

# **Interview Emulator through Video Audio and Text through Deep Learning.**

Jayanta Ghosh

*Department of Computer Science Ramnarian Ruia Autonomous College  
Of Arts & Science Affiliated to University Of Mumbai  
Dombivli, India jayantbghosh98@gmail.com*

## **Abstract**

The Prototype Will be implemented in Python Technology Stack. The Prototype Will have the ability to rate a candidate of the interview According to the past performances of other interviewers and display their emotions along with their Histograms as it will be beneficial for the comparison. In case of the video interview, the candidate's emotions will be mapped and it will display emotions accordingly. All of the above mentioned implementations would be done with the help of Neural Networks and the concept of deep learning which has become a buzzword in the last two decades. There are several use cases which are prevailing of this applications in the near future which are listed as below can Be Used to Train the Interviewer And Better Prepare Them For Job Interview can Be Used To Recruit Candidates Without The Need Of Making The Interviewer Travel to the Workplace For The Interview can Be Used In Unprecedented Times (such as COVID-19 Pandemic etc.) as a cost cutting measure for recruiting. Interview Emulator is a relatively new discipline that aims to include text inputs, as well as sound and video.

## **Keywords**

**Emotion Recognition, CNN, Interview Emulator, ConvNet, LVQ-NN.**

## **Introduction**

Emotion Recognition is been part of many fields now, by detecting emotion of human through by using facial appearance, Program body shape, and etc. Facial image is used for periodically source to identify emotions. The algorithm is composed of three stages are first image processing stage, facial feature extraction stage, and emotion detection stage. The expected feature extraction method consists of three regions: eye region ancillary region and mouth region. Mainly Fuzzy classifier is used in for detecting emotion stages and performing Linear Matrix Inequality Optimization(LMI) Method can be very effective way to overcome this difficulty it finds the vague patterns in input emotion recognition system. For Specifically detecting whole face we have to extract the face region and that can be done by Virtua1 Face Model(VFM) based Histogram method to get information of face region and its components [1].

Emotion Recognition is an vital area of research to enable effective human-computer interaction(HCI). Human-computer interaction is that the study, planning, and style of how people and computers work together in order that a person's needs are satisfied within the best way. Human emotions are often detected using speech signal, facial expressions, visual communication, and electroencephalography (EEG). Humans are capable of manufacturing thousands of facial actions during communication that change in complexity, intensity, and meaning. The addition or absence of 1 or more facial actions may alter its interpretation. Personality Psychology is the major aspect for text emotion analysis its to uniquely identified persons text and their behaviour. The mechanisms should ensure one to one interactions with the social, anyone mind or their personality. Five major models that re used in Personality traits are Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. It is used to depict individuals and each others searches using natural language and theoretically based personality models. Personality traits its an alternative approach to individuals linguistic markers. It showed that multiple correlations between linguistic markers and the Big Five traits like Neuroticism is positively correlated allow the use of negative emotion words and negatively correlated allow the use of Positive emotion words[3]. Emotional Recognition aims at undoubtedly identifies the real state of human voice. The Suggested structure aims at identification of basic emotional states such as anger, joy, happy, neutral and sadness [5]. For to extract any information from any Pre text, LWIC is applied. LWIC indicates as Grammatical analysis and Word results. It is text analysis tool that implement Profitable and Comprehensive approach for studying the various affecting, mental and anatomical present in original samples. It basically targets word at a time by searching through lexicon file. [4]

## Model

CNN Architectures such as Inception, reduced the amount of criterion by including a Global Average Pooling Operation. Global Average Operation performs to extract comprehensive Average Pooling Operation. Global Average Pooling Operation reduces to extract scalar value by taking advantage average data over maps. CNN Models can be validated by creating a real-time vision system and contains approximately such as Xception Architecture for advantage from combination of two of fortunate results comes the use of residual modules and depth-wise separable convolutions. Depth-wise Separable Convolutions curtail promote an amount of framework by separating the mechanism of feature extraction and combination with convolutional layer. The dataset contains of ["angry", "disgust", "fear", "happy", "sad", "surprise", "neutral"] the have drifted out with 66% accuracy model. Basic Implementation of real time guided back- propagation to visualize the pixels [2].

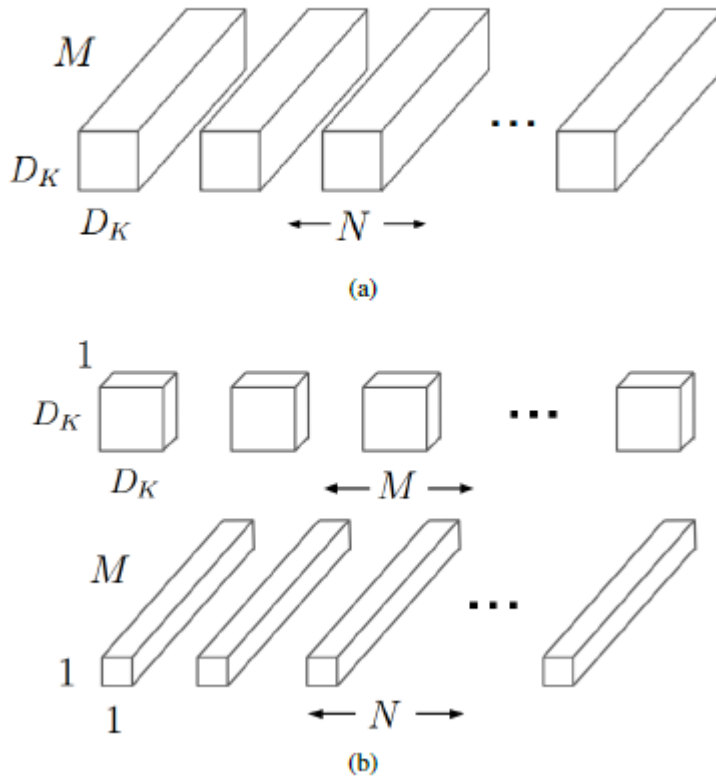


Fig 1[2] Difference between (a) standard convolutions and (b) depth-wise separable convolutions.

A Convolutional Neural Network (ConvNet/CNN) may be a Deep Learning algorithm which may absorb an input image, assign importance (learnable weights and biases) to varied aspects/objects within the image and be ready to differentiate one from the opposite.

For to extract any information from any Pre text, LWIC is applied. LWIC symbolizes as Linguistic Inquiry and Word Count. It is text analysis tool that implement Profitable and Comprehensive approach for studying the various affecting, mental and anatomical present in original samples. It basically targets word at a time by searching through lexicon file. [4]

For Emotion Based Emotion recognition through CNN can be used for classification as document – level inserts for topic Identification [4]. It has various convolution layer at different kernel size that has various through given assertion to effort the information of textual through different images. This paper has carried out e-Vector. e-Vector is a D dimensional vector it corresponds to number of classes. LSTM models has been an asset for neural networks and speech emotion recognition. LSTM have long dependencies they have been vanishing gradient problem. Bidirectional LSTM combines strategies and about average pooling layer [4]. It portrays and analyse the performances of Learning Vector Quantization Neural Network (LVQ NN) and Multiclass Support Vector Machine (SVM). The Overall Preliminary results admits that combination of (LVQ NN-SVM) has greater efficiency than LVQ-NN and SVM.

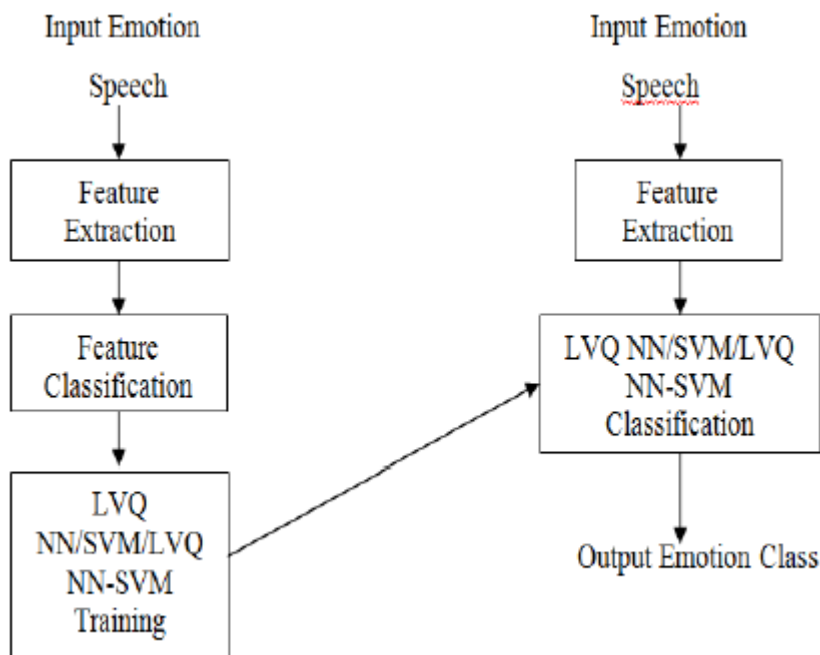


Fig 2[5] Structure of Speech Emotion Recognition

In Block diagram consists of Emotional Speech as input, feature Extraction, feature Selection, Classifier and detection of Emotion. An important step in Emotion Recognition is to through a select o significant feature which carries over Speech signal about emotional information. The Commonly used features such as Prosodic and spectral are obtained. LVQ is a neural net that blends competitive learning with care. The SVM is worn for computer algorithm classification used in pattern recognition for data classification and regression. It performs classification by N-dimensional hyper-plane that optical separates the data into categories. Hybrid Method of LVQ-NN and SVM is comparatively analysis done of single NN, Single SVM and hybrid model is done where in LVQ-NN it classify input vectors automatically whereas SVM constructs a N-dimensional hyper-plane that optimally separates data [5]. Time Distributed CNN(TD-CNN) model will be using for voice recognition through by CNN and LSTM models and apply undulating window all along with log-mel – Spectrogram and the output of each of these CNN networks into recurrent neural network by 2 cells of LSTM to learn the long-term dependent assurance. Certainly, a fully connected layer with softmax activation is used to conclude the emotion identify in the voice.

## Conclusion

Finally concluding the idea extract, by As, in the prevailing and ever growing world, where the technology is booming in each and every field this prototype would be used for recruiting the candidates from all over the world using a single tool. Also, here depicting the use cases of this future Prototype. The tool can be used as an alternative (as an assistive tool not a replacement of the manual interview process) of manual interview in the unprecedented times, so that, the company can be safeguard the health and safety of both their employees and the interviewee candidate. The tool can also be used to train the candidates in order to outperform their budding competitive opposition candidates. In this papers found that emotion recognition can be done by applying psychology, face recognition and behaviour of human though models and technologies. Also, in addition to the paper research works, it is also evident to note that, In these Pandemic Situations the whole world is facing many crisis and the most hard hit are the companies who have

been experiencing uncertainty due to COVID-19. The companies are left to a dilemma such as they cannot recruit candidates even though there is a prerequisite in the vacancies in the companies, because of the COVID-19 contraction fear and dilemma. More over the candidate could use the time productively by preparing for the interviews through this tool, which would be expected to analyse the emotions accordingly and prepare the interviewee for the same.

## References

- [1] Y. H. Joo, *Emotion Detection Algorithm Using Frontal Face Image*, p. 1-7, july 2015.
- [2] O. A. P. G. P. M. Valdenegro, "Real-time Convolutional Neural Networks for Emotion and Gender Classification".
- [3] J. C. R. P. P. K. J. V. Y. C. N. Dehak, *Deep neural networks for emotion recognition*, pp. 1-5, 2018.
- [4] J. C. U. B. A. L. P. C. K. C. R. d. D. B. Edward P. Tighe, *Personality Trait Classification of Essays with the Application of Feature*, pp. 1-7.
- [5] M. A. P. Wanare, "Int. Journal of Engineering Research and Applications," *Human Emotion Recognition From Speech*, vol. 4, no. 7, pp. 1-5, 2014.