iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80521

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Cognify AI: A Multimodal AI-Based Mock Interview System for Real-Time Evaluation of Verbal and Non-Verbal Communication

Aadil Raja[1], Kasa Chanakya[2], Satwik M Badiger[3], M N Kavya[4], Ajil Aylwin[5]
[1, 2, 3, 4]Scholar, [5]Associate Professor, Department of Computer Science and Engineering, Reva University, Bengaluru, Karnataka, India,560064

***Abstract:** Although non-verbal communication critically impacts interview success, traditional mock interview platforms predominantly evaluate isolated speech and text. To overcome this limitation, we introduce Cognify AI, a multimodal system that integrates vision-based behavioral analysis (via DeepFace and OpenCV) with locally executed NLP-based text evaluation within a single real-time processing framework. Unlike single-modality systems that process visual and textual streams in silos, Cognify AI simultaneously correlates eye contact stability,facial emotional patterns, and the presence of key semantic concepts in responses. Our evaluation demonstrates that this inte- grated feedback loop yields measurable improvements in both technical articulation and non-verbal composure across successive candidate sessions, providing actionable insights that tangibly enhance interview preparedness.*

***Keywords:** multimodal AI, interview preparation, facial sentiment analysis, natural language processing, pose estimation, computer vision, real-time feedback*

I. INTRODUCTION

Success in modern recruitment relies heavily on a candidate's non-verbal factors—including gaze behavior, emotional control, and posture—which can significantly influence interviewer perception alongside technical expertise. The transition to virtual recruitment has amplified the challenge of conveying these subtle cues effectively. Despite this, most existing automated interview systems mainly emphasize transcript-based evaluation and verbal performance metrics. Consequently, candidates receive granular feedback on their spoken content but lack actionable insights regarding their behavioral presence.

Cognify AI addresses this fundamental gap by proposing a unified multimodal framework. By synchronously processing real-time computer vision metrics and NLP-driven content evaluation, our system replicates the holistic assessment of a human interviewer. While prior solutions process visual and textual streams in isolation, Cognify AI distinctively fuses these modalities to provide deterministic, actionable feedback on both what a candidate says and how they deliver it. Our primary objective is to equip candidates with specific behavioral and technical insights that systematically reduce communication anxiety and improve overall performance.

II. LITERATURE SURVEY

Automated interview preparation has evolved from static rule-based models to dynamic, context-aware systems driven by large language models. However, despite the rapid advancements in the capture of speech, vision, and text, we observe that many current platforms continue to process different data modalities independently without meaningful integration. This fragmentation drives the architectural need for Cognify AI, which weaves multimodal analysis into a singular, unified evaluation pipeline.

A. Text-Based and Rule-Based Interview Systems

The initial automated evaluation systems were built predominantly on rule-based logic and simple keyword matching techniques to evaluate candidate responses. As natural language processing matured, later approaches adopted machine learning models that interpret underlying semantic meaning instead of relying solely on explicit word patterns [1]. While these text-centric systems successfully automate baseline technical evaluations, our analysis highlights a critical limitation: they are inherently blind to non-verbal behavioral cues [2]. Because human interviewers weigh presence and delivery as heavily as content, systems confined solely to transcript analysis cannot provide a fully representative assessment.

B. Multimodal Interview Systems

To capture the behavioral dimensions of an interview, recent scholarship has increasingly explored multimodal architectures. Foundational work in this domain established that combining facial cues, vocal characteristics, and linguistic features has been shown to produce predictions that closely reflect human evaluation outcomes [3]. Subsequent frameworks have refined this by utilizing deep learning to process these data streams simultaneously [4, 5]. Additionally, emotion recognition models have been deployed to classify candidate confidence and sentiment [6]. However, achieving true real-time multimodal fusion remains a persistent architectural challenge. Many existing solutions ultimately prioritize one modality over the others or fail to synthesize the outputs into a cohesive score. Cognify AI differentiates itself by maintaining parallel, synchronous pipelines for both computer vision and speech-to-text, ensuring neither domain is marginalized.

C. LLM-Based Interview Systems

The introduction of large language models enabled a transition from rigid scoring rules to more flexible, context-sensitive evaluation methods [7]. Recent implementations leverage LLMs to automate academic grading [8], integrate with virtual MetaHuman avatars to create immersive training environments [9], simulate multi-agent interviewer-candidate interactions [10], and generate highly granular pedagogical feedback [11]. Despite these advancements, these models are inherently limited to textual inputs and cannot independently interpret visual behavioral signals. We address this limitation by explicitly pairing an LLM-driven semantic scoring engine with deterministic computer vision metrics.

D. Speech and Audio-Based Interview Systems

Parallel research tracks have isolated specific sensory channels for deeper analysis. Speech-focused platforms have advanced to offer real-time verbal analytics, evaluating fluency, structural coherence, and tone [12–14]. Yet, by focusing exclusively on the verbal channel, they leave out the visual story entirely—a limitation that multimodal frameworks are well positioned to address.

E. Behavioral and Computer Vision-Based Interview Systems

Conversely, computer vision studies have empirically validated the impact of gaze behavior on interviewer perception [15], leading to the development of dedicated eye-tracking assessment tools [16]. Recent applications have also begun integrating facial analysis with voice recognition for localized simulation purposes [17]. Yet, unlike prior work that treats visual and verbal analytics as distinct evaluative outputs, Cognify AI mathematically and structurally synthesizes these metrics. By establishing a unified evaluation layer, our system directly resolves the single-modality bottlenecks prevalent in current literature.

Table 1: Comparison of Existing Interview Systems

Paper	Text Analysis	Speech	Vision	Multimodal
[1]	✓	×	×	×
[4]	✓	✓	✓	✓
[8]	✓	×	×	×
[14]	×	✓	×	×
[16]	×	×	✓	×
Cognify AI (Proposed)	✓	✓	✓	✓

III. SYSTEM DESIGN AND ARCHITECTURE

The architecture of Cognify AI is designed as a modular, multimodal framework capable of processing both visual and textual data simultaneously. The system is structured to handle real-time data efficiently while maintaining low latency, even in typical user environments. To achieve this, the architecture follows a layered design that decouples data processing, analysis, and evaluation.

A. Layer 1: User Interface

The user interface is designed with usability and clarity in mind to ensure that candidates can interact with the system without unnecessary complexity. Since the platform involves recording and transmitting audio and video data, the interface also supports efficient handling of media to maintain smooth operation.

After authentication, candidates are redirected to a dashboard that displays previous interview results, overall performance trends, strengths, weaknesses, and feedback on areas for improvement. The dashboard also provides an option to proceed to the interview setup phase.

During the interview setup phase, candidates can upload a resume, a job description, or both, depending on the desired level of personalization. The system uses these inputs to generate tailored interview questions. Candidates can also configure parameters such as interview duration, difficulty level, focus area, and number of questions. Additional features such as video analysis and speech analysis can be enabled based on user preference.

Once the configuration is complete, the system validates the inputs and sends the selected data to the backend for question generation, ensuring a smooth and efficient setup process.

B. Layer 2: Backend API and Data Management

1) Question Generation Service

The question generation service is responsible for creating personalized interview questions based on candidate-provided inputs. It receives configuration settings and optional documents such as resumes and job descriptions from the setup phase.

The system extracts textual content from uploaded files and combines it with interview parameters to construct a dynamic prompt. A locally hosted language model (Ollama) is then used to generate context-aware interview questions in a structured JSON format.

Each generated question includes metadata such as type, difficulty level, and guidance hints. The system stores the generated questions along with session data, enabling efficient retrieval and continuity across different phases of the interview.

2) Answer Processing and Evaluation Service

This service handles candidate responses during the interview phase and performs multimodal analysis. Video responses are uploaded to the backend, where they are stored in a structured session-based format.

The system extracts audio from each video and processes both modalities independently. Audio is converted into a transcript, while video is analyzed to extract behavioral features such as eye contact and facial expressions.

An AI-based evaluation module then assesses the response by analyzing content accuracy, identifying key concepts, and gen- erating feedback. All results, including transcripts, behavioral metrics, and scores, are stored in a structured database for further analysis.

3) Layer 3: Multimodal Intelligence Engine

Layer 3 performs multimodal analysis by processing both audio and visual inputs. The audio stream is converted into a textual transcript using speech-to-text techniques, which is then used for content evaluation.

Simultaneously, video data is analyzed using computer vision methods to extract behavioral features such as eye contact and posture. Facial expressions are also evaluated to determine the dominant emotional state.

These outputs are generated independently and forwarded for unified evaluation. This modular approach allows each modality to be processed separately while contributing to a combined assessment of candidate performance.

This approach aligns with prior research in multimodal interview analysis, where combining speech, facial expressions, and behavioral cues improves evaluation quality [3].

4) Layer 4: LLM-Based Evaluation and Synthesis

Layer 4 acts as the final evaluation component, integrating outputs from the multimodal intelligence engine to generate a comprehensive assessment.

The primary input to this layer is the transcript of the candidate's response, along with the corresponding question and expected concepts. A large language model is used to evaluate the response in terms of relevance, completeness, and clarity. Leveraging prompt-based capabilities, the system performs flexible evaluation without requiring task-specific training [6].

In addition to textual analysis, behavioral metrics such as eye contact and emotional state are incorporated to provide a more holistic assessment. This allows the system to evaluate both the content and delivery of responses.

The model produces structured outputs, including a holistic score, feedback, strengths, and areas for improvement. It also evaluates concept coverage by comparing expected and identified concepts.

Overall, this layer synthesizes multimodal inputs into a unified and interpretable evaluation, enabling meaningful and action- able feedback for candidates.

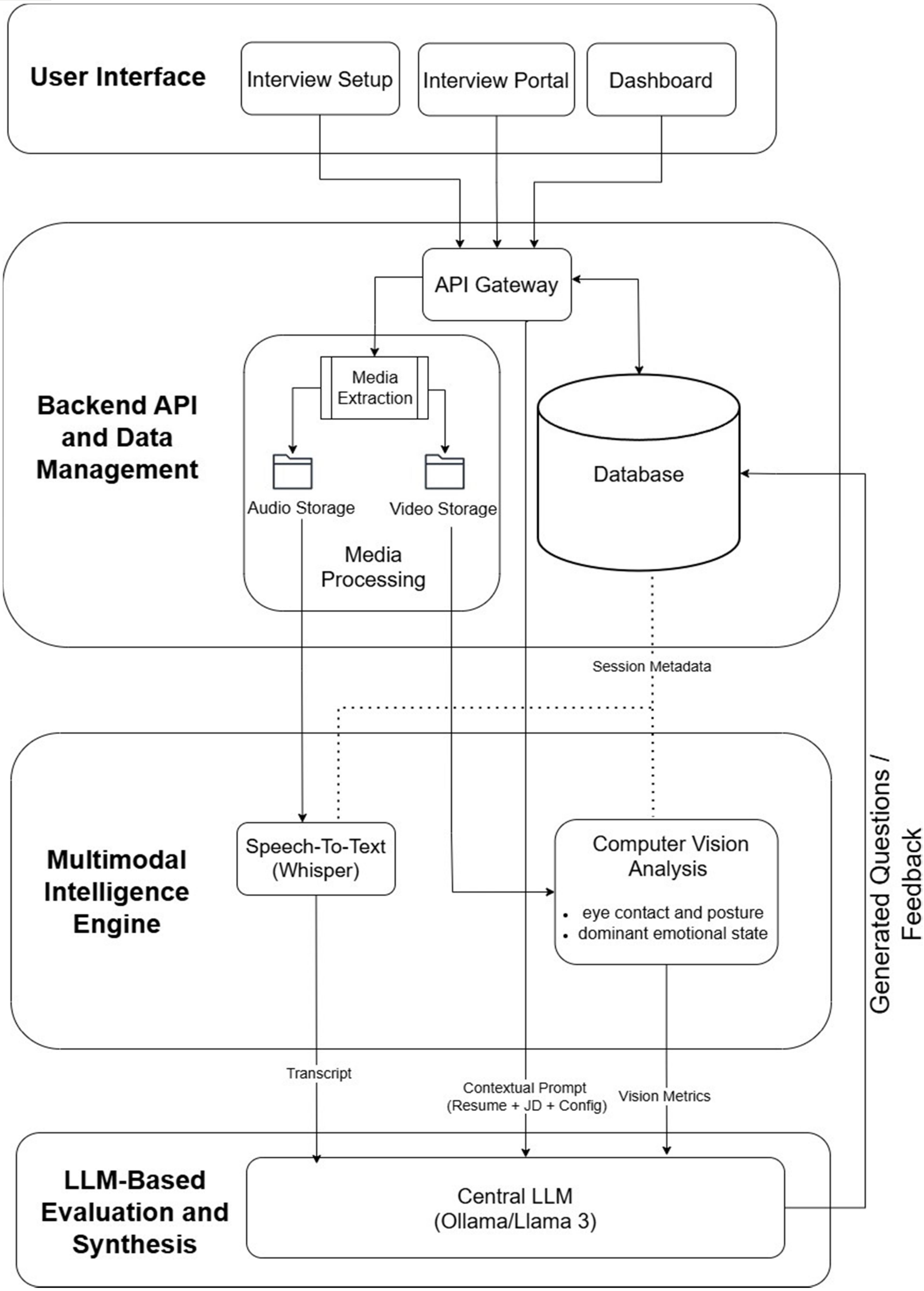


Figure 1: System Architecture of the *Cognify AI*, illustrating the multimodal data flow.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

IV. RESULTS AND EVALUATION

This section presents the evaluation of Cognify AI based on repeated mock interview sessions conducted with a single candidate. The objective was to analyze performance progression over time and assess how multimodal feedback contributes to improvement in both technical and behavioral aspects of interview performance.

A. Evaluation Setup

The system was evaluated over five successive mock interview sessions under controlled conditions. In each session, the candidate responded to a set of technical questions, and the system generated performance metrics using multimodal analysis.

The evaluation was based on the following metrics:

- Concept Detection Accuracy: Percentage of expected technical concepts correctly identified in responses.
- Eye Contact Score: Percentage of frames where stable facial presence and gaze alignment were detected.
- Dominant Emotion: Most frequently detected emotional state during each session.
- Holistic Performance Score: Overall score generated by the LLM-based evaluator.

B. Evaluation Methodology

Cognify AI evaluates candidate performance across two independent pipelines: a verbal analysis pipeline and a non-verbal analysis pipeline. These pipelines run in parallel and their outputs are aggregated into a final report rather than combined into a single numerical formula.

1) Non-Verbal Analysis

The non-verbal pipeline processes the recorded video using OpenCV and DeepFace. Two metrics are extracted:

Eye Contact Score Frames are sampled at one frame per second. For each sampled frame, a Haar Cascade frontal face detector determines whether the candidate's face is visible. The eye contact score is defined as:

$$E = \frac{F_{\text{face}}}{F_{\text{total}}} \times 100$$

where F_{face} is the number of sampled frames in which a face was detected and F_{total} is the total number of sampled frames. $E \in [0, 100]$ represents the percentage of time the candidate maintained camera-facing presence.

Dominant Emotion For each sampled frame, DeepFace performs facial emotion recognition and returns a predicted emotion label. The dominant emotion M is the most frequently occurring label across all sampled frames:

$$M = \arg\max_{e} \text{count}(e), \quad e \in \mathrm{E}$$

where E is the set of emotion categories returned by DeepFace. M is treated as a qualitative behavioral indicator and does not contribute to any numerical score.

2) Verbal Analysis

The verbal pipeline processes the audio track extracted from the recorded video. It consists of two stages: transcription and AI grading.

Transcription The audio is transcribed using OpenAI Whisper. Let T denote the resulting transcript. A filler word count Φ is also computed as:

$\Phi = \text{count}(\text{“}um\text{”}, T) + \text{count}(\text{“}uh\text{”}, T)$ Φ is recorded as a supplementary communication metric.

Concept Accuracy The language model identifies a set of n expected core concepts $\mathrm{K} = \{k_1, k_2, \ldots, k_n\}$ that a complete answer should contain, where $3 \leq n \leq 5$. It then determines the subset $\mathrm{H} \subseteq \mathrm{K}$ of concepts the candidate successfully addressed.The concept accuracy score is:

$$C = \frac{|H|}{|K|} \qquad (1)$$

where $C \in [0, 1]$. For reporting purposes, concept accuracy is expressed as a percentage, i.e., $C \times 100$, as shown in Tables 2 and 3.

Holistic Score The language model produces a holistic score $S \in [1, 10]$ based solely on the question text and the candidate's transcript T:

$$S = f(Q, T)$$

where Q is the interview question. S implicitly captures both technical depth (concept coverage) and communication quality (clarity, coherence, and relevance). The model is prompted with a fixed low temperature ($\tau = 0.2$) to ensure consistent and deterministic grading.

C. *Session-wise Performance Results*

Tables 2 and 3 present the results of two candidates across five successive mock interview sessions, demonstrating measurable improvement in both verbal and non-verbal metrics over repeated practice.

Table 2: Session-wise Performance Results of Candidate 1

Sess.	Concept (%)	Eye Contact (%)	Emotion	Score (%)
1	48%	60%	Neutral	46%
2	50%	75%	Neutral	51%
3	50%	80%	Happy	63%
4	58%	100%	Happy	70%
5	62%	100%	Happy	77%
Avg	54%	83%	–	61%

Candidate 1 began with a holistic score of 46% in Session 1, reflecting limited concept coverage (48%) and a low eye contact score (60%). The dominant emotion in early sessions was *Neutral*, indicating a degree of hesitation. By Session 3, a visible shift was observed as the dominant emotion transitioned to *Happy*. By Sessions 4 and 5, eye contact reached 100%, concept accuracy improved to 62%, and the holistic score peaked at 77%.

Table 3: Session-wise Performance Results of Candidate 2

Sess.	Concept (%)	Eye Contact (%)	Emotion	Score (%)
1	40%	55%	Neutral	42%
2	47%	65%	Neutral	49%
3	53%	75%	Happy	58%
4	60%	85%	Happy	65%
5	67%	95%	Happy	72%
Avg	53%	75%	–	57%

Candidate 2 started from a slightly weaker baseline, with a holistic score of 42% and a concept accuracy of 40% in Session 1. Eye contact improved from 55% to 95% and concept accuracy increased from 40% to 67% across the five sessions. The emotional state similarly transitioned from *Neutral* to *Happy* at Session 3.

D. *Performance Progression Analysis*

The results demonstrate a consistent improvement across successive sessions:

- Concept accuracy improved from 48% to 62% for Candidate 1, indicating better coverage of expected technical concepts. For Candidate 2, it increased from 40% to 67%.
- Eye contact increased from 60% to 100% for Candidate 1, showing improved visual engagement and behavioral stability. Candidate 2 saw an improvement from 55% to 95%.
- Holistic performance score improved from 46% to 77% for Candidate 1, reflecting overall enhancement in answer quality and delivery. Candidate 2's score improved from 42% to 72%.
- Dominant emotion progressed from *Neutral* in early sessions to *Happy* or *Confident* in later sessions, suggesting increased comfort and confidence during interviews.

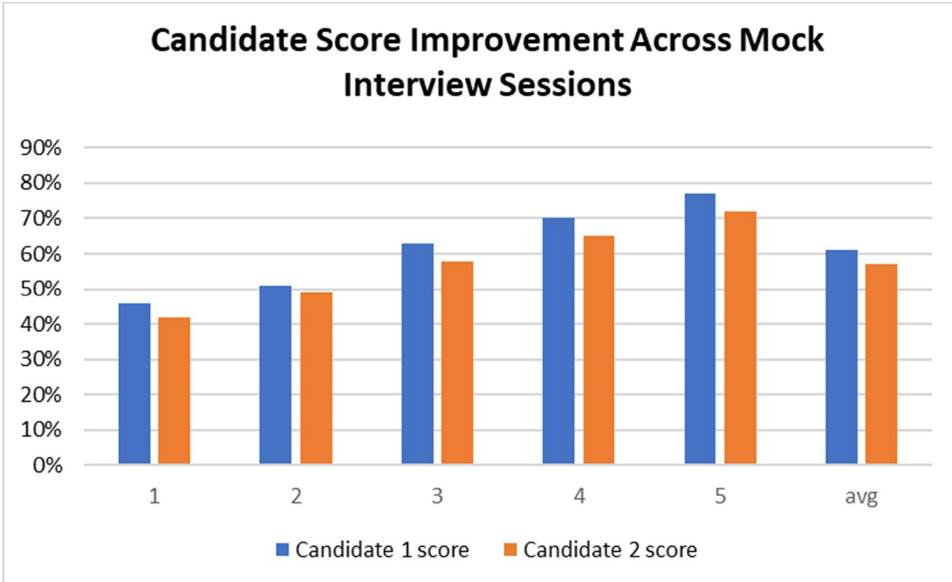


Figure 2: Candidate performance score comparison across mock interview sessions

E. Comparative Analysis

Traditional mock interview systems primarily focus on textual or speech-based evaluation, limiting their ability to assess non-verbal communication. Text-based systems evaluate only content correctness, while speech-based systems focus on delivery aspects such as fluency and tone.
In contrast, the proposed system integrates vision-based behavioral analysis with language model-based evaluation, enabling simultaneous assessment of both verbal and non-verbal performance.
Compared to text-only systems, Cognify AI provides additional behavioral insights such as eye contact and emotional state, which are critical factors in real-world interview performance. This results in a more comprehensive evaluation framework.
Furthermore, unlike standalone LLM-based systems that rely solely on transcript analysis, the proposed approach incorporates multimodal signals, allowing the system to capture both what the candidate says and how it is delivered.

Table 4: Qualitative Comparison with Existing Systems

Feature	Traditional Systems	Cognify AI
Concept Evaluation	✓	✓
Speech Analysis	✓	✓
Eye Contact Analysis	×	✓
Emotion Detection	×	✓
Multimodal Fusion	×	✓
Real-time Feedback	Limited	✓

This multimodal integration enables a more holistic and realistic assessment of interview performance compared to single- modality systems.

F. Discussion

The observed improvements indicate that Cognify AI effectively supports iterative interview preparation by combining multimodal feedback from speech, vision, and language-based evaluation. The system not only evaluates candidate responses but also enables measurable improvement across repeated sessions. While the evaluation was conducted on a limited controlled sample, the results demonstrate the practical potential of the pro- posed system in improving both technical and behavioral interview performance. Future evaluations involving multiple candidates and larger datasets can further validate the system's effectiveness.

G. Limitations

Despite its effectiveness, the proposed system has certain limitations that should be acknowledged.

First, the evaluation was conducted on a limited controlled dataset. While this setup helps demonstrate performance progres- sion, it may not fully generalize to diverse user populations with varying communication styles and backgrounds.

Second, the accuracy of visual analysis is dependent on environmental conditions such as lighting, camera quality, and face visibility. In scenarios with poor illumination or partial occlusion, eye contact estimation and emotion detection may be less reliable.

Third, emotion recognition is inherently challenging, as facial expressions do not always directly correspond to true emotional states. As a result, the detected dominant emotion should be interpreted as an approximate behavioral indicator rather than an exact measure.

Additionally, the holistic scoring mechanism relies on a large language model, which may exhibit variability in evaluation due to prompt sensitivity or inherent model limitations. While structured prompts are used to maintain consistency, slight variations in scoring may still occur.

Finally, the system currently evaluates responses at the session level and does not incorporate long-term learning or adaptive personalization based on historical performance trends. These limitations highlight opportunities for future improvements in scalability, robustness, and personalization.

V. CONCLUSION

This paper presented Cognify AI, a multimodal mock interview preparation system that integrates vision-based behavioral anal- ysis with large language model-driven content evaluation within a unified, real-time processing framework. The system was designed to address a fundamental gap in existing automated interview platforms, which predominantly assess verbal content while overlooking non-verbal behavioral cues such as eye contact and facial emotional state.

The proposed architecture processes audio and video streams through parallel, independent pipelines. Speech is transcribed using OpenAI Whisper and evaluated for concept coverage and holistic quality by a locally hosted Ollama/Llama 3 language model. Simultaneously, video frames are analyzed using OpenCV and DeepFace to extract eye contact scores and dominant emotional states. The outputs of both pipelines are synthesized into a structured, actionable feedback report.

Evaluation was conducted over five successive mock interview sessions with two candidates under controlled conditions. Candidate 1 demonstrated improvement in concept accuracy from 48% to 62%, eye contact from 60% to 100%, and holistic score from 46% to 77% across sessions. Candidate 2 similarly improved concept accuracy from 40% to 67%, eye contact from 55% to 95%, and holistic score from 42% to 72%. In both cases, the dominant emotional state transitioned from *Neutral* in early sessions to *Happy* by Session 3, suggesting increased composure through iterative practice.

These results indicate that multimodal feedback combining verbal and non-verbal signals supports measurable, progressive improvement across repeated sessions — a capability absent in single-modality systems. Qualitative comparison further confirms that Cognify AI extends beyond existing text-only and speech-only platforms by incorporating eye contact analysis, emotion detection, and multimodal fusion within a single evaluation pipeline.

The current evaluation is limited to a small controlled sample, and the accuracy of visual analysis remains sensitive to envi- ronmental conditions such as lighting and camera quality. Future work will focus on expanding the evaluation dataset to include diverse candidate profiles, incorporating long-term adaptive personalization based on historical session performance, improv- ing robustness of emotion recognition under varied recording conditions, and extending the system to support domain-specific interview tracks beyond technical roles.

Cognify AI demonstrates that integrating behavioral and semantic signals into a cohesive evaluation framework is both tech- nically feasible and practically effective for structured interview preparation.

VI. ACKNOWLEDGEMENT

The authors thank all participants who took part in the evaluation for their time and cooperation. This work was carried out at the Department of Computer Science and Engineering, REVA University, Bengaluru, India.

REFERENCES

[1] Caldera et al., “AI Interviewer Chatbot for Technical and HR Brilliance: A Tool for Upskilling Candidate,” International Journal of Engineering Research & Technology (IJERT), 2024.

[2] K. Senthilkumar et al., "AI Based Mock Interview System Using Natural Language Processing," in Proc. 2025 International Conference on Advanced Computing Technologies (ICoACT), 2025, pp. 01–06, doi: 10.1109/ICoACT63339.2025.11005032.

[3] I. Naim, M. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated Analysis and Prediction of Job Interview Performance,"

[4] IEEE Transactions on Affective Computing, vol. 9, no. 2, pp. 191–204, 2018, doi: 10.1109/TAFFC.2016.2614299.

[5] N. Rayasam et al., "Multimodal Sentiment Analysis for Interviews and Proctoring," in Proc. 2024 IEEE 9th Inter- national Conference on Computational Intelligence and Applications (ICCIA), 2024, pp. 115–119, doi: 10.1109/IC- CIA62557.2024.10719163.

[6] C. Kim et al., "Fairness-Aware Multimodal Learning in Automatic Video Interview Assessment," IEEE Access, vol. 11, pp. 122676–122688, 2023.

[7] R. Mandal, P. Lohar, D. Patil, A. Patil, and S. Wagh, "AI-Based Mock Interview Evaluator: An Emotion and Confidence Classifier Model," in Proc. 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), IEEE, 2023, pp. 521–526.

[8] T. Brown et al., "Language Models are Few-Shot Learners," Advances in Neural Information Processing Systems, 2020.

[9] S. Wangwiwattana and P. Tongvivat, "Automating Academic Assessment: A Large Language Model Approach," in Proc. 2023 International Conference on Information Technology and Electrical Engineering, IEEE, 2023.

[10] P. Chokpattarabun et al., "AI-Powered MetaHuman Interviewer: Serious Game for Student Job Interview Skills," IEEE Access, vol. 12, pp. 11275–11285, Dec. 2025.

[11] H. Sun, H. Lin, H. Yan, Y. Song, X. Gao, and R. Yan, "MockLLM: A Multi-Agent Behavior Collaboration Framework for Online Job Seeking and Recruiting," in Proc. 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2025.

[12] Y. Dai et al., "LLM-generated Feedback in Real Classes and Beyond: Perspectives from Students and Instructors," in Proc. 17th International Conference on Educational Data Mining (EDM), 2024.

[13] Y.-C. Chou, F. R. Wongso, C.-Y. Chao, and H.-Y. Yu, "An AI Mock-interview Platform for Interview Performance Analysis," in Proc. 10th Int. Conf. Information and Education Technology (ICIET), 2022, pp. 37–41.

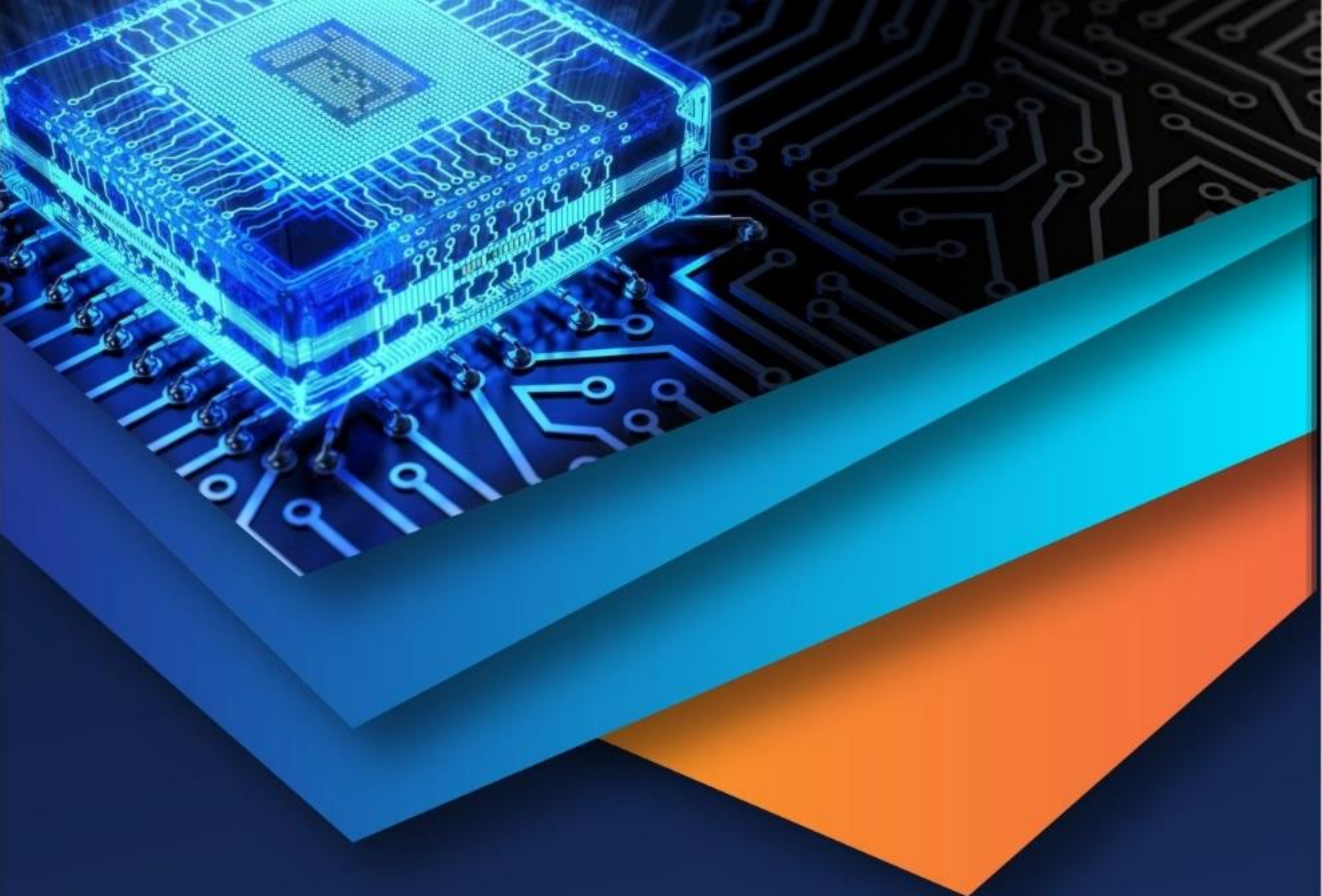
[14] B. J. K. Reddy, B. S. C. Reddy, H. S. Shekhawat, and Indumathy M., "AIVA: AI-powered Interview Verbal Analysis System using Fine Tuned Models," in Proc. International Conference on Intelligent Systems and Digital Transformation (ICISD), Atlantis Highlights in Intelligent Systems, 2025.

[15] S. Yadav et al., "AI Voice Interview Agent for Real-Time Personalized Mock Interviews," IJSAT, 2025.

[16] J. M. Basch and K. G. Melchers, "Here's Looking at You: Does Eye Contact in Video Interviews Affect How Applicants are Perceived and Evaluated?" Journal of Business and Psychology, 2025.

[17] R. Loga Priya, S. R. K. Sri Roshan, G. S. Vidharsana, and P. N. Saravanan, "AI-Enhanced Eye Tracking for Candidate Assessment in Job Interviews," in Proc. 2025 6th Int. Conf. on Mobile Computing and Sustainable Informatics (ICMCSI), IEEE, 2025.

[18] S. R. Jagtap, V. Kulkarni, Y. Pachorkar, O. Taur, S. Gupta, and U. Pujeri, "AI-Driven Real-Time Interview Simulation App with Voice Recognition and Facial Analysis," Indian Journal of Science and Technology, vol. 18, no. 25, pp. 2058–2066, 2025.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY