As we slide our pollutants we ’ll get a 2-D affair for each sludge and we ’ll mound them inclusively and as an check result, we ’ll get affair volume having an depth same to the range of pollutants.

The network will have a look at all of the pollutants. The styles of layers in this network are as follows

Input layer: This sub caste holds the raw enter of the picture with a range of 32, a top of 32, and a depth of 3

Convolution Layer: This sub caste computes the affair volume via the computing fleck product amongstall pollutants and image patches. Suppose we use a aggr egate of 12 pollutants for this deposit we ’ll get affair volume of dimension 32 x 32 x 12.

Activation Function layer: This sub caste will follow an element smart activation point to the affair of the complication sub caste. Some commonplace activation bents are RELU maximum (0, x), Sigmoid 1/ (1e-x), Tanh, Leaky RELU, and so forth. The extent stays unchanged a result affair volume could have dimension 32 x 32 x12. Pool Layer: This sub caste is periodically fitted in the feedback and its foremost specific is to reduce the size of volume which makes the calculation rapid fire reduces reminiscence and also prevents over fitting. Two not unusual styles of pooling layers are uttermost pooling and common pooling. However, the consequent volume might be of dimension 16x16x12, If we use a maximum pool with 2 x 2 pollutants and stride 2.

V RESULT AND DISCUSSION

This includes all step by step process to execute emotion based music player which are given by:
7.1.1 Dataset

The dataset that is being used in this project is Fer2013. Fer2013 is an open-source dataset which is first, created for an ongoing project by Pierre-Luc Carrier and Aaron Courville, then shared publicly for a Kaggle competition, shortly before ICML 2013. This dataset consists of 35,887 grayscale, 48x48 sized face images with various emotions, all labelled. Emotion labels in the dataset:

0 : 4593 images – Angry
1 : 547 images – Disgust
2 : 5121 images – Fear
3 : 8989 images – Happy
4 : 6077 images – Sad
5 : 4002 images – Surprise
6 : 6198 images – Neutral

Fig 7.1: FER2013 dataset

VI SNAPSHOTS

As part of performance evaluation, multiple tests were performed taking into consideration different people expressing emotions in different ways. This way we were able to understand to what level of accuracy the system was able to identify emotions. To our understanding, the probabilities of the system detecting each emotion was considerably good and accurate. Testing and implementation is performed using OpenCV on Windows8/10, 32/64 bit operating system and Intel i3 core processor. The IDE that is used here is PyCharm. Here we have some screenshots depicting the tests carried out various emotions.

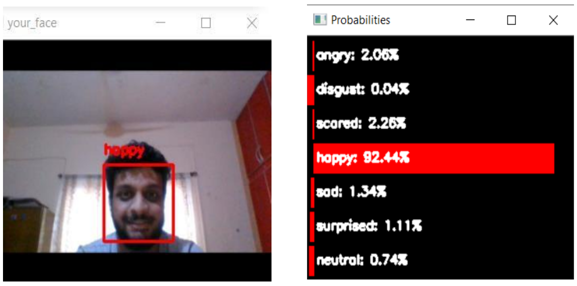

Fig 3. Happy emotion identified and music played

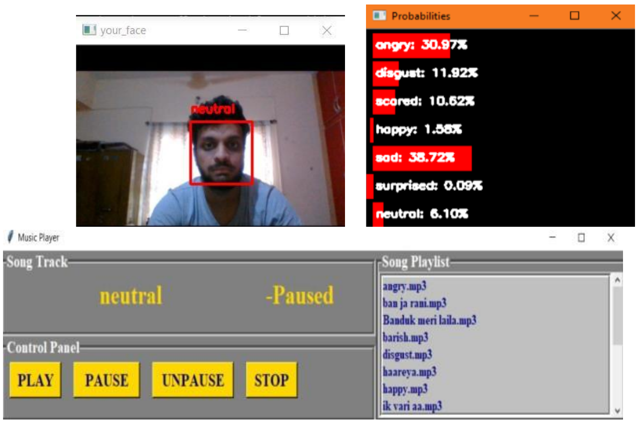

Fig 4 Neutral emotion identified and music played

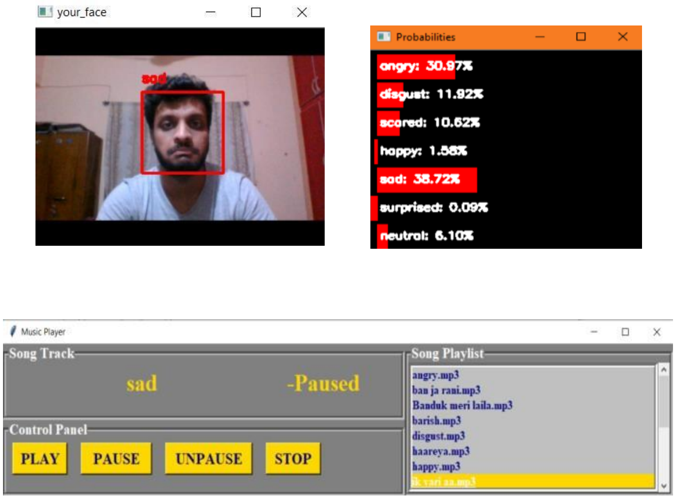

Fig 5 Sad emotion identified and music played

VII CONCLUSION AND FUTURE WORK

To summarize, we were able to create a simple CNN that reduce the Covid-19 patients stress burst based on song selections as per their facial expressions emotion. We extracted specified labels, converted the images to arrays, preprocessed our data, and set aside train data for our model to learn from. Convolutional Neural Networks (CNNs) have made gender prediction easier by removing the need to manually extract and detect which song need to play automatically against spotify me app. On a test set of data, our model gives a 66% accuracy.

The mediocre accuracy is due to the centralization of the facial detections (used for training) to a particular region. The accuracy can be improved by using emotions detections from various regions as train dataset. Studies have shown that a Covid patient's stress free is highly dependent on the counselling and music therapies

Various CNN architectures and approaches have been created in recent years to increase picture classification performance, and there will almost certainly be better ones in the future that will improve the performance of the results obtained in the tests even more. Apart from visual characteristics, the findings imply that emotions and auto play of song plays a role in gender classification.

There are still more ways available now that could potentially increase classification performance. The majority of today's image classification competitions are won by assembling multiple large models and doing model averaging. There are also preprocessing techniques that can be used, such as PCA (Principal Components Analysis) whitening. Song selections with various languages from worldwide can be integrated and extended for future process. For this type of system, information emulsion can do afour separate situations detector posi tion, match score position, point position, and decision position emulsion.

Also, further image data can be acquired for training the deep models, since there are proven results that the increased data ameliorate the bracket performance and widens the features available for the system to learn.The software can be progressed through editing and adding few functionalities. Including other feelings playing language-precise songs. Optimizing the set of rules through which includes extra capabilities which assist the machine to categorize customers based on many different elements like region and suggesting the consumer journey to that area and play songs for this reason. To add a more user-friendly interface. To port into other platforms like Android, IOS, etc

REFERENCES

[1] Zhao, X., Zhang, S., 2016. A review on facial expression recognition : feature extraction and classification. IETE Tech. Rev. 33, 505–517.

[2] Poursaberi, A., Noubari, H.A., Gavrilova, M., Yanushkevich, S.N., 2012. Gauss Laguerre wavelet textural feature fusion with geometrical information for facial expression identification. EURASIP J. Image Video Process., 1–13

[3] Taylor, P., Siddiqi, M.H., Ali, R., Sattar, A., Khan, A.M., Siddiqi, M.H., Ali, R., Sattar, A., Khan, A.M., Lee, S., 2014. Depth camera-based facial expression recognition system using multilayer scheme. IETE Tech. Rev. 31, 277–286. Journal of Computer Applications, 2012.

[4] wusu, E., Zhan, Y., Mao, Q.R., 2014. A neural-ada boost based facial expression recognition system. Expert Syst. Appl. 41, 3383–3390.

[5] Biswas, S., 2015. An Efficient Expression Recognition Method using Contourletm Transform. Int. Conf. Percept. Mach. Intell. pp. 167–174.

[6] Cossetin, M.J., Nievola, J.C., Koerich, A.L., 2016. Facial expression recognition using a pairwise feature selection and classification approach. IEEE Int. Jt. Conf. Neural Networks, pp. 5149–5155.

[7] Happy, S.L., Member, S., Routray, A., 2015. Automatic facial expression recognition using features of salient facial patches. IEEE Trans. Affect. Comput. 6, 1–12..

[8] Hernandez-matamoros, A., Bonarini, A., Escamilla-hernandez, E., Nakano-miyatake, M., 2015., A Facial Expression Recognition with Automatic Segmentation of Face Regions. Int. Conf. Intell. Softw. Methodol. Tools, Tech. 529–540

[9] Dahmane, M., Meunier, J., 2014. Prototype-based modeling for facial expression analysis. IEEE Trans. Multimed. 16, 1574–1584.

[10] Zhang, L., Member, S., Tjondronegoro, D., 2011. Facial expression recognition using facial movement features. IEEE Trans. Affect. Comput. 2, 219–229.

[11] Nigam, S., Singh, R., Misra, A.K., 2018. Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. Multimed. Tools Appl., 1–23

[12] Hegde, G.P., Seetha, M., Hegde, N., 2016. Kernel locality preserving symmetrical weighted fisher discriminant analysis based subspace approach for expression recognition. Eng. Sci. Technol. Int. J. 19, 1321–1333.

[13] Salmam, F.Z., Madani, A., Kissi, M., 2016. Facial Expression Recognition using Decision Trees. IEEE 13th Int. Conf. Comput. Graph. Imaging Vis., 125–130.

[14] G. Shakhnarovich, P. Viola, and B. Moghaddam, "A Unified Learning framework for real time face detection and classification", Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp.16-19, 2002.

[15] X. Leng, and Y. Wang," Improving generalization for gender classification", International Conference on Image Processing, pp. 1656-1659,2008.