An AI-Driven Mock Interview System Using NLP and Computer Vision

Premala[1*], Patki Parth[2], Rakesh[3] ,Veeresh Mulge[4], Vinayak Rao[5]

[1,2,3,4,5]Department of Computer Science and Engineering, Guru Nanak Dev Engineering College Bidar, Karnataka, India.

*Corresponding Author Email: parthpatki4007@gmail.com

Abstract

As the global job market becomes increasingly competitive, many students have realized that strong academic performance alone is insufficient to ensure success in interviews. Recognizing this gap, the present study introduces an AI-driven mock interview system designed to evaluate not onlywhatcandidatessaybutalsohowtheypresentthemselves. The system integrates natural language processing with computervisiontechniquestoassessbothverbalresponsesand nonverbal cues, such as eye contact and facial expressions. To simulate realistic interview scenarios, advanced language models were used to generate domain-specific questions in real time,deliveredthroughaninteractiveaudio-videointerface.In contrast to traditional tools that primarily focus on textual evaluation, the proposed approach provides a more comprehensive understanding of a candidate's communication skills, including the clarity, confidence, and behavioral aspects of their speech. An experimental study conducted with 50 engineering students demonstrated that the inclusion of eye-gaze tracking improved confidence detection accuracy from 72% to 89%, outperforming systems based solely on textual analysis. User feedback further indicated that the system was intuitive and effective, suggesting its potential as a practicaland scalable solution for improving interview readiness.

Keywords- Artificial Intelligence, Mock Interview System, Natural Language Processing, Computer Vision, Eye-Gaze Tracking, Multimodal Evaluation, Real-Time Assessment.

1. Introduction

In the current employment landscape, obtaining a job extends beyond merely holding a degree. While academic credentials may grant initial entry, a candidate's interview performance—encompassing effective communication, stress management, and confidence projection— ultimately securestheposition.Despitethis,manyindividualscontinue to utilize outdated preparation techniques. Traditional methods, such as engaging in mock interviews with acquaintances or memorizing frequently asked questions, often prove inadequate, lacking both objective feedback and the genuine experience of a high-pressure interview.[1]
Artificialintelligence(AI)cansignificantlycontributeinthis regard. By leveraging Natural Language Processing (NLP) and Computer Vision (CV), it is now possible to analyze human interactions in novel ways, considering both verbal expressions and nonverbal cues.[2] Our research introduces an AI-driven mock interview platform designed to address this gap. Instead of solely evaluating the correctness of answers, the system assesses a spectrum of elements, from technical expertise to nonverbal assurance, in real time. Our aim was to create a practical and user-friendly tool that empowers students and professionals to move beyond theoreticalunderstandingandmastertheartofinterviewin g.

A. ProblemStatement

Even with so many interview resources available today, there's still a huge gap when it comes to finding a practice environment that actually feels "real." Most of the platforms outtherererelyonthesameoldstaticquestionbanksorbasic automatedtoolsthatjustcan'treplicatethepressureorthe flow of a live interview.[3]

The biggest issue is that these systems almost entirely ignore nonverbal communication. In a real interview, how you look and act—your eye contact, posture, and overall presence—is just as important as what you say. Without a way to analyze these cues, it is incredibly difficult for a candidate to know if they are actually ready or where they are falling short. [4]Because they do not receive real-time, comprehensive feedback, many students head into their actual interviews feeling underprepared and anxious, rather than confident and coached.

B. Motivation

This study was motivated by the design of an intelligent system that offers a more realistic and comprehensive interview preparation experience. By leveraging advancements in artificial intelligence, the proposed system hasmovedbeyondconventionalapproachesthatfocussolel y on textual and verbal analysis. Instead, it integrates both verbal and behavioral evaluations to provide a more comprehensive understanding of candidate performance.

Another key consideration is the growing adoption of remote hiring processes, in which virtual interviews have become increasingly prevalent. [2]This shift highlights the need for platforms that can effectively replicate such environments while offering meaningful feedback. By enabling real-time analysis and continuous performance tracking, the system aims to help users gradually improve theircommunicationskillsandconfidencelevels.Ultimately , thegoalistoprovideanaccessibleandscalablesolutionthat

supports candidates in preparing more effectively for real-world interviews.

C. Objective

The primary objectives of the proposed system are as follows:

- TodevelopanAI-basedmockinterviewplatformthat can simulate realistic real-time interview scenarios. [5]
- Natural language processing (NLP) techniques are used to evaluate the relevance, clarity, and quality of candidate responses.
- Computer vision methods can be applied to analyze nonverbal cues, such as facial expressions and eye gaze.
- AI-based mock interview platform that can simulate realistic real-time interview scenarios.
- To design a multimodal evaluation framework that integrates both verbal and behavioral analyses.[6]
- To provide real-time personalized feedback along with comprehensive performance scoring.[7]

2. Related Work

Recent developments in artificial intelligence (AI) and machinelearninghaveenabledthedevelopmentofadvanced systemsaimedatenhancinginterviewpreparationandcareer readiness.Numerousstudieshaveinvestigatedplatformsthat simulate real-world interview scenarios and offer personalized feedback to enhance candidates' verbal, nonverbal, and cognitive skills.

A. AI-PoweredMockInterview Platforms

AI-driven mock interview platforms have been developed to enhance interview preparation by integrating natural language processing (NLP) and machine learning (ML) techniques.

- InterviewEaseintegratesresumeanalysisandamock interview module to improve verbal and nonverbal skills.[8]
- AI Mock Interviewer customizes questions based on resumes and provides voice analysis to improve articulation and confidence. [9]
- Platforms such as Aspire-Hub and Automated Interview Evaluation offer tailored practice modules and real-time feedback.[10], [11]

Recent systems have leveraged AI to assess personality and behavioral traits. TensorFlow-based Automatic Personality Recognition evaluates “Big Five” personality traits through asynchronous video interviews, providing an objective,bias-reducedassessment[5].SystemssuchasPrep AI and Full StackAI Mock InterviewApp employ NLPand generative AI to dynamically evaluate communication, confidenceandcontentrelevanceinrealtime.[12],[13],[14]

B. Eye-TrackingandNonverbalSkillAssessment

Nonverbalcuesarecriticalfortheeffectiveevaluationof interviews. Eye-tracking technology integrated with AI enhances behavioral analysis.

- Affordable Sensor-Based Gaze Tracking and MLGaze correct gaze errors and track eye movements effectively using machine learning.[15], [16]
- Deep learning methods improve the robustness and precision in virtual and unconstrained environments by addressing occlusions and reflections.[17], [18]
- Systems such as DeepVisionAnalytics combine eye tracking, facial recognition, and machine learning to objectively infer engagement and decision-making. [19]

C. IntegrationofAIforCareer Readiness

AI-driven platforms are increasingly integrating multimodal analyses for holistic evaluations.

- Platforms that leverage OpenAI LLMs and MediaPipe provide real-time feedback on facial expressions, confidence, and content quality. [20], [21]
- Many systems allow candidates to customize interview types and difficulty levels, thereby improving their skill development and user satisfaction.
- OpenAI LLMs and MediaPipe provide real-time feedback on facial expressions, confidence and content [12][13].
- Future enhancements suggested in prior research include gesture detection, coding environment integrationandimmersivevirtualrealitysimulations. [8], [13], [14]

D. ResearchGap

Despite these advancements, existing systems have their limitations.

- Many studies focus only on verbal or nonverbal evaluations,lackingaunifiedmultimodalframework. [22]
- Real-time interactions and adaptive question generation are often absent.
- Feedback mechanisms are generally generic and inactionable.
- The scalability and usability of diverse user profiles are insufficient.

3. Methodology

The proposed AI-driven multimodal mock interview system simulates live interviews and provides a real-time evaluation of verbal and nonverbal responses. The architecturalframeworkincorporatestechniquesfromnatural languageprocessing(NLP)andcomputervision(CV)within a modular, low-latency structure.

A. SystemArchitecture

The system employs a multi-layered architecture with discrete functional modules:

- Authentication: Users register and login securely via Firebase or OTP-based verification, enabling session tracking and historical performance analysis.
- Profile Setup and Input Configuration: Candidates select their domain (e.g., Software Development,

HR), target company, and professional experience level.

- Dynamic Question Generation: Gemini AI LLM generates five domain-specific interview questions tailored to the candidate's profile.
- Interview Execution: Live video and audio sessions were conducted using WebRTC protocols, ensuring high-fidelity streaming and real-time capture of candidate responses.
- StorageandRecording:Audio,video,andtranscribed responses were securely stored in Firebase for asynchronous analysis and longitudinal tracking.
- AIEvaluationEngine:

a) NLP Module: Analyzes transcribed responses for semantic relevance, conceptual accuracy, and clarity.

b) Computer Vision Module: Tracks facial landmarks and eye gaze to calculate confidence scores.

c) Dual-Stream Fusion: Combines verbal and nonverbal metrics to derive a holistic performance evaluation.

- Feedback and Analytics: An overall answer evaluationscore(outof10)isprovided,withdetailed performance metrics and targeted suggestions for improvement.

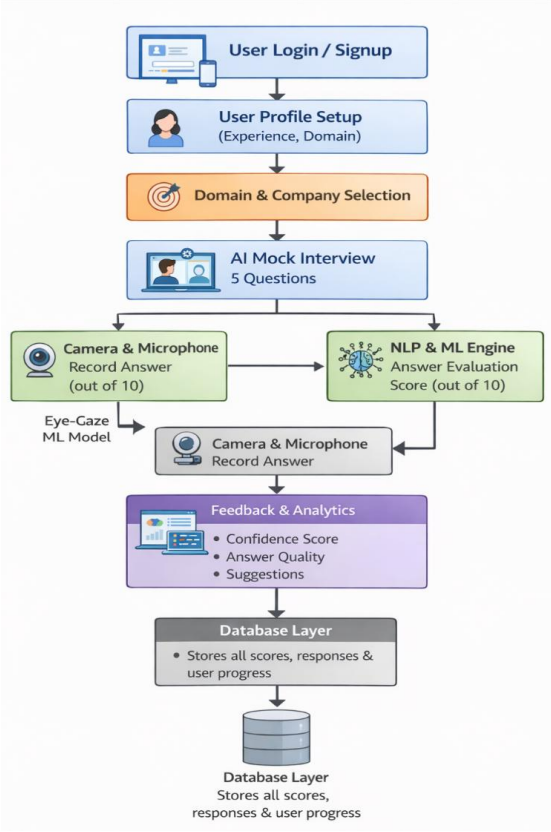

Fig.1.SystemFlowDiagram

E. *System Implementation*

The system was implemented as a scalable web application that enabled real-time interaction.

- Frontend: React.js provides an interactive and responsive user interface.
- Backend: Node.js and Express.js are used to manage API requests, session management, and data processing.
- Database: MongoDB stores user profiles, interview data, and performance metrics.
- Real-Time Communication: WebRTC enables live audio-video streaming between candidates and systems.
- Authentication: OTP-based verification and firebase authentication ensure secure access for users.
- PracticalViability:Withlow-latencyresponsetimeof approximately 460 ms, the system successfully mirrors the fluidity of a live interview, a claim supported by high participant ratings for realism and usability.

Ultimately, these findings substantiate the efficacy of multimodal analysis in interview training. By offering a scalable and intuitive platform, this study provides students and professionals with a powerful tool to cultivate self-confidence and master the complexities of high-stake competitive interviews.

F. *WorkflowOverview*

- Thecandidateloginandsetupprofilepreferences.
- AIdynamicallygeneratestailored,interviewquestions.
- The candidates participated in a live, recorded mock interview.
- TheNLPandCVmodulesanalyzeverbaland nonverbal responses.
- Thesystemdeliversactionablefeedbackandperformance analysis.

4. Results and Discussion

The AI-based multimodal mock interview system was tested on 50 engineering students in the software development, data structure, and human resource domains. Each participant completed a single interview session with fiveAI-generated,role-specificquestionstailoredtotheir domain, target company, and experience level. System performance was assessed based on accuracy, precision, recall, and response latency.

D. *Sample Participant Scores*

User ID	Domain	Answer Quality	Confidence	Overall Score
User01	Software Dev	8.5	7.8	8.15
User02	HR	7.2	6.9	7.05
User03	Data Structures	9.1	8.6	8.85
User04	Software Dev	6.8	6.5	6.65
User05	Frontend Dev	8.9	9.2	9.05

TABLE I. Mock Interview Evaluation Scores

The dual-axis evaluation captures cognitive performance (answer quality) and nonverbal self-confidence (confidence), thereby enabling a holistic assessment of candidates.

E. System Performance Comparison

Metric	NLP-Only	Eye-Gaze + NLP
Confidence Accuracy	72%	89%
Precision	0.74	0.88
Recall	0.71	0.86
Avg. Latency	420 ms	460 ms

TABLE II. System Performance Metrics

F. Observations

- Confidence detection was improved by 17% with gaze cues.
- Participants who maintained direct eye contact scored 20–25% higher on the confidence metric.
- NLP reliably evaluates response relevance, clarity, and conceptual correctness.

The combined NLP and eye-gaze approach distinguishes between technically strong but timid candidates and self-confident yet conceptually weak candidates, a distinction absent that is in NLP-only systems.

G. User Evaluation and Experience

The survey results indicated high user satisfaction.

- Ease of Use: 4.5 / 5
- Accuracy of Feedback: 4.4 / 5
- Real-World Usefulness: 4.6 / 5

Participants found the system intuitive, actionable, and realistic, confirming its effectiveness for interview preparation and confidence-building.

H. Discussion

This study demonstrates that multimodal analysis (verbal NLP + non-verbal eye-gaze ML) provides a more holistic assessment than NLP-only systems.

- Enhanced Accuracy: Visual tracking significantly improves confidence assessment.
- Behavioral Insights: Eye gaze analysis captures subtle cues of self-confidence.
- Actionable Feedback: Combined scoring provides guidance to improve both content and delivery.
- Practical Deployment: Low latency (~460 ms) and positive user experience enable real-time and scalable use.

The dual-axis framework enables a richer and more nuanced evaluation of candidates by bridging the gap between technical proficiency and behavioral competence.

5. User Interface and Output

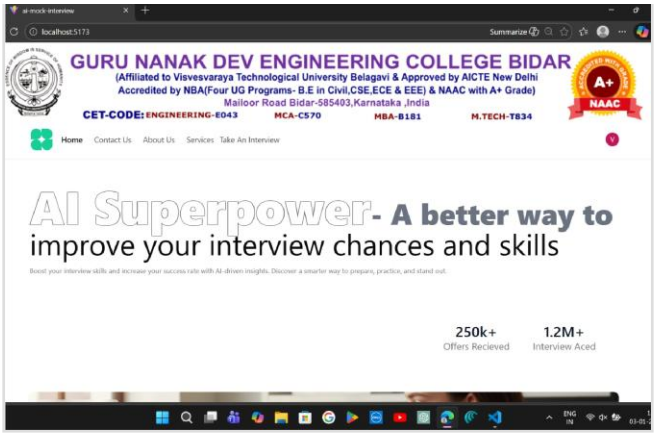

Fig. 2. Home Screen - The home interface serves as the central navigation hub, allowing users to access various features of the system, including profile management, interview initiation, and performance tracking.

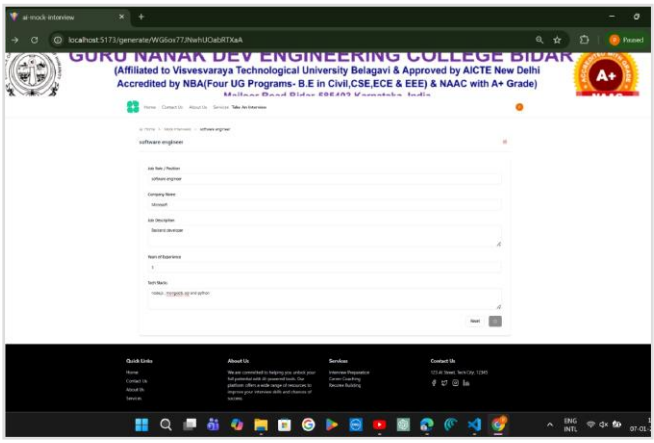

Fig.3. Form Filling Page - This interface enables users to configure their interview session by selecting the target domain, company, and experience level. These inputs are used to generate personalized interview questions.

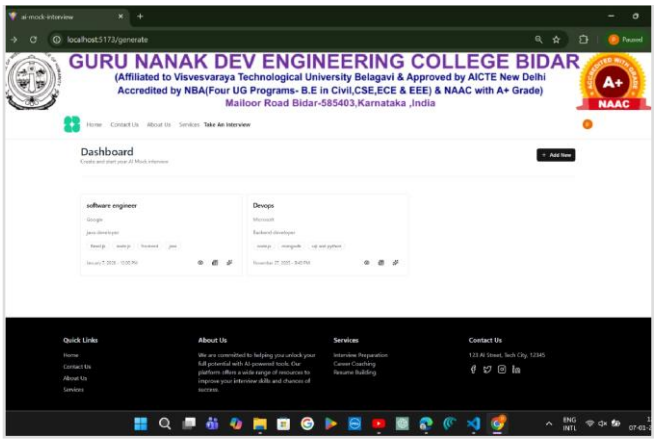

Fig.4. Dashboard - The dashboard provides an overview of user profile details, history of past interview sessions, and performance analytics, enabling users to monitor their progress over time.

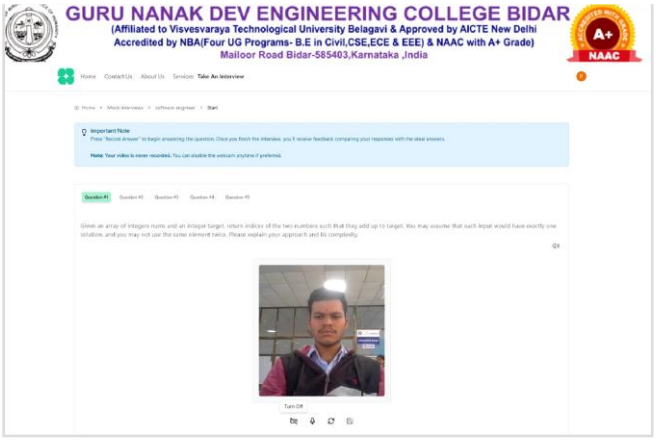

Fig. 5. Mock Interview Interface - During the interview, users respond to dynamically generated questions in real time. The system captures audio and video inputs through an integrated interface.

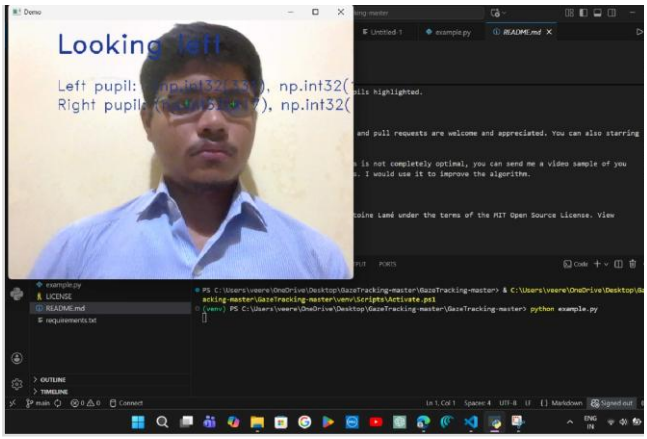

Fig. 6. Eye-Gaze ML Model Analysis - This screen illustrates the real-time processing of facial expressions and eyegaze tracking. These behavioral features were used to compute the confidence score of the candidate.

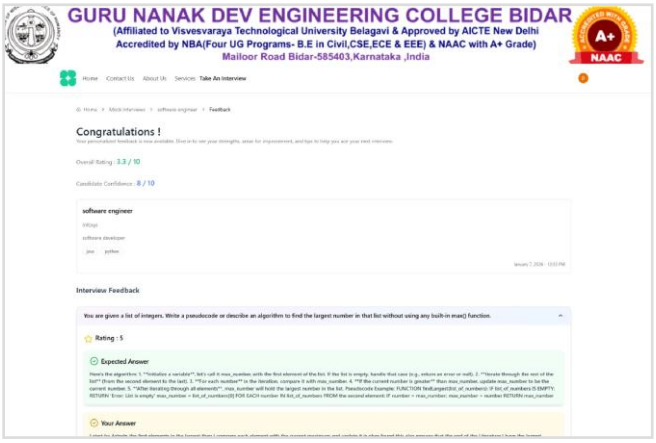

Fig.7. Result and Feedback Page - The final output screen displays the evaluation results, including the answer quality score, confidence score, and detailed feedback

6. Future Work

Although our current results are promising, there is still significant room for improvement in making this platform more immersive. Our next step was to introduce 3D AI avatars into the interface. By moving away from a static screen and toward an embodied "digital interviewer," we can better simulate the actual psychological pressure candidates experienceduring a face-to-face meeting. These avatars won't just sit there—they'll provide real-time visual feedback through micro-expressions and gestures, making the simulation feel far more human.[2]

In addition the visuals, we want to deepen the system's "understanding" of a candidate's behavior. We plan to expand our framework to include gesture recognition, posture analysis, and voice tone (prosody) in the future. This will allow us to provide feedback on how someone's body language and vocal delivery either support or undermine their technical answers.

We also see a major opportunity for adaptive questioning. Instead of a fixed script, we want the AI to adjust the difficulty of its questions based on how the candidate is performing, effectively "stress-testing" their composure in real time. Eventually, we hope to roll this out across specialized fields such as law and healthcare and even integrate it directly with hiring platforms. The ultimate goal is to move from a simple practice tool to a lifelong career coach that tracks a user's growth from their first mock interview to their first big career move.[23]

7. Conclusion

This study introduces a real-time, AI-driven multimodal framework that bridges the critical gap between technical content mastery and professional behavioral presentations. By synthesizing natural language processing with computer vision-based eyetracking, the system provides a holistic evaluation that traditional single-mode preparation tools lack.

Empirical testing with 50 engineering students yielded several key insights.

- Enhanced Precision: Integrating eye-gaze analysis increased confidence detection by 17% over NLP-only benchmarks, validating the necessity of nonverbal data in behavioral assessment.
- Actionable Feedback: The dual-axis scoring system effectively decoupled technical knowledge from communication style, allowing for nuanced, data-driven feedback.[24]
- Practical Viability: With low-latency response time of approximately 460 ms, the system successfully mirrors the fluidity of a live interview, a claim supported by high participant ratings for realism and usability.

Ultimately, these findings substantiate the efficacy of multimodal analysis in interview training. By offering a scalable and intuitive platform, this study provides students and professionals with a powerful tool to cultivate self-confidence and master the complexities of high-stake competitive interviews.

References

1. S. M. Rahman, M. A. Hossain, M. S. Miah, M. M. Alom, and M. Islam, "Artificial Intelligence (AI) in Revolutionizing Sustainable Recruitment: A Framework for Inclusivity and Efficiency," IRJMS, vol. 06, no. 01, pp. 1128–1141, Jan.2025, doi: 10.47857/irjms.2025.v06i01.02698.
2. Y.-C. Chou, F. R. Wongso, C.-Y. Chao, and H.-Y. Yu, "An AI Mock-interview Platform for Interview Performance Analysis," Institute Of Electrical Electronics Engineers, Apr. 2022, pp. 37–41. doi: 10.1109/iciet55102.2022.9778999.

3. A. L. Heimann and A. Schmitz-Wilhelmy, "Observing Interviewees' Inner Self: How Authenticity Cues in Job Interviews Relate to Interview and Job Performance," J Bus Psychol, vol. 40, no. 1, pp. 237–256, May 2024, doi: 10.1007/s10869-024-09949-4.
4. M. P. Martín-Raugh, H. J. Kell, J. G. Randall, C. Anguiano-Carrasco, and J. T. Banfi, "Speaking without words: A meta-analysis of over 70 years of research on the power of nonverbal cues in job interviews," J Organ Behavior, vol. 44, no. 1, pp. 132–156, Oct. 2022, doi: 10.1002/job.2670.
5. C. Pyle, K. Roemmich, and N. Andalibi, "U.S. Job-Seekers' Organizational Justice Perceptions of Emotion AI-Enabled Interviews," Proc. ACM Hum.-Comput. Interact., vol. 8, no. CSCW2, pp. 1–42, Nov. 2024, doi: 10.1145/3686993.
6. L. S. Nguyen, D. Frauendorfer, M. S. Mast, and D. Gatica-Perez, "Hire me: Computational Inference of Hirability in Employment Interviews Based on Nonverbal Behavior," IEEE Trans. Multimedia, vol. 16, no. 4, pp. 1018–1031, June 2014, doi: 10.1109/tmm.2014.2307169.
7. M. A. Rahman, D. J. Brown, N. Shopland, A. Burton, and M. Mahmud, "Explainable Multimodal Machine Learning for Engagement Analysis by Continuous Performance Test," Springer, 2022, pp. 386–399. doi: 10.1007/978-3-031-05039-8_28.
8. P. Kothari, P. Mehta, S. Patil, and V. Hole, "InterviewEase: AI-Powered Placement Assistance," Springer Nature Singapore, 2025, pp. 297–308. doi: 10.1007/978-981-96-0692-4_24.
9. I. Journal, "AI Mock Interviewer: Preparing You Towards Smarter Interview Practice," IJSREM, vol. 08, no. 01, pp. 1–10, Jan. 2024, doi: 10.55041/ijsrem28199.
10. C. Sri Latha, N. Krishna Kalyan, J. Abhilash, O. V. S. N. Kaushik, V. Balmiki, and S. Venukumar, "Automated Interview Evaluation," E3S Web Conf., vol. 430, p. 01026, Jan. 2023, doi: 10.1051/e3sconf/202343001026.
11. S. Kangude, A. Karwande, S. Kulkarni, and R. Pagare, "Aspire-Hub: All in one Career Readiness App," IJSREM, vol. 08, no. 11, pp. 1–8, Nov. 2024, doi: 10.55041/ ijsrem38462.
12. V. Patil, "AI Mock Interview Platform," IJSREM, vol. 09, no. 06, pp. 1–9, June 2025, doi: 10.55041/ijsrem49510.
13. K. D. Bhavani, G. H. S. V. K. Prasad, V. G. S. S. Krishna, M. S. K. Reddy, K. S. Vardhan, and M. Sadik, "Prep AI - Customized Mock Interview Platform Using Gen AI," International Journal of Innovative Science and Research Technology (IJISRT), p. 1981, Jan. 2026, doi: 10.38124/ijisrt/26jan835.
14. U. Bhadane, L. Jagtap, S. Rane, R. Dhonnar, and P. Pawar, "AI Based Interview Mocker System," IJARSCT, pp. 23–30, Apr. 2025, doi: 10.48175/ijarsct-25205.
15. A. Kar, "MLGaze: Machine Learning-Based Analysis of Gaze Error Patterns in Consumer Eye Tracking Systems," May 07, 2020. doi: 10.48550/arxiv.2005.03795.
16. R. Gavas et al., "Affordable sensor based gaze tracking for realistic psychological assessment.," Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, vol. 2017, July 2017, doi: 10.1109/embc.2017.8036932.
17. I. Schuetz and K. Fiehler, "Eye Tracking in Virtual Reality: Vive Pro Eye Spatial Accuracy, Precision, and Calibration Reliability," JEMR, vol. 15, no. 3, Sept. 2022, doi: 10.16910/jemr.15.3.3.
18. K. Barkevich, R. Bailey, and G. J. Diaz, "Using Deep Learning to Increase Eye-Tracking Robustness, Accuracy, and Precision in Virtual Reality.," Proc. ACM Comput. Graph. Interact. Tech., vol. 7, no. 2, pp. 1–16, May 2024, doi: 10.1145/3654705.
19. J. A. Rodrigues, A. V. D. Castro, and M. Llamas-Nistal, "Advanced Consumer Behaviour Analysis: Integrating Eye Tracking, Machine Learning, and Facial Recognition.," JEMR, vol. 19, no. 1, p. 9, Jan. 2026, doi: 10.3390/jemr19010009.
20. V. P, P. K. S. Rakunde, P. N U, S. Doggalli, and I. T. Yeptho, "An AI-Powered Platform for Mock Interviews and Career Preparation Using OpenAI (LLM's) and Google MediaPipe," IJSREM, vol. 09, no. 12, pp. 1–9, Dec. 2025, doi: 10.55041/ijsrem55442.
21. R. Jadhav, S. Kapse, S. Mane, and N. Bhagwat, "Advancements in AI-Enabled Mock Interview Platforms: A Review of Interviewello's Web Application," ITSI-TEEE, vol. 13, no. 2, pp. 30–35, Mar. 2025, doi: 10.65521/itsi-teee.v13i2.51.
22. M. Krishna Pasupuleti, "AI-Enhanced Virtual Reality: Transforming Immersive Training and Simulation for Real-World Skill Development," National Education Services, 2024, pp. 96–113. doi: 10.62311/nesx/66241.
23. H. Kumazaki et al., "Job interview training targeting nonverbal communication using an android robot for individuals with autism spectrum disorder.," Autism, vol. 23, no. 6, pp. 1586–1595, Feb. 2019, doi: 10.1177/1362361319827134.
24. H. Ashraf, M. H. Sodergren, N. Merali, G. Mylonas, H. Singh, and A. Darzi, "Eye-tracking technology in medical education: A systematic review," Medical Teacher, vol. 40, no.1, pp. 62–69, Nov. 2017, doi: 10.1080/0142159x.2017.1391373.P. Kothari et al.,"InterviewEase : AI-powered interview assistance," Research Square Platform LLC, Feb. 28, 2024, doi: 10.21203/ rs.3.rs-3964944/v1.