AI Interview Agent for Predicting Communication Skills and Personality Traits

Prajwal R[1], Divyansh Divya[2], Hitesh Gowda K[3], Prof. Hemalatha K.L, Professor and HOD[4]
[1]Department of ISE, Sri Krishna Institute of Technology, B'lore-560090, India
[2,3,4] ISE Department, Sri Krishna Institute of Technology, B'lore-560090, India

Abstract — In organizational and industrial psychology along with affective computing, it's crucial to anticipate interpersonal communication abilities and personality traits. AVI-AI is an Asynchronous Video Interview platform with an AI decision agent based on Tensor Flow Convolutional Neural Network (CNN). The AI Interview Agent detects critical personality traits like agreeableness, openness, neuroticism, extraversion, and conscientiousness by assessing the candidate's answers. It helps employers make better hiring decisions by identifying candidates with the necessary communication abilities and personality traits for a particular role. By automating the interview process, the AI Interview Agent can save time and reduce hiring biases. It also offers consistent support and objective for evaluations each candidate's, Personal traits and communication abilities.

***Keywords—AI, AVI, CNN* and Facial Expression Recognition Dataset (2013).**

I. INTRODUCTION

Organisational success and work effectiveness depend on having strong interpersonal skills and personality qualities. effectiveness. Interpersonal communication involves the exchange, sharing, and feedback of information using verbal and nonverbal messages. Verbal messages convey exact words, while nonverbal messages such as gestures, facial expressions, posture, and tone of voice, help interpret underlying emotions, attitudes, and feelings. To provide an alternative to traditional interviews, asynchronous video interviews (AVI) have been developed. AVI allows job candidates to answer webcam-based pre-determined interview questions. evaluating one's personality and communication abilities traits in job candidates can be a challenging task for employers, given the subjective and time-consuming nature of traditional interviewing methods, which can lead to biased hiring decisions. To overcome these challenges, AI Interview Agent has emerged as a promising solution. This software application employs artificial intelligence and natural language processing to evaluate a candidate's communication skills and personality traits. By automating the interview process, AI Interview Agent offers an objective and consistent evaluation of each candidate's personality traits and communication skills unaffected by personal biases or subjective opinions of interviewers. The software predicts language fluency, verbal and non-verbal communication skills and also identifies candidate's personality traits emotions such as conscientiousness, agreeableness, openness, extraversion, and neuroticism for employers to make informed hiring decisions.

II. RELATED WORKS

Intelligent video interview agent used to predict communication skill and perceived personality traits

According to a research done on industrial and organisational psychology (I/O) in both organizational and industrial psychology A crucial duty is to forecast each person's interpersonal communication skills and personality characteristics. An Interview is scheduled for 20 people to this study, including 77 interviewers and 7 participants. To gather data on interviewees interpersonal qualities and communication process as observed by actual interviewers controlled interview environment. Then by creating AVI-AI, an

asynchronously video interviewing system with an artificially intelligent (AI) decision agent built on a tensor flow convolutional neural net (CNN), which can be utilised to predict a job the applicant's communication abilities and personality traits and replace some of the work done by human raters during the first phase of employment screening.

"A Deep Learning Based Approach to Measure Confidence for Virtual Interviews

In interviews, confidence is crucial in showing trust and faith in one's abilities and can be the deciding factor in getting hired. However, virtual interviews during the pandemic rely solely on a candidate's facial expressions to gauge their self-confidence. To address this, Deep Learning can be used to classify confident and unconfident images based on facial expressions. This can help interviewees practice confident expressions and improve their skills. The model is trained using two optimizers, Adam and SGD, to develop a confident and unconfident image classifier. Trust and a firm belief in something or someone are the definitions of confidence. Being self-assured during an interview, that is, having faith in your own abilities, can make the difference between you succeeding or failing. Candidates. It is just as crucial as demonstrating your capabilities since without it, you won't be able to persuade the interviewer to believe in your abilities to learn, complete the work, and use your expertise. Positivity about one's appearance and a strong, self-assured demeanor can leave a lasting impression on an interviewer and possibly other future coworkers as well. Since interviews are now conducted online in the age of the coronavirus, it is only possible to assess a candidate's level of confidence by looking at his or her face. Deep Learning has potential benefits.

Personality Recognition and Video Interview Analysis

As the employment agency sector currently needs AI to be implemented, we can anticipate human skill and its recommended job type in certain help of Artificial Intelligence (AI) systems. utilizing all machine learning techniques including Naive Bayes, Random Forest, SVM, and Convolutional Neural Network (CNN). This project will involve the conversion of an employee's voice into text, their emotional state, and a face-based company verification process. In the future, system users will be able to examine user records and analyze their effectiveness. Students can also profit from this approach by using it to analyze their own emotions and personalities, as we may one day see it used in interview processes. Future system users will have access to user records that have been evaluated.

III. METHODLOGY

1) Facial Expression Recognition Dataset

Facial Expression Recognition, a dataset from Kaggle, is used for facial expression recognition. Using this dataset, the emotions that the system detected are ‘Angry’, ’Fear’, ’Happy’, ’Sad’, ’Surprise’ and ‘Neutral’. The dataset Consist of 36,986 data which is divided into training dataset and testing dataset. The system uses 29,000 Training dataset is used to train the model and 7,896 Testing dataset is used. Whenever the system gets any image, it will detect the facial expression of that image by comparing it with the images of the training dataset. The system reads an image using the values at each pixel, and using CNN these pixels are processed further for generating results. Images are fed to the system. Then they are converted into greyscale images. From this greyscale image, the face is then detected.

2) Feature Extraction and Modeling

Created a three-stage model, as shown in Fig. 1, to create an AVI-AI programme that could be utilised to forecast interpersonal communication abilities including traits of personality as evaluated by the human raters. Using we created an AVI to extract the interviewees' facial expressions from each frame

during the video data processing step. By recording 86 landmark locations on the face each frame, the open-CV and Keras were used to identify the facial characteristics. Preprocessing was required to reduce

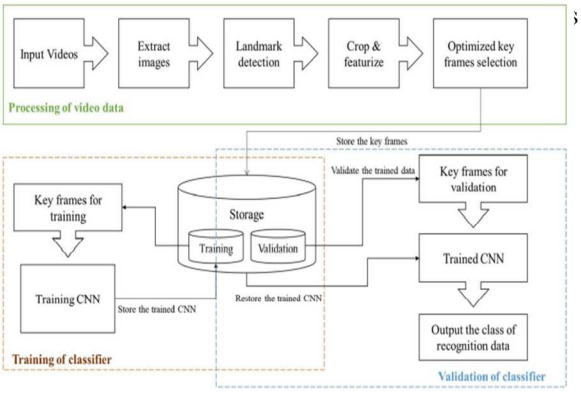

Fig.1: Traing Video Processing Model

3) Facial Emotion Detection Model

Facial Expression Recognition, a dataset from Kaggle is used for facial expression recognition. Using this dataset, the emotions that the system detected are 'Angry', 'Fear', 'Happy', 'Sad', 'Surprise' and 'Neutral'. Each face must be assigned to a single category of seven categories, with 0 denoting anger, 1 disgust, 2 fear, 3 happiness, 4 sadness, 5 surprise, and 6 neutralities. The dataset is divided into training and testing datasets. Whenever the system gets any image, it will detect the facial expression of that image by comparing it with the images of the training dataset. The system reads an image using the values at each pixel, and using CNN these pixels are processed further for generating results, images are given to the model. Then they are converted into greyscale images. Grayscale portraits of faces measuring 48 × 48 pixels make up the data. The faces were automatically registered such that each face roughly fills an identical quantity of space in each image and is cantered. Then we resized the images into the scale of 49*49*1. After resizing the detected face, scaling of an image is done. This image then undergoes multiple layers to detect the emotion of that image as shown in Fig 2 then face emotion is detected.

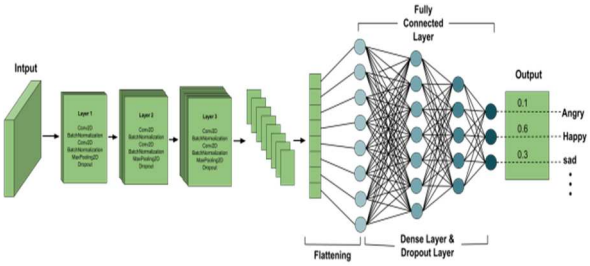

Fig.2: Facial Emotion Detection Model

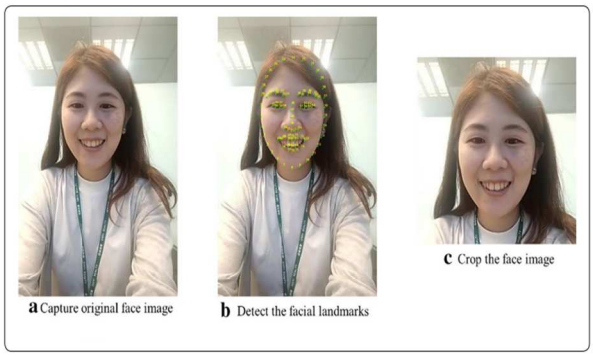

Fig. 3: Face Identifying and Cut Shot

To train the classifier for predicting communication skills and personality traits, we utilized a process of detecting and cropping face images, as shown in Fig. 2, where we obtained the original image, cropped the picture after facial landmark detection for training. Later, a gray-scale model was created from the cropped images to emphasise the features of movement and facial expression while lessening the impact of lighting.

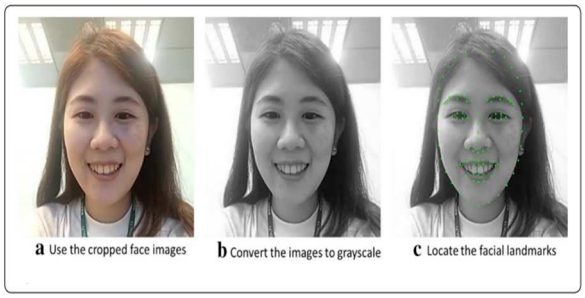

Fig.4: Finding the Facial Landmarks and Grayscale Conversion of the Compressed Face Images

The next step involved locating the 86 facial landmarks, displayed in highlighted in Fig. Any frames where these landmarks could not be detected were discarded. For the divider training phase, then combined Kaggle dataset with the extracted features.

4) CNN Model

As shown in the Fig, the model consisted of a Tensorflow Flow based CNN framework, in which the CNN (convolution neural network) structure is the artificial neural network which consist of in order to classify images, four layers of convolution, three layers for pooling, 10 mixture layers, a layer that is fully connected, and soft max layers are used.

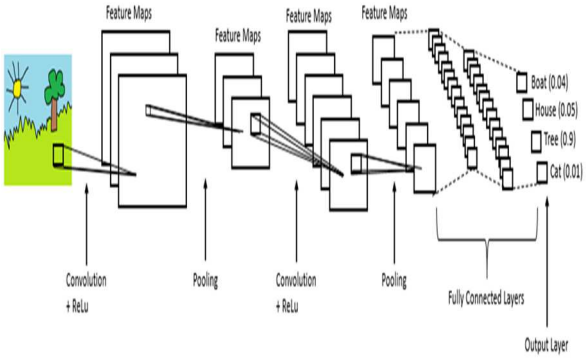

Fig.5 : The Architecture of the CNN Model

In CNN, neuron of single layer is connected to the small region. Hence using CNN, weights that needs to be handle or process will be less and it will require less numbers of neurons. CNN classifies the image by considering pieces, called as Features. For this, recorded videos are converted into number of Key Frames to get the multiple images.

Fig 6: Providing the Featured Images into the Model

The input images of 640 pixels wide and were normalized to account for variations in rotation and shifting, as well as to prevent distortion of the original facial image. The interviewee's extracted facial expression features were used as inputs, while the communication ability score and big personality traits were the output of the neural net. Each layers of the trained model contained training parameters, and the Rectified Non-linear Unit (ReLU) was used to address the vanishing gradient problem that could occur in a sigmoid function. A soft max level with was the last layer with 70 possible outputs. Randomly selected participants from both the training set eighty percent and the verification set (20%) during the classifier verification stage. Each interviewee had different features, including a single communication ability score plus five personality attributes. With a rate of learning of 0.02, a test rate of twenty, and a training a batch size of 256, the algorithm underwent training for 2,250 iterations. After then, the model was checked for correctness.

5) HAAR Cascade

HAAR Cascade is an object detection algorithm that utilizes machine learning to identify objects in images or videos. It employs a cascade function that is trained with both positive and negative images to detect objects in other images. One of the primary applications of HAAR Cascade is in facial and eye detection.

The algorithm uses a set of HAAR-like features, which are simple rectangular filters that are applied to an image or video frame to extract features. These features are then combined to form a strong classifier that can detect objects in the image. The cascade classifier detect multi-scale method is then used for detection, returning boundary rectangles for detected faces or eyes. The mathematical equation of HARR Cascade is given by

$$F(x) = af1(x) + a2f2(x) + a3f3(x) + \ldots. \quad (1)$$

6) Local Binary Patterns

The feature extraction method is called LBP. The initial LBP operator encodes the local framework around each pixel by pointing the pixels in a picture with decimal numbers, often known as LBPs or LBP

codes. By deducting the value of the centre pixel, each of the pixels is contrasted with its eight neighbours in a 3X3 neighbourhood. In the final code, negative values are represented by 0 and all other values by 1. A binary number is created for each given pixel by adding all of these binary values clockwise, starting from one among its top-left neighbours.

The LBPs, which or LBP codes are used to refer to the resulting binary integers. LBP is further expanded to include consistent patterns. Binary string is thought of as circular, an LBP is deemed to be uniform if it has no over two bit-wise changes from zero to one or vice versa. E.g.,00000000, 001110000 and 11100001 are uniform patterns. A histogram of a labelled image f1(x, y) can be defined as shown in Equation (2) and (3)

$$Hi = \sum x,y\ I\ (fl(x, y) = i),\ \ i = 0 – n\text{-}1 \qquad (2)$$

Where n is the number of different labels produced by the LBP operator

$$Hi = \sum x,yI\ (fl(x,y) = i)I((x,y)\epsilon Rj \qquad (3)$$

IV. RESULTS AND DISCUSSION

To assess the candidate's emotion features, we utilized the Facial Expression Recognition Dataset (2013) available on Kaggle, comprising over 37,000 images categorized into 256 classes. To enhance image representation and minimize interference from hair and features, we converted all the photos to grayscale.

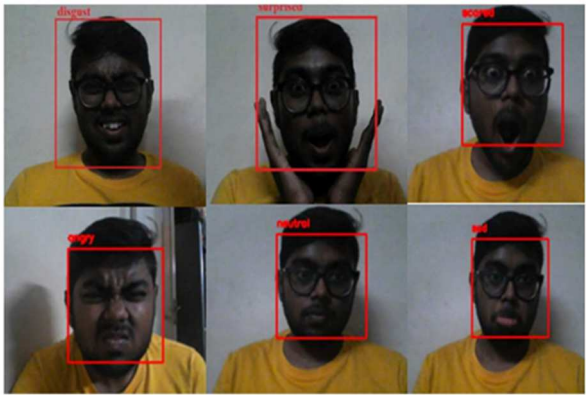

Fig 7: Emotion Features

The AVI-AI was able to learn and predict with success the facial emotions, communication abilities, and personality qualities including openness, agreeableness, and expected neuroticism as evaluated by human raters, but not conscientiousness and extraversion. then stating that a candidate's facial expressions represent his or her communication abilities as determined by HR specialists based on organised behavioural interviews.

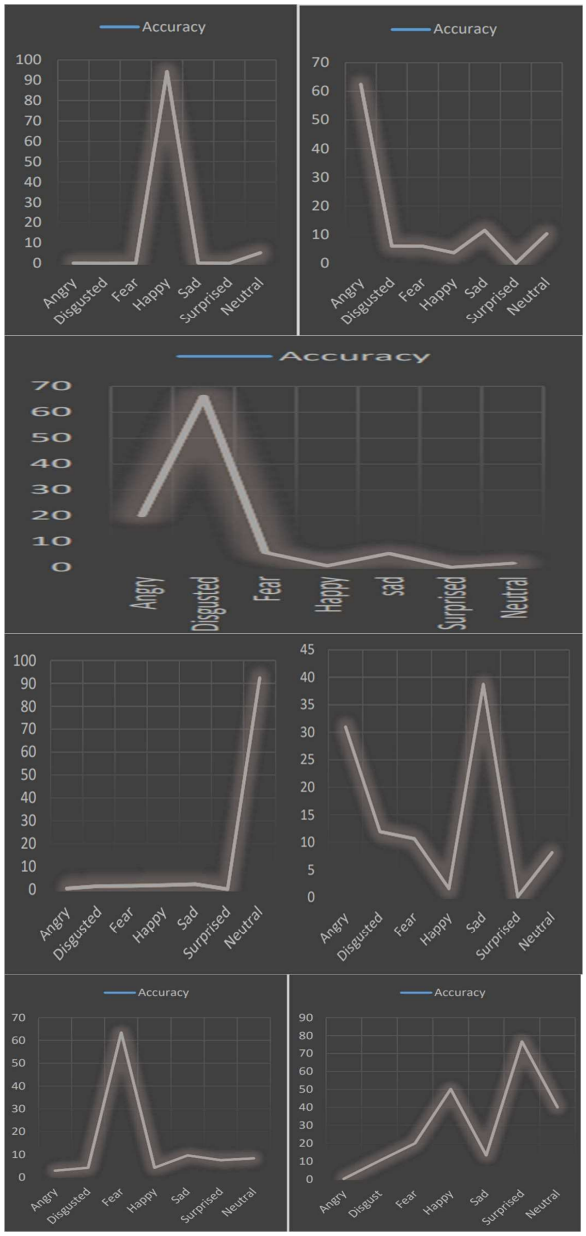

Fig 8:Accuracy Predictions for Happy, Angry, Disgusted, Surprised, Sad, Fear and Neutral

The Accuracy of the proposed model, VGG-16 Res-Net model, Emotion detection model, Image-Net Pre-trained model as shown the below table 1.

Table 1: Performance of Comparison model

Model	Performance			
	Precision	Recall	F1 Score	Accuracy
Proposed Model	86.09	82.31	79.23	89.32
VGG-16 Res-Net Model	90.33	60.23	74.21	74.89
Emotion Detection Model	88.11	77.2	55.12	76.25
Image Net Pre Trained Model	74.89	71.70	75.29	66.45

The AVI-AI's concurrent validity was measured using the Pearson correlation coefficient (R), explained variation (R^2), and mean square error (MSE), as shown in Table 2. R^2 represents the proportion of variance in the dependent variable that can be predicted by the independent variables, with higher values indicating better estimation. Conversely, a lower MSE (0 being perfect) indicates a smaller estimation error.

Table 2: Output Factors of AVI-AI

Output factors	R	R^2	MSE	ACC %
Communication skills	0.972	0.945	0.045	99.5
Openness	0.987	0.970	0.031	98.8
Conscientiousness	-	-	-	-
Extraversion	-	-	-	-
Agreeableness	0.978	0.957	0.038	96.9
Neuroticism	0.982	0.964	0.037	98.4

V. CONCLUSION

The usage of AI Interview Agent Platform for predicting communication abilites and personality traits has the potential to revolutionize the recruitment process. The ability of AI interview agents to analyze personalities in a standardized and objective manner can help remove human biases and streamline the recruitment process. However, it is important to address the ethical concerns associated with the use of AI interview agents. Companies must ensure that the algorithms used by AI interview agents are transparent and free from biases, and that applicants are fully informed about the use of AI interview agents and their data privacy rights. Despite these concerns, the benefits of using AI interview agents for predicting communication abilities and personalities traits are clear. They offer a powerful tool for companies to efficiently screen job applicants and identify the most suitable candidates for further consideration.

VI. ACKNOWLEDGEMENT

We would like to thank our guide and HOD, **Dr Hemalatha K.L Department of Information Science and Engineering, Sri Krishna Institute of Technology Bangalore**, for her valuable suggestion, expert advice and moral support in the process of preparing this paper.

REFERENCES

[1] Conrad D and Newberry R, "24 Business Communication Skills: Attitudes of Human Resource Managers Versus Business Educators",International Communication Journal 13:4–23 2, 2016.

[2] Prasetio, B. H., Tamura, H., and Tanno, K., "Support Vector Slant Binary Tree Architecture for Facial Stress Recognition Based on Gabor and HOG Feature", In 2018 International Workshop on Big Data and Information Security (IWBIS) (pp. 63-68). IEEE, 2018.

[3] Kazemi, V. and Sullivan, J., "One Millisecond Face Alignment with an Ensemble of Regression Trees" In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1867-1874), 2019.

[4] Zhang, K., Zhang, Z., Li, Z., and Qiao, Y., "Joint Face Detection And Alignment Using Multitask Cascaded Convolutional Networks", IEEE Signal Processing Letters, 23(10), 1499-1503, 2016.

[5] Spitzberg BH and Adams TW CSRS, "The Conversational Skills Rating Scale: Instructional

Assessment of Interpersonal Competence", National Association, Washington, DC, 2018.

[6] Mohammadi G and Vinciarelli A, "Towards a Technology of Nonverbal Communication: Vocal Behavior In Social and Afective Phenomena", In: Gokcay D, Yildirim G (eds) "Afective Computing and Interaction: Psychological", Cognitive and Neuroscientifc Perspectives. IGI, Pennsylvania, pp 133– 156, 2019.

[7] Levashina J, Hartwell C. ., Morgeson, F.P and Campion, "The Structured Employment Interview: Narrative and Quantitative Review of the Research Literature", Personnel Psychology, vol:1 - pp: '241-293', 2019.

[8] Salminen J., Miekka-Niemi S., and Vainio L., "What if the Candidate Were a Robot? Evaluating Human-Human and Human-Robot Job Interviews", vol:8 - pp:'39-48', 2020.

[9] Dery R, Tziner A, and Eden D., "The Dark Side of AI in HR Selection: Opportunities and Challenges", vol:IX - pp:'223-232', 2020.

[10] Zide J , Chua E, Golbeck J and Stone P, "Predicting Interview Outcomes with Natural Language Processing and Machine Learning", Proceedings of the International Conference on Human-Computer Interaction, vol:IV - pp:'384-399', 2018.

[11] Sriramprakash S, Prasanna V.D, and Murthy O. R, "Stress Detection in Working People", Procedia Computer Science, 115, 359-366, 2017.

[12] Supriya Anand , Nihar Gupta , Mayesh Mulay , Abhimanyu Sherawat and Rupali Gangarde, Personality Recognition and Video Interview Analysis, International Journal Of Engineering Research and Technology (IJERT) Volume 10, Issue 05 (May 2021).

[13] Iftekhar Naim, M. Iftekhar Tanveer, Daniel Gildea, and Mohammed (Ehsan) Hoque. "Automated Prediction and Analysis of Job Interview Performance: The Role of What You Say and How You Say It" ROC HCI, Department of Computer Science, University of Rochester, ROC HCI, Department of Electrical and Computer Engineering, University of Rochester,2015.

[14] Anna Lena Hunkenschroer and Christoph Luetge, "Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda," Journal of Business Ethics, Springer, vol. 178(4), pages 977-1007, July.

[15] Hung-Yue Suen , Kuo-En Hung and Chien- Liang Lin "Intelligent Video Interview Agent used to Predict Communication Skill and Perceived Personality Traits" Suen et al. Human. Central. Computer. Information. Science. 10:3 (2020).