Archives available at journals.mriindia.com

International Journal of Recent Advances in Engineering and Technology

ISSN: 2347 - 2812
Volume 14 Issue 01s, 2025

INTERVIEWX: AI/ML POWERED INTERVIEW SIMULATOR USING NLPAND CNN

[1]Bhor Akshada, [2]Kute Manjusha, [3]Mate prachi, [4]Prof. Mundhe Bhalchandra
[1 2 3 4]AIDS Engineering Jaihind College of Engineering, Kuran, Pune, India
Email: [1]akshadabhor210@gmail.com, [2]kutemanjusha4@gmail.com, [3]mateprachi7@gmail.com, [4]maauti.jcoe@gmail.com

Peer Review Information

Submission: 1 Sept 2025

Revision: 28 Sept 2025

Acceptance: 12 Oct 2025

Keywords

AI, ML, CNN ,NLP, Embedding system, Recognition.

Abstract

Interview-X is an advanced AI and ML-powered simulator designed to mimic real-world interview scenarios, enabling users to effectively prepare for interviews. The platform evaluates both verbal and non-verbal responses, offering comprehensive feedback that helps users refine their performance. By integrating Natural Language Processing (NLP) and Convolutional Neural Networks (CNN), Interview- X delivers precise insights into the user's communication skills and confidence levels. For speech evaluation, the system uses NLP to transcribe and analyze spoken answers. By leveraging the Google Speech-to-Text API, it converts audio inputs into text and evaluates the accuracy andrelevanceof theresponses based on the interview questions. A predefined accuracy threshold determines whether the answer meets the desired standard, offering actionable feedback for improvement. On the non-verbal front, Interview-X uses CNNs for facial expression analysis. With technologies like MTCNN for facial detection and Face Net/VGG Face for feature extraction, the system assesses the user's confidence by comparing facial landmarks and expressions with predefined confident face templates. This analysis ensures that both verbal content and non-verbal cues are aligned for effective interview performance. The platform aggregates results from both NLP and CNN evaluations, providing users with detailed feedback on their performance. This holistic approach helps individuals enhance both their answers and their presentation, boosting their interview confidence and readiness. Interview-X is a valuable tool for anyone looking to improve their interview skills in a realistic, data-driven environment.

INTRODUCTION

Interview-X is an innovative AI and ML-based platform designed to simulate realistic interview environments, helping users enhance their interview performance through comprehensive feedback. As the job market becomes more competitive, mastering both technical skills and interpersonal communication is essential for candidates aiming to succeeding interviews. Interview-X addresses this need by combining advanced Natural Language Processing (NLP) and Computer Vision (CV) techniques to evaluate users' verbal and non- verbal responses in real time. Interviews not only test a candidate's technical knowledge and problem-solving skills but also assess their communication, confidence, and ability to manage stress. Traditionally, interview

preparation has involved mock interviews with human interviewers or static question-answer sessions, which often lack dynamic feedback on key elements such as verbal communication effectiveness, body language, and confidence. As the demand for more personalized and efficient preparation methods grows, leveraging Artificial Intelligence (AI) and Machine Learning (ML) offers a promising solution. The platform's goal is to provide an immersive interview experience that goes beyond traditional preparation methods.

DATASET

There are special datasets on Indian Sign Language (ISL) and some of them are gathered from Kaggle which were used in previous papers which were used for reference purposes. Those images were already in the 128 * 128 resolutions. There is another dataset which was created by ourself for the training process. The images were 1280 * 720 resolution. These images were directly used to train the landmark model for better detection of the hand landmarks with better quality and then those images were converted to128 * 128 resolution for training the image model which reduces the time complexity for the model. The images were even processed with different augmentation and the different processing. The processing included adding more brightness and adding different kinds of rotations in the image. This could help in making the model more robust even if used in different situations. The imageswere captured with different backgrounds to make the model more reliable. There is image (3.1),which is a representation of what kind of dataset used to train the models.

OBJECTIVE

- To Enhanced user experience with AI-Powered feedback.
- To boost user Confidence with real time analysis.
- To ensure accessibility for all experience levels.
- To continuously update with new modules and features.

LITERATURESURVEY

1. Dr.SurendraMahajanet.al."Generative AI-BASED Interview Simulation and Performance Analysis",5 May 2024.

In today's job market, interviews are essential for evaluating a candidate's skills, qualifications, and fit with an organization. However, many job seekers struggle with anxiety and stress during interviews, often due to insufficient preparation or lack of familiarity with the process. To help with these challenges, we propose a Generative AI-based interview preparation app. This app offers a personalized way to practice interviews through voice and text inter actions, providing realistic simulations, tailored feedback, performance tracking, and easy accessibility. It aims to improve candidates' interview readiness. This paper outlines the application's flow, from creating a profile to receiving an interview scorecard, and describes the app's features, functionality, and technology. Keywords: Generative AI- based interview, job market, feedback and guidance, mock interview, candidate skills.

2. Pankaj Rambhau Patil et.al. "Elevating Performance through AI-Driven Mock Interviews", 2024

This paper introduces an advanced AI-powered mock interview platform aimedatim proving interview preparation. The platform assesses candidates in three main areas: emotions, confidence, and knowledge. Using deep learning techniques, it analyzes facial expressions to read emotional responses, while speech recognition and natural language processing gauge confidence levels. It also evaluates knowledgebycomparingcandidates'answerstore levanton line information through semantic analysis. This all- encompassing method is designed to lower anxiety, build confidence, and sharpen interview skills, of faring a more effective practice tool than traditional approaches. Keywords:AI-based interview,speech recognition,emotion recognition, FFM PEG

3. Prof. Sakharam Kolpe, et.al. AI Based Mock-Interview Behavioral Recognition Analyst", May 2024

In today's competitive job market, interviews kills are crucial for graduates. Many candidates lack adequate preparation for real interview situations. To address this, our approach uses an AI-driven mock interview system with a virtual recruiter to help candidates improve social and communication skills. Using real-time cues like facial expressions, head nodding, and speaking rate, it provides feedback to enhance performance. A speech-to-text tool evaluates grammar, and visual result comparisons track progress over sessions. This system aims to boost candidates' readiness for job interviews through realistic practice and feedback

METHODOLOGY

Algorithm

i) Speech Recognition Algorithm (Google Speech-to-Text)

Input: audio data (from user) Output: transcribed text

Step 1: Initialize Google Speech-to-Text API
Step2: Send audio data to the API
Step3: Receive the transcribed text from API response Step 4: Return transcribed text

ii) Text Evaluation Algorithm (Gemini Pro)
Input: transcribed text, interview question
Output: accuracy score
Step1: Initialize Gemini Pro NLPAPI
Step 2: Send transcribed text and interview question to the Gemini Pro API
Step 3: Receive accuracy score from the API response (0 to 100
Step4: If accuracy score¿=80, mark answer as correct
Step 5: Return accuracy score.

iii) Facial Detection Algorithm (MTCNN):
Input: video frame
Output: face coordinates facial landmarks
Step1: Initialize MTCNN for face detection
Step 2: Pass video frame to MTCNN for processing
Step 3: MTCNN detects face coordinates in the frame
Step4: MTCNN extracts facial and marks from the face
Step 5: Return face coordinates and facial landmarks.
Step 5: Return accuracy score.

iv) Feature Extraction Algorithm (FaceNet/VGGFace):
Input: face coordinates, facial and marks
Output: face embedding
Step1: Initialize Face Net or VGG Face model
Step2: Crop the face from the video frame using face coordinates
Step3: Pass the cropped face and facial and marks to the model
Step4: The model generates face embedding (numerical representation
Step5: Return face embedding

v) Graph Comparison Algorithm (Euclidean Distance):
Input: face embedding, stored confident embeddings, stored unconfident embeddings
Output: Confidence score
Step1: Initialize comparison metric (Cosine Similarity or Euclidean Distance)
Step2: Compare face embedding with to red confident embeddings
Step3: Compare face embedding with to red unconfident embeddings
Step4: Calculate similarity scores for both comparisons
Step5: Calculate confidence score based on higher similarity with confident embeddings
Step6: Return confidence score.

vi) Decision Logic Algorithm:
Input: accuracy score, confidence score
Output: pass or fail (boolean)
Step1: Set pass thresh old=80
Step2: If accuracy score =pass threshold AND confidence score ¿= pass threshold: pass or fail = PASS Else: pass or fail = FAIL
Step3: Return passor fail.

vii)Data Storage and Notification:
Input: userid, accuracy score, confidence score, pass or fail Output: success message (for notification)
Step1: Initialize connection to Fire base Fire store
Step2: Store userid, accuracy score, confidence score, and pass or fail in Fire store
Step3: Send success message to front end using Socket.IO
Step4: Notify user of their result (passor fail)
Step5: Return success message.

SYSTEMARCHITECTURE
The front end captures live audio and video streams using React, Vite, and Web RTC. Real-time data is sent to the backend via Socket.IO, where the audio is processed using an NLP pipe line, and video is analyzed by a CNN-based facial detection system. Results are aggregated and sent back to the front end for reels are aggregated and sent back to the front end for real-time display.

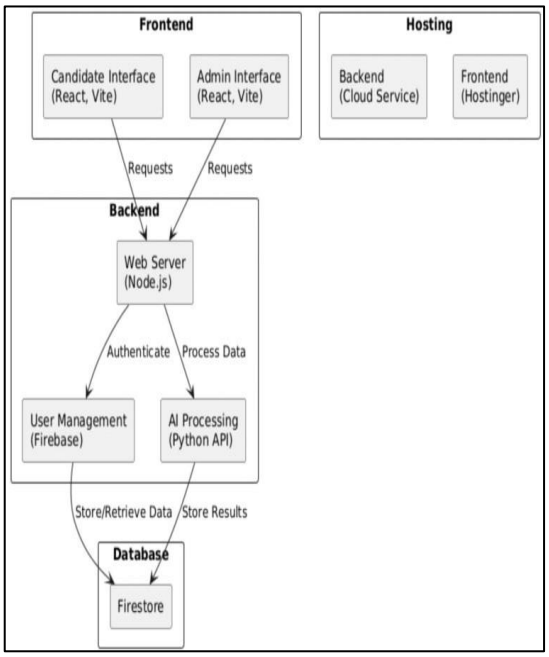

Fig1: Architecture of Interview X

CNN Pipe line for Facial Detection Recognition:

- **Facial Detection:** MTCNN (Multi-task Cascaded Convolutional Neural Networks): This is a deep learning model specifically designed for detecting faces in images. It works by generating bounding boxes around detected faces and identifying key facial landmarks, such as eyes, nose, and mouth. Themodel uses a series of cascaded CNNs to process images at multiple stages, improving accuracy by refining the face location and landmark points at each stage. MTCNN is highly effective for detecting faces in various lighting conditions, angles, and backgrounds, making it suitable for real-world applications.
- **Facial Recognition: Face Net:** Once faces are detected, the next step is facial recognition. Face Net is a deep neural network that generates facial embeddings unique numerical representations of faces by analyzing facial features. These embeddings capture essential details and distinguish one face from another. In this process, each face is converted into a 128 vector. Comparing Embeddings: To recognize a face, Face Net compares the embeddings of the detected face with thoseinadatabaseofknownfaces.Euclideandistance is typically used to measure the similarity between embeddings. A smaller distance indicates a closer match, meaning the more similar.
- **Confidence Scoring: Similarity Threshold:** During recognition, a confidence score is assigned to each match based on the similarity of the embeddings. This score represents how confident the system is that the detected face matches a known face in the database. The confidence is calculated by comparing theEuclideandistancebetweenembeddingsagainst a predefined threshold. If the distance is below the threshold, it indicates a match with high confidence; otherwise, it is likely not a match. Assigning Scores: The closer the distance to zero, the higher the similarity, and thus the higher the confidence scores. This scoring system helps determine if the detected face is an accurate match to a known face or if it is unfamiliar

CONCLUSION

In this paper the Interview X project creates a realistic interview environment using AI to help users improve their interview skills. By analyzing answers and facial expressions, it gives feedback on verbal and non- verbal communication. This project combines modern technologies like React, Firebase, and AI algorithms to provide a useful tool for job seekers to practice and for companies to streamline their hiring process. Overall, Interview X shows how AI can make interview preparation and assessment more efficient and effective.

FUTURESCOPE

- Expansion of the system to evaluate group interviews or panel interviews.
- Integrating additional behavioral cues, such as tone analysis or eye contact detection.
- Development of industry-specific interview simulations tailored to technical, HR, or case- based interviews.

REFERENCE

Xinyi Luo1, Yuyang Wang, LikHang Lee, Zihan Xing,ShanJin,BoyaDong,YuanyiHu,ZemingChen,Jing Yan, Pan Hui, “Usingavirtualrealityinterviewsimulatorto explore factors infuencing people’s behavior”,28 February 2024

Dr. Surendra Mahajan, Prof Nilesh Sonawane, Rohit Bagal, Saket Kulkarni, AyushSuryawanshi, YadnyeshMhaisne,“GenerativeAI-BASEDInterviewSimulationand Performance Analysis”,International Research Journal of Modernization in Engineering Technology and Science Volume:06/Issue:05/May-2024

Pankaj Rambhau Patil, Shinde Rushikesh Rajendra, Gosavi Vishakha Mahendra, Bhamare Bhagyashri Jijabrao, Patil Paresh Dilip, “Elevating Performance Through AI-DrivenMockInterviews”,InternationalJournalfor Research in Applied Science Engineering Technology (IJRASET)ISSN:2321-9653;ICValue:45.98;SJImpact Factor: 7.538 Volume 12 Issue VI June 2024 Available at www.ijraset.com

Amitha Caldera, SathiraHettiarachchi,Bandara H.M.R.M, Abeywickrama Y.S “Automating Recruitment Processusing Natural”, International Research Journal of Engineering and Technology(IRJET)e-ISSN:2395-0056Volume:10Issue: 10-Oct2023 www.irjet.net

RohanPatil , Akash Butte , Sahil Temgire , Varun Nanekar,Asst.Prof.ShivgangaGavhane“RealTime Mock Interview using Deep Learning”,International Journal of Engineering Research Technology (IJERT) ISSN: 2278- 0181 IJERTV10IS050213 Vol. 10 Issue 05, May2021