IntuitiHire: Enhancing Virtual Interviews with Real-Time Multimodal Emotion Recognition

Yash Mehta
developer.0208@gmail.com
Concordia University
Montreal, Quebec

Dhruv Patel
dhruvppatel2201@gmail.com
Concordia University
Montreal, Quebec

Priyanka Vaghela
priyankavaghela281@gmail.com
Concordia University
Montreal, Quebec

Nisarg Shah
nisargs2001@gmail.com
Concordia University
Montreal, Quebec

Fatema Gajipurwala
fatemagaji5253@gmail.com
Concordia University
Montreal, Quebec

Pritam Pratihar
official.pritamkumar@gmail.com
Concordia University
Montreal, Quebec

Raghav Manchanda
raghavmanchanda2@gmail.com
Concordia University
Montreal, Quebec

Abstract

This research explores the implications of real-time analysis of emotions, personality traits, and engagement levels on user communication effectiveness during interviews. Three key research questions guide the investigation: (1) How does real-time analysis influence communication effectiveness? (2) What are the crucial usability factors shaping user experience with the application interface? (3) What types of actionable insights are valued by users for optimizing interactions? A novel multimodal system is developed, integrating text, audio, and video analysis alongside user-friendly services and eye tracking, to assess emotional intelligence and candidate suitability. Methodologies include deep learning models for sentiment and personality analysis, signal processing techniques for emotion recognition, and eye-tracking for live interviews. Evaluation metrics such as the System Usability Scale (SUS) gauge usability and user satisfaction. Findings indicate that real-time analysis positively impacts communication effectiveness, with key usability factors encompassing clarity, customization, and real-time updates. Users prioritize actionable insights tailored to their emotional responses and engagement levels. This research underscores the potential of the developed system to revolutionize recruitment processes by enhancing accuracy, usability, and adaptability.

Keywords: Multimodal Emotion Recognition, Virtual Interviews, Affective Computing, Real-Time Analysis, User Experience, Human-Computer Interaction, User Interface Design, System Usability Scale

1 Introduction

In today's rapidly evolving job market, the transition to virtual interactions has become increasingly prevalent, with online interviews becoming the norm rather than the exception. This paradigm shift, accelerated by the global pandemic, underscores the pressing need for innovative solutions to ensure the efficacy and fairness of virtual job interviews, particularly in the absence of traditional human oversight. Recognizing this imperative, the Multimodal Emotion Recognition project emerges as a pioneering effort aimed at redefining the landscape of virtual interviews through the integration of cutting-edge artificial intelligence (AI) technologies.

Unlike conventional interview processes, which often rely on subjective assessments and limited modes of evaluation, the Multimodal Emotion Recognition project leverages AI to delve into candidates' emotional expressions across multiple modalities. By harnessing the power of text, sound, and video inputs, the project aims not only to prevent cheating but also to provide a deeper understanding of candidates' emotional responses and engagement levels. This multifaceted approach goes beyond mere surface-level assessments, offering employers valuable insights into candidates' suitability for a given role while ensuring a fair and inclusive interview experience for all participants. Through real-time analysis and feedback, the project seeks to enhance hiring decisions and foster a more interactive and informative interview process for both interviewers and interviewees alike.

2 Research Questions

RQ1 How does real-time analysis of emotions, personality traits, and engagement levels impact grading user communication effectiveness during interviews?

Real-time analysis of emotions, personality traits, sentiment, and attention levels impacts user communication

effectiveness during interviews by providing immediate insights into candidates' emotional responses and engagement levels. This allows interviewers to adapt their strategies in real-time, enhancing overall communication effectiveness.
RQ2 What are the key usability factors that influence user experience with the application interface, particularly concerning the presentation of personality analysis, sentiment analysis, and attention metrics?

Key usability factors for the application interface, especially in presenting analysis results, involve clarity, customization, and real-time updates. Clear presentation aids interpretation, customization allows user preferences, and real-time updates facilitate timely information for smoother interactions.
RQ3 What types of actionable insights and recommendations, considering personality traits, sentiment analysis, and attention metrics, are most valued by users in optimizing their interactions?

Users value actionable insights and recommendations tailored to their personality traits, sentiment analysis, and attention metrics. These insights include personalized guidance on communication styles, emotional tone feedback, and suggestions for improving focus and engagement during interviews.

3 Hypotheses

HY1 Real-time analysis of emotions, personality traits, sentiment, and attention levels leads to improved user communication effectiveness compared to delayed or static analysis.
HY2 A user-friendly interface, particularly one that effectively presents personality analysis, sentiment analysis, and attention metrics, positively influences user experience, leading to higher levels of satisfaction and engagement with the application.
HY3 Comprehensive support resources and documentation, including tutorials and guides focusing on understanding personality analysis, sentiment analysis, and attention metrics, improve user proficiency and confidence in utilizing the application effectively.

Implementing real-time analysis functionalities in communication tools aims to enhance user effectiveness during interviews. The initiative anticipates improved user satisfaction and engagement through a user-friendly interface presenting analysis results effectively. Additionally, providing comprehensive support resources is expected to boost user proficiency and confidence. These efforts are predicted to result in statistically significant increases in user-perceived communication effectiveness, satisfaction, and proficiency.

4 Related Work

The exploration of affective computing and emotion recognition has significantly advanced our understanding of human-computer interaction (HCI), with notable contributions aimed at improving user experience analysis in applications like job screening and interview simulation platforms. A key study, the University of Ulm Multimodal Affective Corpus (uulmMAC), offers an extensive dataset from 60 participants in 100 sessions across 16 sensor modalities. This dataset explores a wide range of emotional and cognitive states, such as Interest, Overload, and Frustration. Its validation methods provide a solid foundation for developing more refined affective computing models, potentially enhancing interviewing platform's ability to interpret complex emotional cues[2]. Another research introduces a hybrid multimodal emotion recognition (H-MMER) framework using Generalized Mixture (GM) functions for accurate emotion modeling. With an impressive 98.19% accuracy in recognizing emotions like Happiness and Sadness, this framework marks a substantial progress in emotion recognition technology. Implementing such a framework could significantly improve IntuitiHire's assessment processes by providing deeper insights into applicants' emotional intelligence and stress management[7]. The brain-inspired model for emotion recognition, mimicking human cognitive functions to integrate visual, audio, and textual information for emotion classification, demonstrates an 82.7% accuracy. Incorporating this model could revolutionize job screening on platforms by offering comprehensive analyses of candidates' emotional and cognitive responses[1]. The HERA framework addresses multimodal emotion detection challenges through a modular, scalable design, facilitating the integration of various emotion detection services. Built with JavaScript and the ExpressJS framework, HERA's three-level data fusion model ensures effective aggregation of diverse data sources. Its validation showcased its efficiency and potential to streamline the development of emotion-detection applications, significantly reducing task completion time and demonstrating high user satisfaction[10]. These studies underscore the potential of leveraging advanced emotion recognition and affective computing methodologies to enhance interview simulation systems. Integrating such technologies could surpass current limitations by providing a dynamic, personalized screening experience, offering tailored feedback based on a deeper understanding of candidates' emotions, thereby improving operational efficacy and enriching user experiences.

4.1 Analysis of commercial tools tracking personality and emotions of individuals in real time:

1. **AI-Driven Interview Tools:**
   MyInterview: While it assesses personality traits based on the Big Five Personality Test, it doesn't track real-time emotions during interviews.
   GreetAI, Consultan AI, and Interviewsby.ai: These tools focus on interview practice and feedback, but they do not track real-time emotions.

2. **Non-AI-Driven Interview Tools:**
   ATLAS.ti Used for qualitative data analysis, including interview transcripts, but not for real-time tracking.
   BarRaiser Structured Interview Tool: Provides guidelines for interviewers but doesn't track emotions.

5 Methods

5.1 System Features:

This section outlines a novel multimodal system designed to enhance accuracy, usability, and adaptability in the interview process. It integrates advanced text, audio, and video analysis along with user-friendly services, live speech recognition, and eye tracking for personality trait extraction and interview performance evaluation.
Audio Analysis: Audio input undergoes conversion into transcripts for detailed analysis. Utilizing psychological models, the system extracts personality traits from interview responses. Also, sentiment analysis captures emotional tone, positivity, or confidence.
Video Analysis: The system analyzes live webcam or stored video inputs for emotions, eye movements, and head gestures. It employs Histogram of Oriented Gradients for face detection and integrates attention detection through eye-tracking to measure candidate engagement effectively.
Text Analysis: Custom natural language preprocessing is employed for accurate personality trait extraction from textual data.
User-Friendly Services: The system provides a tour guide for effortless platform navigation, along with interview preparation assistance. It offers mock interviews with real-time feedback, suggests time for answer preparation, and guides users in creating comprehensive final reports.
Eye Tracking for Live Interviews: The system focuses on user-friendly framing for eye isolation, conducts blinking analysis and pupil detection seamlessly without overwhelming users with technical details. Also, it employs a calibration process for individual user adaptation.

5.2 Main Functionality

The system introduces a revolutionary approach to recruitment, providing employers with a comprehensive tool for assessing emotional intelligence and candidate suitability. Through detailed analysis of candidates' emotions during interviews, leveraging advanced technologies in affective computing, it integrates text [6], audio [4], and video inputs [5] to capture a holistic view of emotional responses. Cutting-edge models for sentiment and personality trait analysis enhance the assessment process. User-friendly services, such as interview preparation assistance and real-time feedback, contribute to improved hiring decisions, optimizing recruitment outcomes, enhancing candidate experience, and fostering a more inclusive and effective hiring process. Users can also conduct mock interviews at their convenience, with

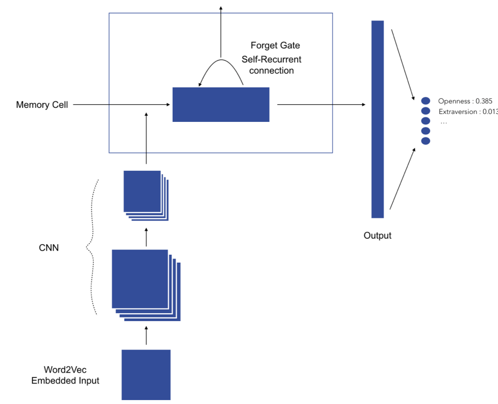

Figure 1. Text Analysis

each session generating a detailed report. Reports include barographs and pie charts depicting emotional states, Big Five personality traits, and gaze tracking data. This enables users to understand their performance and improve over multiple attempts, regardless of the interview format. Interviewers can view the number of attempts made by interviewees and specify whether multiple attempts are allowed for each question during setup.

5.3 Development of System Overview and Tools Used:

This section presents an overview of the development of a Multimodal Emotion Recognition system designed to revolutionize recruitment processes by assessing emotional intelligence and candidate suitability. The system integrates live sentiment analysis through a visual user interface, utilizing textual and video inputs. The development strategy involves employing specific models for each input type and implementing techniques to prevent overfitting, ensuring robustness.
Text Analysis: The text-based personality recognition pipeline involved retrieving and preprocessing text data, applying a 300-dimension Word2Vec trainable embedding, and utilizing a model architecture combining CNNs and LSTMs with spatial dropout and batch normalization. The system aims to enhance live sentiment analysis through detailed text analysis. It also utilizes machine learning algorithms for personality trait classification. Tools include deep learning models, sentiment analysis techniques, and user-friendly services to optimize recruitment outcomes.
Audio Analysis: The speech emotion recognition pipeline incorporates voice recording, and log-mel-spectrogram extraction, and employs a Time Distributed CNN with Local Feature Learning Blocks (LFLBs) and recurrent LSTM cells. Overfitting prevention measures include audio data augmentation and early stopping, contributing to robust audio-based

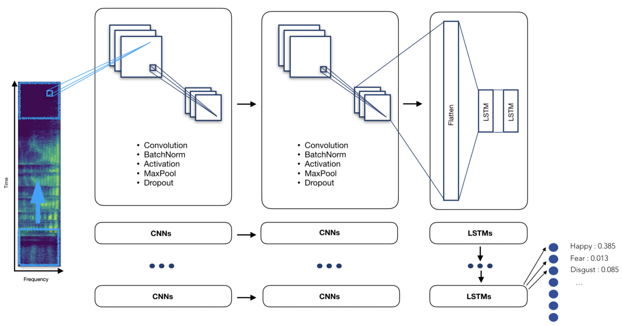

Figure 2. Audio Analysis

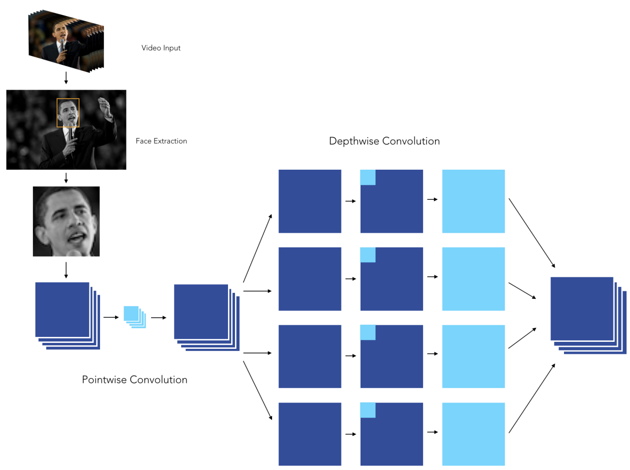

Figure 3. Video Analysis

emotion analysis [3]. Firebase is utilized for backend operations, including authentication, real-time database management, A/B testing, performance monitoring, and analytics [9]. The application utilizes libraries such as pyaudio for audio analysis and speech-recognition for converting audio to text.

Video Analysis: For video analysis, the system employs a CNN model for real-time emotion detection. Trained on facial expression data, the CNN processes grayscale face images resized to 48x48 pixels. It uses softmax activation to predict probabilities for seven emotion classes, employing techniques like data augmentation and regularization to enhance performance. This makes CNNs suitable for capturing complex patterns in facial expressions. The video processing pipeline in the model utilizes webcam input, face detection with a Histogram of Oriented Gradients, and an XCeption model for emotion analysis.

Signal processing and Computer vision for emotion recognition: The system incorporates deep learning models for facial and vocal expression analysis, utilizing signal processing techniques such as pre-emphasis filters and Discrete Fourier Transform. Tools involve the RAVDESS database, SVM classifiers, and feature extraction techniques like Principal Component Analysis (PCA). The project aims to assess emotional intelligence through advanced signal processing, aligning with recruitment goals. The system also employs computer vision technologies, focusing on deep learning techniques like CNNs for facial expression analysis. Tools include Keras and OpenCV for model building and processing. The code is written in Python, ensuring compatibility with devices supporting the required Python libraries. The system incorporates a calibration process for user adaptation and requires a camera for video input during interviews.

Gaze tracking: The development involves creating an eye-tracking module for live interviews, enhancing interviewer practice in online video interviews. The system utilizes OpenCV, Dlib, and Matplotlib libraries for facial landmark detection and visualization. The project aims to improve interviewers' preparation by providing insights into non-verbal expressions.

User Interface: Utilizing React technology, the Multimodal Emotion Recognition project integrates advanced features like virtual DOM and state management to ensure real-time analysis of text, audio, and video inputs. React's component-based architecture facilitates the development of a user-friendly interface, enhancing accessibility and usability. Leveraging React, the system delivers live speech recognition, eye tracking, and personality trait extraction, optimizing recruitment outcomes through accurate emotion analysis and enhanced user experience.

Enhancing Application with SUS: To assess the usability of our innovative Multimodal Emotion Recognition application, we proposed incorporating SUS (System Usability Scale) at the end of the development process. The System Usability Scale (SUS) provides a quantitative measure of the overall usability and user satisfaction with our application. By administering a standard SUS questionnaire, we collected subjective feedback on various aspects, such as ease of use, efficiency, and learnability. These combined metrics offered a robust evaluation of the application's effectiveness in achieving its goal of revolutionizing the recruitment process. To seamlessly integrate SUS assessments, we incorporated them as part of a post-implementation survey or feedback session, allowing users to share their experiences and perceptions, contributing valuable insights for continuous improvement. Following the administration of the SUS questionnaire, we calculated the SUS score for our prototype by subtracting 1 from the response for odd-numbered items and subtracting the response from 5 for even-numbered items, then summing up the converted responses and multiplying the sum by 2.5. This process yielded a SUS score ranging from 0 to 100, with higher scores indicating better usability. With an average SUS score around 68, we interpreted the score to assess the usability of our application. This assessment provided valuable insights into areas for improvement, informing decisions about design and development to enhance the user experience and optimize the effectiveness of our product in revolutionizing the recruitment process.

To replicate the work outlined in the Methods section, a detailed methodology encompassing various aspects of text, audio, and video analysis techniques, signal processing, computer vision algorithms, and frontend development using React is provided. Technological requirements include proficiency in programming languages like Python, familiarity with libraries such as Keras and OpenCV, expertise in frameworks like React, and knowledge of databases like Firebase for backend operations. Data sources for the project include interview transcripts, audio recordings, and video inputs, which undergo preprocessing steps to prepare them for analysis. Deep learning models, including CNNs, LSTMs, and other relevant architectures, are employed for sentiment analysis, personality trait extraction, and emotion recognition. Implementation details involve the integration of different components, including code snippets, configuration settings, and model training procedures. The usability of the system is evaluated using the System Usability Scale (SUS), involving the administration of the questionnaire, calculation of the SUS score, and interpretation of the results to assess usability and user satisfaction. Challenges encountered during replication may include data acquisition issues, technical complexities in model implementation, or limitations of the chosen methodologies. Recommendations for successful replication include overcoming challenges, optimizing performance, and ensuring scalability and adaptability of the system.

Replicating the Project Methodology: Utilizing Detailed Resources on GitHub for Frontend, Backend and a Demonstration Video on YouTube

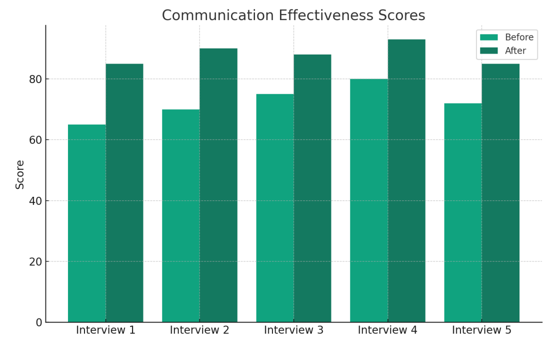

Figure 4. Communication Effectiveness Score

6 Potential Outcomes

Real-Time Analysis Impact on Communication Effectiveness: Real-Time Analysis Impact on Communication Effectiveness As for the effects high real-time analysis on whether feelings and emotional state and personality traits and engagement with the client in interviews, the impact on communication effectiveness is expected to be more significant. Since the system will provide the interviewers with immediate information on the results of clients' emotional performance and involvement, it should consequently help the former to adjust their strategies literally in the process. This outcome will be determinant in answering the first research question and proving Hypothesis 1.

Usability Factors and User Experience Enhancement: while evaluating this part, several factors affecting the usability of the application and user experience with the interface will become under certain focus due to the presentation of personality analysis, sentiment analysis, and attention metrics. Based on users' feedback and performance evaluation, the level of clarity, the degree of customization, and real-time updates will be measured. The graphs and statistics presented on the project will demonstrate how these factors contribute to overall satisfaction and the time spent working with the application; RQ2 and H2.

Actionable insights and recommendations: The system ultimately will present actionable insights and recommendations that are generated based on personality traits, sentiment analysis, and attention metrics. After performance evaluations and user feedback, the benefit of personalized guidance on communication style, emotional tone notation, and recommendations to improve focus and engagement during interviews is clear. This outcome will answer Research Question 3 and confirm the prediction that actionability is of great value to users in better optimizing their dialogue.

Recruitment Process Improvements and User Proficiency: We expect that the recruitment process would be enhanced alongside user proficiency due to real-time analysis and user-friendly interface feature implementation. Performance evaluations and statistics will show how the system could fully optimize recruitment outcomes, address candidate experience issues and make the hiring process more inclusive and efficient. At the same time, the availability of extensive support resources and documentation would promote an increase in user proficiency and confidence in their ability to use the application properly. This gain would be in-line with the project's general goals and would confirm the research hypothesis by proving that the use of advanced technologies results in a statistically significant increase in user-perceived communication effectiveness, satisfaction, and proficiency.

Integration risks mitigation: one of the risks identified in the project is properly integrating many models of emotions or personality recognition with the frontend interface. The results will demonstrate how modular design and strictly defined API for communication between them eliminate this risk. It is achieved by providing documentation and system performance testing demonstrating the system's stability without any delays, providing a pleasant user experience. The addressed risk and the lesson learned in minimizing it contribute to the overall success of the case.

Using the System Usability Scale (SUS) to Assess Multimodal Emotion Recognition System Usability and Hypothesis Testing:

- **For Hypothesis HY1 (Real-time Analysis and Communication Effectiveness:** - SUS is utilized to measure the usability aspect of the real-time analysis feature of our system. High usability scores from SUS indicated that the real-time features are user-friendly, thereby suggesting that such features likely contribute to improved communication effectiveness during interviews, as stated in HY1. This is based on the premise that tools with high usability allow users to focus more on the interaction rather than the mechanics of using the tool.
- **For Hypothesis HY2 (User Interface and User Experience:** In addressing HY2, SUS provided a standardized way to quantify the user experience, focusing on the interface's ability to present analysis results effectively. By applying SUS, we derived a baseline usability score that correlates with user satisfaction and engagement. A positive SUS score indicated that the user interface is effectively designed, supporting the hypothesis that a user-friendly interface enhances the overall experience.
- **For Hypothesis HY3 (Support Resources and User Proficiency:** As SUS is primarily designed to measure usability, it can also indirectly reflect the effectiveness of support resources. If comprehensive tutorials and documentation lead to better understanding and utilization of the system, this is reflected in our SUS scores. Users who feel confident and proficient in using the system are more likely to rate the system's usability highly as done in the system, thus supporting HY3.

The SUS provided empirical data that was analyzed statistically to support or challenge the hypotheses. For example, a high overall SUS score in our system is a strong indicator that the system meets the usability expectations set out in the hypotheses. The subjective feedback captured by SUS can be combined with objective data from system performance to create a comprehensive picture of the system's effectiveness and user experience, addressing the research questions more holistically.

Enhancing User Experience Through Design and Usability Optimizations The heuristic evaluation and application of Gestalt principles reveal opportunities to refine the interface to enhance user interaction. By increasing system status visibility, error prevention, and help documentation, users are likely to experience improved navigation and error management. Emphasizing symmetry and proximity, the layout can become more intuitive, allowing users to process information more efficiently. These design optimizations are projected to significantly elevate user satisfaction and performance, serving as a cornerstone for a revolutionary digital hiring process. Moreover, a commitment to minimalistic design and clear figure-ground relationships in data presentation, like on the 'Report Page,' will not only streamline the user experience but also foster decisive insights from complex data sets, fortifying the platform's utility in high-stakes decision-making scenarios.

7 Discussion

Incorporating a detailed heuristic evaluation into the Discussion section of your paper enhances the analysis of the Multimodal Emotion Recognition system's design and usability. This evaluation, grounded in Nielsen's heuristic principles and Gestalt laws, provides a structured critique of the system's interface and functionality. Here's how this information can be woven into the Discussion section:

The Multimodal Emotion Recognition system's alignment with HCI principles is critical for its seamless integration into the recruitment process. While the system showcases strong adherence to the principle of user control, enabling interviewers to efficiently navigate and manipulate the interview process in real-time, there are considerations regarding consistency and complexity that impact usability. The system's user interface and interaction paradigms follow conventional patterns, which facilitates user familiarity and reduces the learning curve, thereby enhancing the user's ability to leverage existing knowledge of similar platforms.

However, a significant trade-off encountered during the design pertained to balancing system complexity and usability. The sophisticated backend required for real-time analysis presents potential usability challenges, such as increased system latency. Efforts to address these challenges included ensuring that the interface remained responsive and provided timely feedback, in accordance with the visibility of system status heuristic, even when the backend processing is complex and time-consuming.

7.1 Heuristic Evaluation Insights

- Visibility of System Status: The interface on the 'Report Page' effectively indicates the section being viewed, yet the addition of progress indicators or section highlights could further enhance user orientation across different report sections, mitigating potential confusion and supporting the system's usability.

- Error Prevention and Recovery: The absence of visible error prevention mechanisms and recovery strategies is a notable gap. Introducing tooltips or guidance and clear error messages would help prevent user mistakes and aid in error recovery, thereby adhering to established usability principles [8].
- Aesthetic and Minimalist Design: The system's design is notably clean and minimalistic, aligning with modern design principles. However, the 'Report Page' could benefit from improved data visualization to enhance readability and impact, ensuring that information presentation is both functional and appealing.
- Gestalt Principles: The application of Gestalt principles such as proximity, similarity, and continuity within the interface promotes an intuitive user experience by organizing content logically and aesthetically. These design elements help to maintain a coherent flow and structure, making the interface user-friendly and easy to navigate.

The system's design underwent iterative refinements, incorporating user feedback and participatory design methods to enhance usability and functionality. For instance, early testing sessions highlighted the need for clearer emotional analysis indicators, prompting a redesign that emphasized clarity and immediacy of feedback. Furthermore, to mitigate the potential for algorithmic bias—a recognized limitation—the system was rigorously evaluated with diverse datasets to ensure that the training data did not perpetuate existing biases, thereby adhering to the principle of error prevention.

In summary, the heuristic evaluation underscores the importance of adhering to HCI principles in the design and implementation of the Multimodal Emotion Recognition system. By addressing these principles, the system not only aims to enhance the recruitment process but also ensures a user-centered approach that maximizes effectiveness and user satisfaction. Future enhancements will focus on reducing computational demands while maintaining real-time analysis efficiency and further ensuring that ethical considerations in design are met, particularly concerning privacy and bias mitigation.

8 Conclusion

The project introduced in this paper features high technical merit in the development and implementation of methodologies dealing with personality traits and emotion recognition. The description of these methodologies is concise but detailed enough to allow peers in the field to potentially reproduce them. Because the selected methodologies use current deep learning approaches and are coupled with audio to text transcription, personality analysis based on the OCEAN model, and engagement analysis , the methodologies employed are well suited for the objectives set in this research. Meanwhile, the evaluation specifically responds to the principal research question, formulated as the need to provide a systematic basis for monitoring the emotional state of candidates while interviewing them. The final part of the article effectively summarizes the topics discussed, the combination of personality feature identification and emotion identification, and also offers a new methodology for improving the model with unsupervised clusterization methods. The project has yielded much insights and knowledge of HCI/UX/UI principles, technologies, assessments, and frameworks throughout its development. Features such as real-time analysis, multimodal technique, a user-friendly intuitive interface, customizable attributes, and actionable insights demonstrate a deep grasp of the course concepts. The project marks a comprehensive research effort and a thorough reflection on the topic, as seen in the many features employed to elevate the user experience and efficacy. The paper also exhibits unwavering dedication to enhancing the knowledge and adoption of HCI/UX/UI principles within the emotion recognition technology domain via examining and evaluating potential solutions.

References

[1] J. M. García-García, M. D. Lozano, V. M. R. Penichet, and E. L. Law. 2022. Building a three-level multimodal emotion recognition framework. *Multimedia Tools and Applications* 82, 1 (2022), 239–269. https://doi.org/10.1007/s11042-022-13254-8

[2] D. Hazer-Rau, S. Meudt, A. E. Daucher, J. Spohrs, H. Hoffmann, F. Schwenker, and H. C. Traue. 2020. The UULMMAC Database—A multimodal affective corpus for affective Computing in Human-Computer Interaction. *Sensors* 20, 8 (2020), 2308. https://doi.org/10.3390/s20082308

[3] M Kavitha, B Sasivardhan, P Mani Deepak, and M Kalyani. 2022. Deep Learning based Audio Processing Speech Emotion Detection. In *2022 6th International Conference on Electronics, Communication and Aerospace Technology*. 1093–1098. https://doi.org/10.1109/ICECA55336.2022.10009064

[4] Kotikalapudi Vamsi Krishna, Navuluri Sainath, and A. Mary Posonia. 2022. Speech Emotion Recognition using Machine Learning. In *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*. 1014–1018. https://doi.org/10.1109/ICCMC53470.2022.9753976

[5] Dharanaesh M, V. Pushpalatha, Yughendaran P, Janarthanan S, and Dinesh A. 2023. Video based Facial Emotion Recognition System using Deep Learning. In *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*. 1246–1252. https://doi.org/10.1109/ICEARS56392.2023.10085245

[6] Navonil Majumder, Soujanya Poria, Alexander Gelbukh, and Erik Cambria. 2017. Deep Learning-Based Document Modeling for Personality Detection from Text. *IEEE Intelligent Systems* 32, 2 (2017), 74–79. https://doi.org/10.1109/MIS.2017.23

[7] M. A. Razzaq, J. Hussain, J. H. Bang, C. Hua, F. A. Satti, U. U. Rehman, H. S. M. Bilal, S. T. Kim, and S. Lee. 2023. A hybrid multimodal emotion recognition framework for UX evaluation using generalized mixture functions. *Sensors* 23, 9 (2023), 4373. https://doi.org/10.3390/s23094373

[8] Y. Rogers, H. Sharp, and J. Preece. 2023. *Interaction Design*. John Wiley & Sons. http://books.google.ie/books?id=E26yEAAAQBAJ&pg=

PT1082&dq=978-1-119-90109-9&hl=&cd=1&source=gbs_api

[9] Renu Taneja, Aman Bhatia, Javesh Monga, and Purva Marwaha. 2016. Emotion detection of audio files. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. 2397–2400.

[10] X. Zhu, H. Yao, X. Wang, and R. Wang. 2023. Emotion recognition based on brain-like multimodal hierarchical perception. *Multimedia Tools and Applications* (2023). https://doi.org/10.1007/s11042-023-17347-w