Journal of Global Research in Electronics and Communication
Volume 1, No. 9, September 2025
Available Online at: www.jgrec.info
RESEARCH PAPER

AI-Powered Presentation Assistants: A Review of Tools for Real-Time Feedback and Content Optimization

Dr. Bal Krishna Sharma
Professor
Department of Computer Sciences and Applications
Mandsaur University
Mandsaur
bksharma7426@gmail.com

***Abstract*—Artificial Intelligence (AI)-powered presentation assistants are revolutionizing how users create and deliver presentations by offering real-time, personalized feedback on speech, gestures, and content organization. This paper reviews and classifies these tools into five key categories: speech analysis tools, facial expression and gesture recognition systems, content structure and visual optimization platforms, real-time audience engagement analytics, and automatic summarization modules. The study examines core AI technologies such as CTENet for emotion detection, transformer-based models for semantic coherence, and explainable AI (XAI) methods to enhance feedback interpretability. Key evaluation criteria include real-time processing, personalization, voice integration, privacy, and accessibility. Applications are explored across education, corporate training, and public speaking, demonstrating the tools' ability to provide unbiased, scalable, and time-efficient performance evaluations. Despite their strengths, challenges remain, including limited adaptability, insufficient qualitative feedback, and context-blind emotion recognition. The review makes reference to future objectives of increased inclusivity, user experience, and pedagogical effectiveness. In general, presentation assistance AI can be a beneficial tool in the workplace and in general education, as it demonstrates substantial potential to improve communication processes and streamline content delivery.**

Keywords—AI-Powered Presentation Tools, Remote Software Collaboration, Natural Language Processing (NLP), Educational Technology, Real-Time Feedback.

I. INTRODUCTION

Presentations are now a ubiquitous communication and knowledge dissemination tool, used in a wide variety of contexts in education, business, law, healthcare, and engineering. The use of technology-based presentation especially PowerPoint (PPT) in teaching and learning has gradually superseded the use of chalkboard, whiteboard, or overhead projectors in teaching and learning, especially in the tertiary level of education [1]. Using LCD projectors and laptops, such slide-based techniques are now the norm within classrooms, and they frequently supersede the traditional chalk and talk approach. As more textbook-based slides are made available and with institutional support, PPT integration has now been made more available and time-saving to teachers [2]. However, despite their popularity, concerns have been raised about their actual performance in the student engagement and performance domains.

Artificial intelligence (AI) tools are software and systems designed to mimic human intellect. These two aspects of artificial intelligence (AI) are significantly impacting the education sphere, specifically in the optimization of presentations and the implementation of student feedback. The emergence of AI-enhanced presentation assistants has enabled instructors and students to benefit from the advantages of real-time feedback, customized assistance, and content analysis. These types of approaches can be implemented through natural language processing (NLP), machine learning and sentiment analysis to dynamically react to the needs of learners and offer cognitive and emotional feedback to tailor the learning experience.

Sentiment analysis is a highly specific use of NLP, which helps assess the emotional involvement of learners using their feedback expressed in the form of text or verbal means. Positive feelings are typically related to knowledge and interest, but negative emotions might suggest disorientation or disinterest [3]. These clues enable AI-powered systems to modify the delivery of presentations or deliver specialized interventions in real time.

Moreover, AI features incorporated into presentations make them more productive and engaging to users through automated routine tasks, smooth slides, and suggestions before delivery on how to improve a presentation [4]. These systems are able to read through the content of a slide, summarize important information, enable smart search and help organize the content in an effective way, all of which lead to a more efficient knowledge transfer.

To further examine the pedagogical effect of these tools, cross-disciplinary research projects have examined AI-based e-learning systems that assist self-directed and unmonitored learning. Such platforms can not only help to conduct the evaluation of performance in real-time but also to compare AI-generated feedback with human ones to prove its usefulness. Initial research is also positive, indicating the possibility and consistency of using AI in assessing student presentations and improving presentation skills in general [5]. In summary, the convergence of AI and presentation technology offers a new paradigm in educational delivery where real-time feedback, emotional analysis, and intelligent content optimization can significantly enrich the teaching-learning experience. This review delves into the landscape of AI-powered presentation assistants, examining the tools and technologies designed to support real-time feedback and content enhancement in higher education and beyond.

A. Structure of the Paper

In structure of the paper, Section 1 discusses Introduction to Presentation Assistants for Real-Time Feedback. Section II reviews the Categories of AI-powered presentation Tools. Section III AI-Powered Presentation Assistant Technologies. Section IV provides Applications and Use Cases. Sections V and VI present a literature review and conclusion.

II. Categories of AI-Powered Presentation Tools

AI-powered presentation tools can be broadly categorized into speech analysis tools, gesture recognition systems, and content optimization platforms, as shown in Figure 1. These categories focus on enhancing verbal delivery, non-verbal communication, and slide design, respectively. Each type leverages specific AI techniques such as NLP, computer vision, or machine learning for tailored feedback.

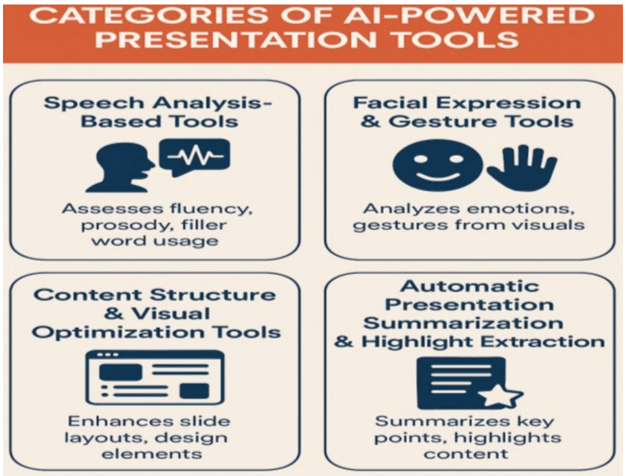

Fig. 1. Categories of AI-Powered Presentation Tools

- **Speech Analysis–Based Tools:** Speech analysis is a foundational component of AI-powered presentation assistants, enabling real-time assessment of fluency, prosody, and filler word usage. Since speech signals are continuous waveforms, short-time analysis (10–30 ms) is used to extract key features like pitch and intonation. While real-time processing poses challenges for visual feedback, recent advances in AI and signal processing have made it possible to deliver actionable feedback during live presentations. Other systems use NLP-augmented speech transcription to analyze speech rate, tone, and clarity across multiple languages for personalized coaching feedback [6][7]. Automated transcription methods can be used to get the transcript from a presentation video. On top of it, to back up word-phrase analysis [8]. Modify the textacy2 natural language processing library so it can extract prepositional phrases (PP), subject-verb object phrases (SVP), verb phrases (VP), and noun phrases (NP).
- **Facial Expression & Gesture tools:** The data related to the facial expression is readily available; hence they are a common visual emotion ID tool. Whether an individual is a member of a particular cultural group or not, it does not matter its cultural background or origin as far as facial expression perceptions and expressions are concerned. There is a minimal influence of culture on expression of emotion on the face as well as basic human emotions are relatively universal [9]. The facial expression-based approach has three sub-elements: video pre-processing, emotion recognition, and spatiotemporal feature extraction. As a first step, the pre-processing module cuts the original video into specific segments, preserving just the facial expressions in each segment [10].
- **Content Structure & Visual Optimization Tools:** The SynLecSlideGen framework, driven by large language models and synthetic slide datasets, has demonstrated improved detection and retrieval of slide elements (text regions, graphics, layout), facilitating visual optimization workflows. AUTOPRESENT demonstrates how AI can automate and enhance slide design by translating natural language instructions into structured slide layouts [11].
- **Real-Time Audience Engagement Analytics:** AI-powered tools now offer real-time audience feedback analysis by tracking engagement metrics such as gaze direction, attention span, and sentiment. Using computer vision and emotion recognition, these systems can provide presenters with live insights into audience interest levels, enabling dynamic adjustments in content delivery. Tools like Tactiq and Mentimeter AI analytics use emotion classifiers and gaze-tracking algorithms to visualize audience responses in real time [12][13].
- **Automatic Presentation Summarization & Highlight Extraction:** Some AI tools now provide automatic summarization of presentation content for quick review or archiving. Using NLP models like BART or T5, they identify key takeaways, create slide-level summaries, or even generate follow-up questions based on presented material. These tools help both presenters and viewers revisit the most important parts efficiently [14].

A. Different Types of AI Tools

Artificial intelligence (AI) tools cover a wide range of technologies, each developed to tackle particular kinds of problems and activities [15]. As illustrated in Figure 2, these can be classified into the following categories:

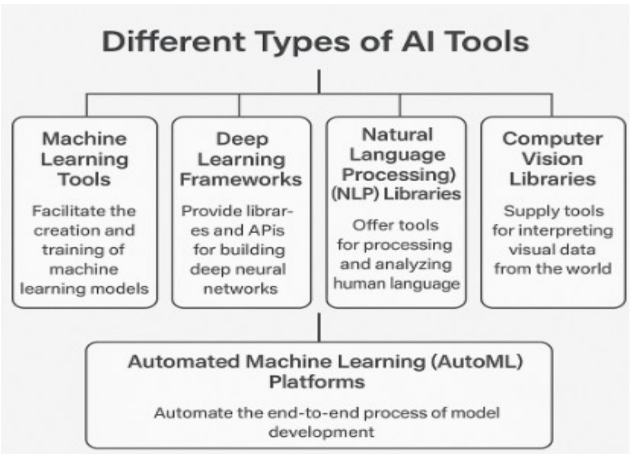

Fig. 2. Classification of Different Types of AI Tools Based on Their Functional Capabilities

- **Machine Learning Tools:** These are employed in the creation and training of machine learning models. As a branch of artificial intelligence, machine learning

(ML) teaches computers to learn by doing and to make decisions or predictions on their own without assistance from humans. ML tools facilitate several paradigms of learning. These are reinforcement learning, supervised learning and unsupervised learning.

- **Deep Learning Frameworks:** A subdivision of machine learning known as deep learning uses multi-layered neural networks to approximate complex data patterns. To build and train these types of networks, several frameworks offer powerful libraries and APIs, including TensorFlow, PyTorch, and Keras. The advanced models that can be developed with these frameworks are only natural language processing, picture and audio recognition, and playing games.
- **Natural Language Processing (NLP) Libraries:** A number of natural language processing (NLP) libraries provide ready-prepared models and human language analysis techniques, such as NLTK, SpaCy and Hugging Face Transformers [16]. Natural language processing (NLP) is used to process and communicate with human language in many applications, including chatbots, sentiment analysis, language translation, and information retrieval, among others.
- **Computer Vision Libraries:** The list of visual data interpretation tasks enabled by libraries such as OpenCV, TensorFlow, and PyTorch is long: picture recognition, object detection, and segmentation are only several examples of computer vision. These libraries are used in many industries such as driverless vehicles, surveillance, and telemedicine (medical imaging).
- **Automated Machine Learning (AutoML) Platforms:** Google Cloud AutoML, H2O.ai, and DataRobot are just a few examples of autoML platforms that aim to automate the entire model development process, making machine learning accessible to those without specialized training [17]. Feature engineering, model selection, data preparation, and hyperparameter adjustment are necessary for users to install high-end ML models with little manual labour.

III. AI-Powered Presentation Assistant Technologies

This section talks about fundamental AI technologies used to generate real-time feedback during the presentation, such as speech analysis, gesture recognition, and slide content coordination to evaluate the whole performance. Figure 3 shows the technologies.

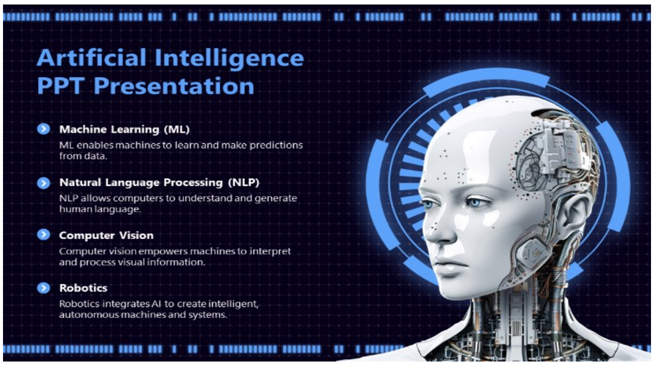

Fig. 3. Technologies in AI-powered Presentations

A. *Algorithms used in Emotion Detection (CTENet)*

The developed CTENet model recognizes emotions on MFCC spectrograms as 2D grayscale images, and it is tested using the RAVDESS and IEMOCAP datasets. It consists of two parallel CNN blocks that include Conv2D, batch normalization, leaky ReLU, max pooling and dropout layers as Figure 4 shows [18]. It maintains a specific dimension on its inputs through zero-padding, and the architecture is optimized using batch size 32 and dropout probability 0.20. The design of the model is robust such that it acquires space and time features essential to producing reliable real-time feedback of emotion activations.

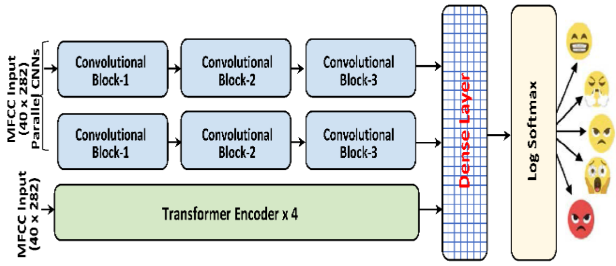

Fig. 4. Architecture of the CTENet Model for Speech Emotion Recognition

B. *Transformer-Based Models for Semantic Coherence in Slide Content*

This approach takes advantage of transformer model-based language models like BERT to produce embeddings with a great deal of contextual information on the text material. Then it clusters them to induce semantically coherent discourse. When used with slide content, this method improves semantic consistency by representing each slide text as document embeddings in the high-dimensional space, which enables an effective grouping of captions and edges in the form of flow [19]. This allows the flow of theme within a presentation to continue, enabling the inclusion of more logical and context-related information than that based on conventional topic modelling approaches.

C. *Use of Explainable AI (XAI) in Interpretability of Feedback*

Explainable AI (XAI) approaches have emerged as the only ways of making the feedback generated in the context of presentation assistants transparent and trustworthy. Model-agnostic methods like LIME and SHAP can be used to see the important effect of an aspect like pitch range or pause rate on sentiment measurement and so enable the user to appreciate why a given fluent score has been allocated. In data-specific situations, one can employ such methods in the form of attention mechanisms or layer-wise relevance propagation (LRP) to explain how the prosody of the speech or the content of the slide entered the picture in terms of the predictive outcome [20]. Survey results suggest that XAI enables trust, comprehensibility and fairness especially when explanations are considered using measures of fidelity, compactness and user-friendly user-centred usability judgement measures.

D. *Improvements Needed for AI Presentation Assistants*

Although the users have already been satisfied with the features offered by AI presentation assistants, they provided some improvement suggestions [21].

- **Enhanced Accessibility and Ease of Use:** The users discussed the necessity to simplify and open AI tools more often. Making the interface and the functionality easy can be very useful in reaching a wider audience even those that might not be tech-savvy. Increased availability of such functions as voice commands, screen readers, and adjustable display settings would make the AI tools more inclusive. With the complexity lowered and the interface made more intuitive, developers can make sure that the AI presentation assistants accessible to all users and able to use them efficiently regardless of their degree of skill.

- **Feedback Mechanisms:** In addition, it is demanded to have strong feedback mechanisms, in which the users are to report problems and offer improvement options. Establishment of proper mechanisms for user feedback is necessary to inform constant enhancements [22]. In-app feedback options, periodic surveying of users, and user forums can be established to give the company a better sense of user experiences and likes/dislikes. The feedback mechanisms can assist the developers in detecting problems and correcting them in a short period, which make the AI tools develop by user needs and demands.
- **Integration into Daily Life:** Several respondents feel that AI would be part of the day-to-day lives in many professions implying that there more acceptability and even dependency on AI tools to ease the work and make it more efficient. This view is a subset of the broader trend of portraying AI technologies in the working process.
- **General Satisfaction:** A small number of users reported that they were happy with the existing capabilities of AI tools, which means the current features satisfy their needs to a certain degree. But with this satisfaction comes the expectation of being able to evolve further and create innovations that enhance the user experience.
- **Future Outlook:** The respondents see the emerging role of AI and its potential impact in a positive way. They are forecasting that in the future, see further advancements that fully integrate AI into various settings and make it an even more necessary instrument. On the basis of this optimism, AI is an innovation, and it is increasingly getting more sophisticated. Future AI tools are expected to give even greater abilities, more predictive analytics, new user interface and even more technology integration by users.

IV. Applications and Use Cases

Presentation is now also used in education, business and oratory to enhance the communication behaviour of the presenter and is often supported by AI. These tools complement speech, gesture, and interaction analysis in real time and enable the presenter to receive actionable feedback. Some of the most intriguing applications of use cases include student presentation training, executive coaching, optimization of virtual meetings, and the list goes on and on, hence it can be used in both any learning and business environment.

A. Evaluation Criteria

Real-time presentation feedback systems are typically evaluated across four key dimensions: real-time feedback capability, which assesses how promptly the tool analyzes speech and non-verbal cues; UI/UX, focusing on ease of use, clarity of the dashboard and feedback delivery; personalization, meaning how well the tool tailors feedback based on user's unique speech habits or improvement goals; and privacy, addressing data retention, confidentiality, and user control over recorded sessions [23]. These criteria derive from literature and detailed reviews highlighting their centrality to effective AI-assisted coaching (e.g., NLP and ASR for timely filler-word detection, dashboards for user-friendly learning, and data policy transparency).

1) Real-Time Feedback System

The system provides real-time feedback by capturing and analyzing presenter behavior using the Kinect sensor, without requiring user interaction. It records RGB, depth, skeleton, and audio data streams, enabling automatic gesture and speech analysis through Microsoft's Kinect SDK and System [24]. Speech API. Although minor synchronization delays may occur, multi-threaded decoding and time-stamp alignment ensure smooth playback. This allows visualization of posture, movement, and speech dynamics for performance evaluation. Figure 5 shows the graphical user interface used for data capture and feedback visualization.

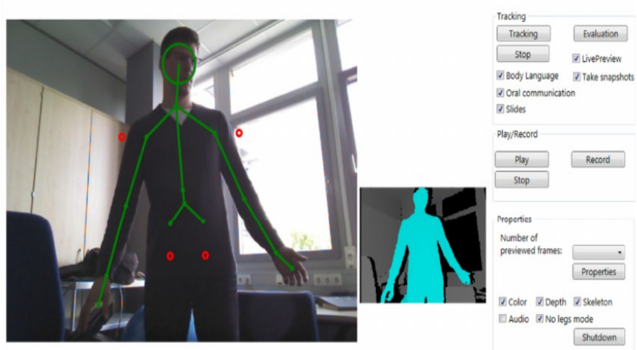

Fig. 5. GUI (Main View) of the Feedback System

2) Personalization

Recent frameworks such as MOSAIC-F demonstrate how multimodal learning analytics, combining human rubrics, gaze tracking, speech, and physiological signals, can produce highly personalized feedback tailored to individual performance patterns [25]. Yoodli also adapts feedback based on user input, asking about known weaknesses and tracking progress across sessions to customize future coaching snippets.

3) Privacy

Yoodli prioritizes user privacy via default private sessions, optional deletion of recordings, and enterprise-grade data encryption. Enterprise clients like coaches can opt out of using their uploaded content to train models, ensuring content stays under their control. However, reviews note that advanced privacy options are only available in enterprise tiers, and concerns remain over terms of service granting broad usage rights unless explicitly restricted.

4) Voice Integration

Another notable suggestion was the integration of voice recognition capabilities. The user think of a world that allows them to write what they want and the AI convert it into presentation material to make the content creation process simplified. Integration of voices can make the AI presentation assistants far more desirable and usable, particularly to those who consider typing and talking to be optional combinations. The tools can interpret the interactive and verbal speech and produce long-lasting presentations with the help of the power of artificial intelligence that is combined with a high level of efficiency and convenience.

5) Accessibility and User-Friendly Interface

The respondents have highly emphasized on the need to design a more user friendly and accessible interface. The need for improved accessibility features demonstrates the significance of increasing the inclusion and accommodation of diverse user needs by means of developing AI tools. The solutions suggested included: simplicity of user interface and voice command and assistive technologies. These advancements allow AI tools to be more accessible to a wider audience, including individuals with disabilities, and help the online environment feel more humanized, as well. Moreover, by simplifying the interface, the developers of the AI able to assure that individuals at various stages of technical expertise more eager to use these tools more productively.

B. Strengths of Current Tools

The existing AI-based presentation software is capable of identifying vocal fillers, demonstrates a steady ability to recognize emotions, and moderate scoring of pronunciation. They are objective, time-efficient, fatigue-free and can be scaled. These benefits increase self-directed learning and allow errors to be corrected instantly and in a manner that can be exploited.

1) Effective Detection of Vocal Fillers

The ability of the AI tools to identify the majority of word-based vocal fillers such as, so, it has been found to be high, and this is normal among most non-native English speakers who tend to use the tools during the presentation process. This aspect makes the job of human raters much easier since they would otherwise have to count manually these interruptions. The on-screen feedback also makes it possible to identify and eradicate filler words during the practice, which means that the tool is especially helpful for self-guided language enhancement.

2) Impartial and Consistent Emotion Recognition

Emotion scoring tools AI can be used to offer objective and unbiased assessments of the facial manifestations of presenters that may be difficult to detect or measure by human evaluators because of tiredness or bias. Even with the low level of illumination on the face, or when the window where video signal is shown is narrow, the AI system able to analyze the facial expression. It can also detect a broader range of emotions, including anger, contempt, happiness, and surprise, and provides extensive and objective information that a human rater would not easily identify continuously.

3) Moderately Accurate Pronunciation Scoring

Accuracy scores calculated by AI, depending on the similarity of pronunciation to native pronunciation on a phonemic level, were correlated with the ratings given by humans with a medium strength. This is also an objective test and, therefore a good addition to learners who want to perfect their pronunciation.

4) Scalable and Time-Efficient Evaluation

The rapidity and consistency in assessing many presentations are one of the most useful capabilities of AI presentation assistance. AI tools can assess on a large scale, and they differ from human raters. They do not get tired, biased, or slow as they are doing the work. This scalability best suits education institutions or training programs that deal with a very high number of students.

5) Support for Self-Directed Learning

Presentation assistants powered by AI can provide real-time and actionable feedback to promote self-directed learning in a self-paced manner. Without being monitored at all times, students able to practice their presentation, and their performance critiqued by automated systems, allowing them to make iterations. Vocal filler detection, pronunciation scoring and elementary emotional analysis are other features, that allow learners to identify and resolve certain points of weakness in the long run.

C. Challenges of AI tools

The limitations of presentation assistants powered by AI include low accuracy in a variety of linguistic and cultural conditions, potential user dependency, and data confidentiality issues. The quality of feedback might be hindered by technical issues, i.e. delays in processing in real-time, ambivalence in interpreting the gestures. Also, there are still moral issues related to the use of data and opacity thereof.

1) Limited Adaptability to Individual Learning Styles

The AI feedback system does not have a customization option and adaptive learning module to address diverse needs and organizational level of the students. It administers standardized criteria, which do not necessarily relate to the speaking styles, cultural communication, or individual learning objectives and restrict its pedagogical flexibility.

2) Lack of Qualitative Feedback

In (AI) the output is usually in numbers or categories rather than the rich and descriptive feedback that can be obtained with human raters. Such comments as fluent with very few distracting pauses or exhibited interest by talking in varying tones give the learners a clue as to how they can work on their performance. Such feedback can help learners interpret their scores or identify areas for improvement, which they may not be able to do without it.

3) Emotion Recognition Lacks Contextual Depth

The AI tool was only able to recognize basic facial expressions, but it does not take into account context-dependent signals such as enthusiasm, engagement, or rhetorical tone of speech. The human raters would be able to give more subtle information such as self-absorbed and unaware of the audience or uplifting voice because the AI could not understand that. This restricts the AI's ability to provide significant feedback regarding presentation delivery and audience interaction.

4) Inaccurate Fluency Assessment for Low-Proficiency Speakers

The AI tool mainly estimates the fluency by nonsensical words with silent gaps in between in utterance which is a poor parameter of human judgment of a non-fluent speaker. Although it can very well identify the hesitation among fluent users, it is incapable of the more rudimentary elements like tempo of the sentence, chunking, or associating, which is why it does not correlate very well with human scoring. It renders the tool ineffective in detecting weak speakers or providing practical fluency reports to learners.

V. Literature Review

This part is a literature analysis of AI-driven presentation assistants. It examines different mechanisms and structures that help in strengthening user performance, interactions and provision in the course of presentations.

Markowitz et al. (2025) strive to combine these approaches by using data from external knowledge graphs and interactions in other domains to provide predictions in a new domain that would be hard to achieve using just one of these sources. Create a new model and demonstrate how these techniques operate together using a dataset derived from millions of users' requests for content in three categories—books, music, and movies—all via a real-time virtual assistant app. Show considerable improvement in both domain-wide and user-specific recommendation quality [26].

Kumar et al. (2024) presented a hand gestures approach that not only can control presentations but is also useful for facilitating drawing and managing system volume, which enriches the user experience and engagement by offering an interface that eliminates the need for keyboard and mouse. With the help of this algorithm, the computer system accurately responds to hand movements in real time. With this model, one can easily navigate through slides, play, or pause videos, and perform multiple presentation-related tasks using hand gestures. It boosts mobility and enhances audience engagement by removing the limitations posed by traditional input methods. It also enables users to draw electronically on the system [27].

Miladinovic and Schefer-Wenzl (2024) examined the purpose of final theses and the actions to be taken to achieve this goal in the context of modern technological innovations. By employing the approach proposed by Dettmer, it is necessary to identify and examine the root causes of challenges caused by GenAI tools. The findings reflect the necessity of changes at the curriculum level to address the mentioned challenges. These involve transferring the priority responsibilities of thesis elements that can be easily created by AI to the input of learners into a given research study. Such adaptations are expected to improve the quality and originality of final theses, and they also promote better institutionalization of the students into their research projects, as well as generating new and novel forms of presenting theses in a multimedia form [28].

Nguyen et al. (2023) presented a range of possibilities to present shorelining information using spatial-audio auditory sensory augmentation to help individuals with low vision or blindness. Shorelining is the practice of following a path or paths in relation to the contours of structures within a constructed environment. It is typically associated with the usage of a white cane. The goal is to support legible, rational modes of transportation. One method uses a series of two auditory ear cons that are spatially portrayed along the beach and at the participant's left or right side to deliver aural shorelining indications [29].

Lee et al. (2023) presented three study tasks that utilize principles of psychology and multimedia learning to assess The ability of a vision-language model to understand multimodal material includes development of slide explanations, figure-to-text retrieval, and text-to-figure retrieval. To help with the implementation and baseline-setting of these activities, manual annotations are offered. A comparison between baselines and human student outcomes reveals that state-of-the-art vision-language models (zero-shot and fine-tuned) struggle in four areas: learning novel visual mediums, technical language, long-range sequences, and weak cross-modal alignment between slides and spoken text. PolyViLT, a novel multimodal transformer learnt using a multi-instance learning loss, is offered to enhance current retrieval techniques [30].

Lai et al. (2023) aimed to demonstrate how well SmartPresenter served as a tool for both students' and instructors' use in developing their skills in giving oral presentations in English. In addition to specialized lesson plans and helpful hints for English as a second language learners, SmartPresenter provides a wealth of AI-generated feedback for independent practice of oral presentations, including comments on the presenter's facial expressions, eye contact, clarity of sight, verbal fillers, fluency, and pronunciation. Students and instructors alike valued AI comments more highly; therefore, after survey results showed that students required help with presentation abilities, Smart Presenter was improved and tested in a database course [31].

Neretin and Kharchenko (2022) propose markers of the accuracy of this data, particularly its completeness and dependability, and critically examine the methods and resources currently used to collect data on vulnerabilities in order to evaluate the cybersecurity of AI systems. An IDEF (Integrated Definition Framework) architecture, with multiple levels, explains how to gather and examine vulnerabilities in AI systems. The steps that comprise the procedure are as follows: group processing, arranging data from several sources, and coordinating the outcomes with Big Data tools; choosing and presenting data in a specific format regarding the AI system's vulnerabilities; and finally, evaluating the severity of those vulnerabilities using a variety of methods, including IMECA (Intrusion Modes and Effects Criticality Analysis) [32].

Table I presents a structured overview of key research on AI-powered presentation assistants, highlighting focus areas, key findings, challenges, and contributions.

TABLE I. SUMMARY OF LITERATURE REVIEW BASED ON AI-POWERED PRESENTATION ASSISTANTS

Reference	Focus Area	Key Findings	Challenges	Key Contribution
Markowitz et al. (2025)	Cross-domain recommendations using external knowledge	Combining interaction data and knowledge graphs improves predictions	Cross-domain data integration and the new user cold-start problem	Unified model that boosts recommendations across videos, music, and books
Kumar et al. (2024)	Real-time hand gesture control system	Gesture-based system for controlling presentations and system volume	Eliminating dependence on traditional inputs	Enables touchless, real-time interaction, enhancing user mobility
Miladinovic and Schefer-Wenzl (2024)	Impact of GenAI tools on thesis quality	Curricula should emphasize student-driven research	Overreliance on AI-generated content	Suggests curriculum changes to improve originality and engagement
Nguyen et al. (2023)	Assistive spatial-audio technology for the visually impaired	Spatial auditory cues help blind users follow logical paths	Providing accurate spatial-audio feedback	Shorelining cues via audio improve independent navigation

Lee et al. (2023)	Vision-language model for slide-based learning tasks	Vision-language models struggle with complex multimodal tasks	Poor cross-modal alignment and handling of long text	Introduces PolyViLT: improved multimodal transformer for educational retrieval tasks
Lai et al. (2023)	AI feedback tool for oral presentation skills	SmartPresenter supports autonomous learning with multimodal feedback	Need for personalized and scalable feedback	AI tool improves EAL students' oral presentation skills and grading efficiency
Neretin and Kharchenko (2022)	Cybersecurity evaluation of AI systems	The multi-level IDEF model helps assess vulnerabilities	Ensuring data completeness and reliability	Framework for structured vulnerability information collection using Big Data

VI. CONCLUSION AND FUTURE WORK

The use of AI is introducing a paradigm shift in various industries, offering a new solution to old, challenging problems. AI presentation assistants are transforming the face of oratory, rhetoric and training by providing multimodal feedback on speech, gesture and visual presentations in real-time. A combination of new AI-powered tools, including emotion detection (deep learning), content summarization (NLP), and explainable AI (feedback transparency), has enabled the creation of scalable, efficient, and objective tools. Such systems have demonstrated great potential in learning, business and work environments through the provision of self-directed learning and enhanced presentation. Notwithstanding these advancements, several limitations remain. Current tools are likely to fail to adapt to new linguistic and cultural contexts, provide limited information with little value in a qualitative form and face technical and ethical challenges related to data privacy and situational awareness. It is also concerning that low-proficiency speakers are required to take the fluency test, and that emotional nuance can be detected.

Future research is required to improve personalization and flexibility of reinforcement learning-based and user-cantered design feedback systems. The next way to improve user experience is to introduce richer, qualitative feedback, enhanced emotive recognition with context awareness, and make it available to all. Additionally, the advancement of privacy-saving AI technology and the development of universal assessment indicators will play a crucial role in the increased uptake and trust in AI-based presentation technology.

REFERENCES

[1] D. Sewasew, M. Mengestie, and G. Abate, "A comparative study on PowerPoint presentation and traditional lecture method in material understandability, effectiveness and attitude," *Educ. Res. Rev.*, vol. 10, no. 2, pp. 234–243, 2015, doi: 10.5897/ERR2014.2027.

[2] J. Prakash, R. Swathiramya, G. Balambigai, R. Menaha, and J.S. Abhirami, "AI-Driven Real-Time Feedback System for Enhanced Student Support: Leveraging Sentiment Analysis and Machine Learning Algorithms," *Int. J. Comput. Exp. Sci. Eng.*, vol. 10, no. 4, Dec. 2024, doi: 10.22399/ijcesen.780.

[3] A. R. Bilipelli, "Application of AI and Data Analysis for Classification of Student Success in Large-Scale Educational Dataset," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 4, no. 6, pp. 428–441, Nov. 2024, doi: 10.48175/IJARSCT-22564.

[4] S. Maniyar, V. Trivedi, A. Mondal, A. Mishra, and C. V. Jawahar, "AI-Generated Lecture Slides for Improving Slide Element Detection and Retrieval," pp. 1–40, 2025.

[5] J. Chen, P. Lai, A. Chan, V. Man, and C.-H. Chan, "AI-Assisted Enhancement of Student Presentation Skills: Challenges and Opportunities," *Sustainability*, vol. 15, no. 1, Dec. 2022, doi: 10.3390/su15010196.

[6] S. Bhuvaneshwaran, "The Fluent Speaker' S Guide," *Int. J. Nov. Res. Dev.*, vol. 9, no. 12, pp. 301–310, 2024.

[7] R. Q. Majumder, "Machine Learning for Predictive Analytics: Trends and Future Directions," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 4, pp. 3557–3564, 2025.

[8] H. Zeng, X. Wang, Y. Wang, A. Wu, T. C. Pong, and H. Qu, "GestureLens: Visual Analysis of Gestures in Presentation Videos," *IEEE Trans. Vis. Comput. Graph.*, vol. 29, no. 8, pp. 3685–3697, 2023, doi: 10.1109/TVCG.2022.3169175.

[9] S. C. Leong, Y. M. Tang, C. H. Lai, and C. K. M. Lee, "Facial expression and body gesture emotion recognition: A systematic review on the use of visual data in affective computing," *Computer Science Review*. 2023. doi: 10.1016/j.cosrev.2023.100545.

[10] J. Wei, G. Hu, X. Yang, A. T. Luu, and Y. Dong, "Learning facial expression and body gesture visual information for video emotion recognition," *Expert Syst. Appl.*, 2024, doi: 10.1016/j.eswa.2023.121419.

[11] J. Ge *et al.*, "AutoPresent: Designing Structured Visuals from Scratch," 2025.

[12] C. Scheirer and C. Harrison, "DynaTags: Low-Cost Fiducial Marker Mechanisms," *ACM Int. Conf. Proceeding Ser.*, pp. 432–443, 2022, doi: 10.1145/3536221.3556591.

[13] N. Prajapati, "The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications," *ESP J. Eng. Technol. Adv.*, vol. 5, no. 2, pp. 16–22, 2025, doi: 10.56472/25832646/JETA-V5I2P103.

[14] F. Jiao, Y. Guo, X. Song, and L. Nie, "MERIt: Meta-Path Guided Contrastive Learning for Logical Reasoning," *Proc. Annu. Meet. Assoc. Comput. Linguist.*, pp. 3496–3509, 2022, doi: 10.18653/v1/2022.findings-acl.276.

[15] S. Akhand *et al.*, "Exploring Ai Tools: Types, Applications, Challenges, And Future Trends," *Educ. Adm. Theory Pract.*, vol. 30, 2024, doi: 10.53555/kuey.v30i5.7038.

[16] D. Patel, "AI-Enhanced Natural Language Processing for Improving Web Page Classification Accuracy," *ESP J. Eng. Technol. Adv.*, vol. 4, no. 1, 2024, doi: 10.56472/25832646/JETA-V4I1P119.

[17] R. Patel and P. Patel, "A Survey on AI-Driven Autonomous Robots for Smart Manufacturing and Industrial Automation," *Tech. Int. J. Eng. Res.*, vol. 9, no. 2, 2022, doi: 10.56975/tijer.v9i2.158819.

[18] R. Ullah *et al.*, "Speech Emotion Recognition Using Convolution Neural Networks and Multi-Head Convolutional Transformer," *Sensors*, vol. 23, no. 13, p. 6212, Jul. 2023, doi: 10.3390/s23136212.

[19] B. Patel, H. Mallisetty, and K. M. Rao, "Artificial Intelligence Helper Application for Delivering Effective Presentations," 2024.

[20] L. Qudus, "Resilient Systems: Building Secure Cyber-Physical Infrastructure for Critical Industries Against Emerging Threats," *Int. J. Res. Publ. Rev.*, vol. 6, no. 1, Jan. 2025, doi: 10.55248/gengpi.6.0125.0514.

[21] M. Pari, "Artificial Intelligence (AI) Presentation Assistants: Current Trends, Experiences, and Future Directions for Language Pedagogy," *Int. J. Pract. Pedagog. Issues English Educ.*, vol. 2, no. 3, pp. 54–66, 2024, doi: 10.22034/ijpie.2024.469410.1030.

[22] R. Patel and P. Patel, "A Machine Learning-Based Detection and Recognition Approach for 1D Barcode Scanning Through Visible Light Communication," *Int. J. Res. Anal. Rev.*, vol. 8, no. 2, 2021, doi: 10.56975/ijrar.v8i2.315834.

[23] N. Thakkar *et al.*, "Can LLM feedback enhance review quality? A randomized study of 20K reviews at ICLR 2025," pp. 1–30, 2025.

[24] S. Kopf, D. Schön, B. Guthier, R. Rietsche, and W. Effelsberg, "A Real-time Feedback System for Presentation Skills," no. June, pp. 1686–1693, 2015.

[25] A. Becerra, D. Andres, P. Villegas, R. Daza, and R. Cobos, "MOSAIC-F: A Framework for Enhancing Students' Oral

Presentation Skills through Personalized Feedback," vol. 2682, 2025, [Online]. Available: http://arxiv.org/abs/2506.08634

[26] E. Markowitz *et al.*, "Knowledge Enhanced Multi-Domain Recommendations in an AI Assistant Application," in *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2025, pp. 1–5. doi: 10.1109/ICASSP49660.2025.10889248.

[27] A. Kumar, R. Pandey, K. Alam, H. Anandaram, K. Joshi, and S. Chaudhary, "Hand Gesture based AI Controller for Presentation, Virtual Drawing and System Volume Management," in *2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN)*, Jul. 2024, pp. 620–624. doi: 10.1109/ICIPCN63822.2024.00107.

[28] I. Miladinovic and S. Schefer-Wenzl, "From Learners to Contributors: Redefining Bachelor's and Master's Theses in the Context of Generative AI Tools," in *2024 15th International Conference on Distance Learning and Education (ICDLE)*, 2024, pp. 15–19. doi: 10.1109/ICDLE63439.2024.00010.

[29] M. Nguyen *et al.*, "An Evaluation of the Presentation of Acoustic Cues for Shorelining Techniques," in *2023 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2023, pp. 1–9. doi: 10.1109/I3DA57090.2023.10288622.

[30] D. W. Lee, C. Ahuja, P. P. Liang, S. Natu, and L.-P. Morency, "Lecture Presentations Multimodal Dataset: Towards Understanding Multimodality in Educational Videos," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2023, pp. 20030–20041. doi: 10.1109/ICCV51070.2023.01838.

[31] P. Lai, J. Chen, V. Man, and C. H. Chan, "A New Frontier in AI-Assisted English Oral Presentation Assessment," in *2023 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*, Nov. 2023, pp. 1–8. doi: 10.1109/TALE56641.2023.10398320.

[32] O. Neretin and V. Kharchenko, "Model for Describing Processes of AI Systems Vulnerabilities Collection and Analysis using Big Data Tools," in *2022 12th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Dec. 2022, pp. 1–5. doi: 10.1109/DESSERT58054.2022.10018811.

