JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Face Detection and Emotion DetectionUsing Deeplearning

Vaibhav Gosavi[1], Snehal Jadhav[2], Shivam Singh[3], Gayatri Londhe[4], Monica Gunjal[51,2,3]Student, Dr. D.Y. Patil Institute of Technology
[4,5]Assistant Professor, Dr.D.Y. Patil Institute of Technology, Pimpri, Pune
[1]gosavivh123@gmail.com,
[2]jsnehal2000@gmail.com,[3]shivam9102snp@gmail.com,
[4]gayatri.londhe19@gmail.com, [5]monica.gunjal@dypvp.edu.in

***Abstract* -** The facial emotion recognition system assumes a lot of significance in this period since it can capture mortal behavior, heartstrings, intentions, etc. We propose a fast face detector using an effective architecture predicated on a hierarchical cascade of neural network ensembles with which we achieve enhanced discovery delicacy and effectiveness.

We propose a way to form a neural network ensemble by using several neural network classifiers, each of which is specialized in a subregion in the face-pattern space. These classifiers round each other and, together, perform the discovery task. In this way, simpler and further effective ensembles used at earlier stages in the cascade are suitable to reject a maturity of nonface patterns in the image background

Index Terms – CNN (Convolutional Neural Network), Haar classifier, Haar Cascade frontal face, Deep learning, HOG (Histogram of Oriented Gradients), R-CNN (Region-based Convolutional Neural Networks), etc.

I. INTRODUCTION

In recent years, with the rapid development of artificial intelligence, more and more research has been conducted in the field of human-computer interaction technology. Deep learning, a subset of machine learning uses an artificial neural network, which is an algorithm inspired by the human brain. Convolutional Neural Network (CNN) is a class of deep neural networks that uses convolution as the mathematical operation [1].

As we all know everything is being interconnected so security is also need to be improved precisely. Many people have been working on face detection to improve the security in the application area. We are doing this project to contribute to the changes happening in security. CNN has achieved great results in face detection. With advanced network architecture and keen learning methods, CNN has improved face recognition ability to unprecedented levels.

As we know artificial intelligence is very crucial nowadays, and a very interesting topic so we decided to do the project on this stream. Our topic is mood analysis using deep learning, by using Haar Cascade andCNN model we detect the face of a person, and its scope is also vast.

The system can be used in an educational institutes, universities, etc., it can also be used in government offices and IT companies. Most categories have minimum images which typically isn't enough for a neural network to learn with high accuracy. Therefore, we are using a pre-built and pre-trained model applying transfer learning. Transfer learning is an optimization that permits rapid progress or improved performance when modeling the second task. The system captures video through the camera and these are then converted to image frames for further processing of recognition.

The system is initially trained with the training images that extract the facial features from the image. Different features enable recognizing an individual and then classifying the emotions of that person successfully. The images framed out ofvideos are cropped, focusing on the face and removing the background to have good accuracy in recognizing the emotions. We are looking forward to working on emotion detection.

II. EXISTING WORK

CNN has achieved great results in face and emotion recognition. With advanced network architecture and keen learning methods, CNN has improved face recognition ability to unprecedented levels.

The facial emotion recognition system assumes a lot of importance in this era since it can capture human behavior, feelings, intentions, etc. The conventional methods have limited speed and have less accuracy while the facial emotion recognition system using deep learning has proved to be the better one. This system aims to build a deep convolutional neural network model that recognizes different human facial emotions such as happiness, surprise, sadness, and neutrality and this can be used for applications such as customer feedback analysis, face unlocking, etc.

ShimaAlizadeh, et.al. [1], have presented convolutional neural networks (CNN) for a facial expression recognition task. The goal is to classify each facial image into one of the seven facial emotioncategories. In addition to the networks performing supported raw pixel data, and employed a hybrid feature strategy by which they have trained a unique CNN model with the combination of Histogram of Oriented Gradients (HOG) features and raw pixel data. To reduce the overfitting of the models, they have utilized different techniques including dropout and batch normalization in addition to L2 regularization.

AyaHassouneh, et.al. [2], have presented an emotion recognition system using facial expression and EEG. This paper aims to classify a physically disabled (deaf, bedridden, and dumb)person's emotional expressions based on facial landmarks and electroencephalograph (EEG) signals to employ a convolutional neural network (CNN) and long short term memory (LSTM) classifier. The result shows that the system can recognize emotion in 99.81% of facial landmarks and 87.25%of EEG signals. Suraj Kamal, et.al. [have presented the two-layer convolution network model for facial emotion recognition. The model classifies 5 different facial emotions from the image dataset. Adam optimizer is used to reduce the loss function and it's tested to have an accuracy of 78.04%. The model has comparable training accuracy and validation accuracy which convey that the model has the best fit and is generalizedto the info.

III. METHODOLOGY

OpenCV provides the Cascade Classifier class that can be used to create a cascade classifier for face detection. The constructor can take a filename as an argument that specifies the XML file for a pre- trained model.

OpenCV provides several pre-trained models as part of the installation. These are available onour system and are also available on the OpenCV Github project.

Haar Cascade is an Object Detection Algorithm used to identify faces in an image or a real-time video.

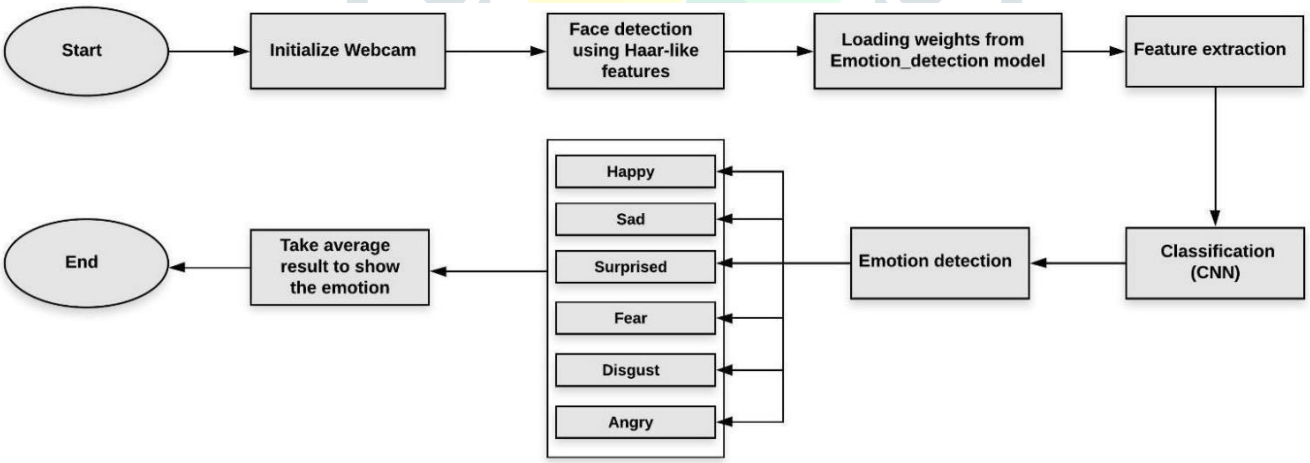

Fig.1 The system flow for the CNN model

Video Capture

Using System's local webcam, video is captured and with the help of Haar Cascades using OpenCV face of a person is detected, in each frame, important features like wrinkles, folds, and other edges which are important for facial expressions are found.

Frames Pre-processing

The Face detected in each frame is then converted from RGB to GGrayscalewhich would reduce the computational power also we need only wrinkles and specific edges from a person's face which can be found easily in GrayGrayscalege.

Frames are then set to detect images of different sizes and accordingly draw a rectangle around the face of a person with the message of emotion beside that rectangle.

CNN Model

For detecting or analyzing the mood of a person, a specific CNN model is created and is trained on the FER Dataset which contains approximately 35000 images of different emotions of different persons.

Also, data augmentation is performed on the dataset to bring a wide variance for an image in the dataset which would be a helpful trail.

By following the CNN architecture, hidden and fully connected layers are created explicitly and, in these layers, neurons are trained by forwarding and backward propagations which would fire with the help of activation functions when specific emotions are detected.

A Fully-Connected layer learns the non-linear combinations of high-level features as represented by the output of the convolutional layer. The image is then flattened to a column vector and is fed to a feed-forward neural network and backpropagation is applied to every iteration of training. After a series of epochs, the model can distinguish between dominating and low-level features, and then they have to be classified using the Softmax Classification technique. ResNet, VGGNet, and AlexNet are certain architectures of CNNs available which have been used in building certain powerful algorithms.

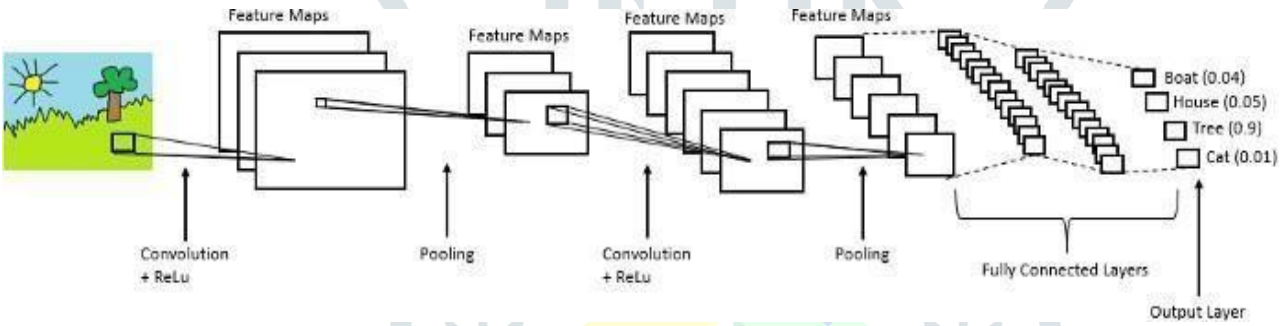

Fig.2 Working of CNN layers

Feature Extraction

With the help of several different Convolutional filters, different features from facial expressions in the images are extracted or convolved and considering the error, weights of the filter are updated with backpropagation and hence checks for the validation loss accordingly and continue to update until validation loss reaches to global minima which can be determined with the help of gradient descent technique.

Hence after sufficient epochs of training, now the filters in the CNN model are capable enough to detectthe features and classify the type of emotions accordingly.

IV. EXPERIMENTATION

First, all the libraries required have to be imported and then the image file is read. The training features and labels are fetched from the pixels of the image and are converted to NumPy arrays. An additional dimension is added to the feature vector by using the np.expand_dims() function. This is done to make the input suitable for the CNN model. The features and labels are stored in .npy files which will be used later.

For developing the CNN model, the required libraries are imported and then all the variables which will be needed for training the CNN are declared. The image resolution is 48*48 pixels. Five emotions (0=Angry, 1=Happy, 2=Neutral, 3=Sad, 4=Surprise) will be predicted by the model so there are 5 labels for those. The inputs will be processed with a batch size of 32 i.e 32 images will be passed through the CNN layers at a time.

The features and labels are loaded and standardized by subtracting the means and dividing by the standard deviation. Then the data is divided into training and testing sets, and the training data is further divided to obtain the validation data. The features would then be passed through the CNN to train the model and then be tested using the test features.

Systems Requirements:

- PC/Laptop
- Webcam
- 32 /64-bit (preferable) processor
- Internet connectivity via Wi-fi /Ethernet

Platform/Software requirement:

- Google Colab
- Jupyter notebook
- Image dataset (FER dataset)
- Python with its neural network libraries
- Keras

V. RESULTS AND DISCUSSION

We have completed the experimentation of this project and we have collected output for both the phasesof our project (Face recognition and Emotion Detection). Below are the photos.

The output of Face detection (Part 1):

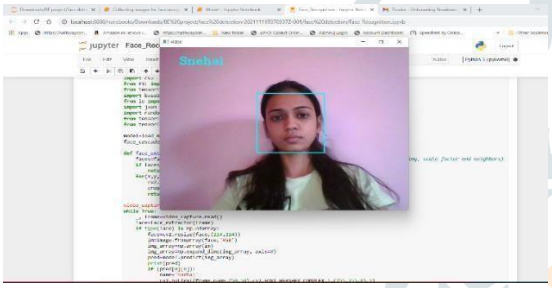

Figure 3.1

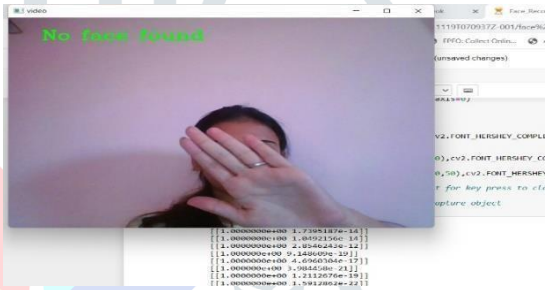

Figure 3.2

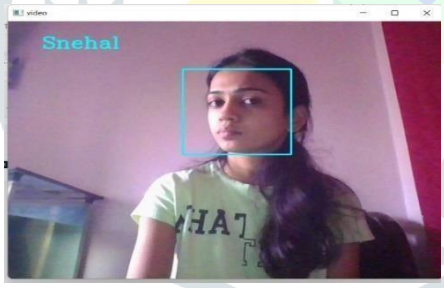

Figure 3.3

In part one, the algorithm gives the command to the system to initialize the webcam and take pictures asper the count. Then, it will recognize the person by comparing the face with the database.

The output of Emotion detection (Part 2):

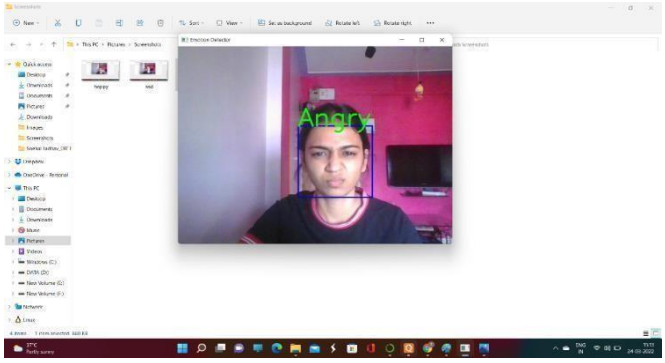

Figure 4.1

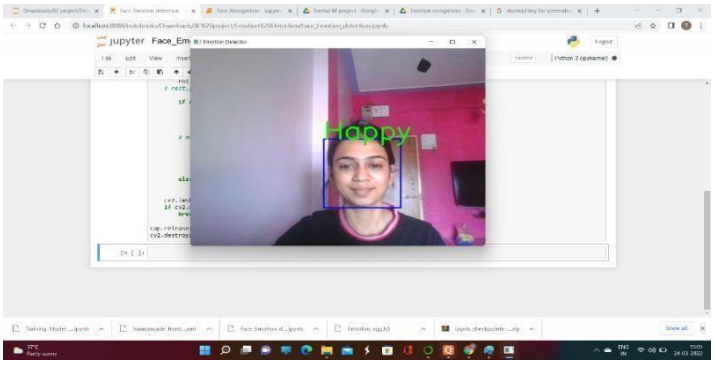

Figure 4.2

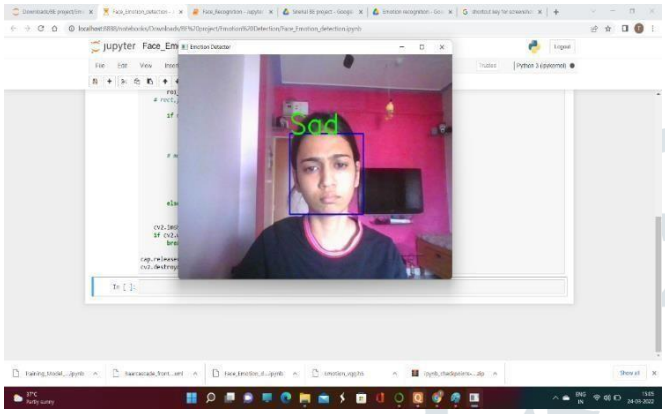

Figure 4.3

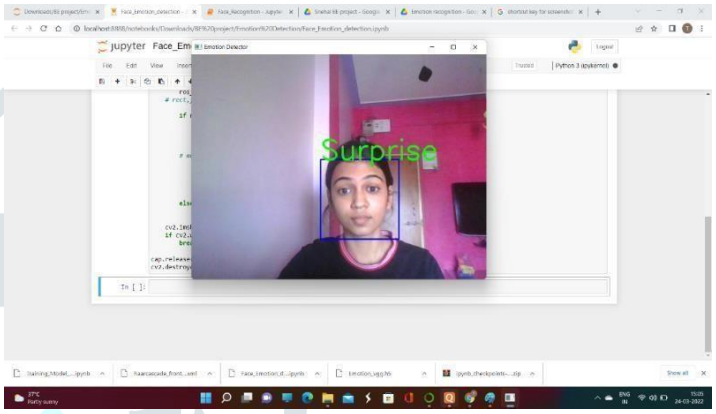

Figure 4.4

In part 2, the human faces are recognized and the system will check the emotion of each individual usingthe algorithm. We have trained basic human emotions viz. Happy, Sad, Surprise, and Angry.

Conclusion

In this paper the accuracy of the CNN model is around 74% and the number of filters and other parameters in the different layers of the CNN model has been changed to improve the accuracy of prediction. The Transfer learning model has also been developed to improve the accuracy and faces are identified with improved accuracy in this model. As the future scope is concerned, we can add face verification to this model which will help to verify the person.

REFERENCES

[1] E. Pranav, S. Kamal, C. Satheesh Chandran, and M. H. Supriya, "Facial Emotion Recognition Using Deep Convolutional Neural Network," *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2020, pp. 317-320, DOI: 10.1109/ICACCS48705.2020.90743021

[2] Z. Rzayeva and E. Alasgarov, "Facial Emotion Recognition using Convolutional Neural Networks," *2019 IEEE 13th International Conference on Application of Information and Communication Technologies (AICT)*, Baku, Azerbaijan, 2019, pp. 1-5, DOI: 10.1109/AICT47866.2019.8981757

[3] I. Lasri, A. R. Solh, and M. E. Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network," *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, 2019, pp. 1-6, DOI: 10.1109/ICDS47004.2019.8942386.

[4] D. Wang, H. Yu, D. Wang, and G. Li, "Face Recognition System Based on CNN," 2020 International Conference on Computer Information and Big Data Applications (CIBDA), Guiyang, China, 2020, pp. 470-473, DOI: 10.1109/CIBDA50819.2020.00111

[5] H. Zhang, A. Jolfaei and M. Alazab, "A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing," in IEEE Access, vol. 7, pp. 159081-159089, 2019, DOI: 10.1109/ACCESS.2019.2949741

[6] G. Aggarwal, A. K. R. Chowdhury, and R. Chellappa, "A system identification approach for video-based face recognition," Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., Cambridge, 2004, pp. 175-178 Vol.4, DOI: 10.1109/ICPR.2004.1333732.

[7] M. K. Lee, D. H. Kim, D. Y. Choi and B. C. Song, "Deep Transfer Learning for Emotion Recognition Networks,"

2018 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia), Jeju, 2018, pp. 206-212, DOI: 10.1109/ICCE-ASIA.2018.8552131.

[8] P. Khorrami, T. Le Paine, K. Brady, C. Dagli and T. S. Huang, "How deep neural networks can improve emotion recognition on video data," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 619-623, DOI: 10.1109/ICIP.2016.7532431.

[9] Y. Zhou, H. Ni, F. Ren, and X. Kang, "Face and Gender Recognition System Based on Convolutional Neural networks," *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, Tianjin, China, 2019, pp. 1091-1095, DOI: 10.1109/ICMA.2019.8816192.Changxing Ding, Dacheng Tao. Robust Face Recognition via Multimodal Deep Face Representation, IEEE Transaction on Multimedia(2015).

[10] I. Berle, Face Recognition Technology, Law, Governance and Technology Series 41, https://doi.org/10.1007/978-3-030-36887-6_5

[11] Ali, W., Tian, W., Din, S.U. et al. Classical and modern face recognition approaches a complete review. Multimed Tools Appl (2020). https://doi.org/10.1007/s11042-020-09850-1

[12] B. Krishnaveni and S. Sridhar, "Face Recognition using a greedy based warping algorithm," 2019 11th International Conference on Advanced Computing (ICAC), Chennai, India, 2019, pp. 253- 258, DOI: 10.1109/ICoAC48765.2019.247133.

[13] Anwarul S., Dahiya S. (2020) A Comprehensive Review on Face Recognition Methods and Factors Affecting Facial Recognition Accuracy. In: Singh P., Kar A., Singh Y., Kolekar M., Tanwar S. (eds) Proceedings of ICRIC 2019. Lecture Notes in Electrical Engineering, vol 597. Springer, Cham. https://doi.org/10.1007/978-3-030-29407-6_36