

AI Twin – 3D Avatar for Automated Meeting Presence

Shoib Vakkund*, Suhana Banu**, Shoyab Ahamad***, Shaikhshavali****,

*(CSE-AI, Ballari Institute of Technology & Management and Ballari

Email: shoibvakkund@gmail.com)

** (CSE-AI, Ballari Institute of Technology & Management and Ballari

Email: suhanabanu7624@gmail.com)

*** (CSE-AI, Ballari Institute of Technology & Management and Ballari

Email: shoyabahamed.r@gmail.com)

**** (CSE-AI, Ballari Institute of Technology & Management and Ballari

Email: shaikhani841@gmail.com)

Abstract:

In today's rapidly evolving digital work environment, virtual meetings have become an essential part of communication, collaboration, and decision-making. However, individuals may not always be able to attend every meeting due to scheduling conflicts, workload, or other commitments, which can result in missed discussions and important information. Traditional meeting tools mainly provide basic video conferencing and recording facilities but do not provide an intelligent virtual presence that can represent an absent participant and capture meeting information automatically. This project presents an AI Twin – 3D Avatar for Automated Meeting Presence, an intelligent virtual meeting system designed to provide a human-like digital presence on behalf of a participant during online meetings.

The proposed system integrates Artificial Intelligence, 3D avatar technology, speech recognition, natural language processing, and real-time communication technologies to create an AI-powered digital twin capable of attending meetings automatically. The AI Twin can represent the participant through a 3D avatar, capture and convert meeting conversations into text, identify important discussions, generate structured meeting notes, and provide an AI-generated summary of the meeting. The system is designed to maintain a continuous record of meeting information so that the participant can review important conversations, decisions, and action items even when they are unable to attend the meeting personally. By combining automated meeting presence, real-time communication, speech-to-text processing, and intelligent meeting analysis within a single platform, the proposed system reduces the dependency on manual note-taking and conventional meeting recordings. The platform also supports features such as multiparticipant virtual meetings, live transcription, AI-generated summaries, meeting chat, screen sharing, and personalized 3D avatar representation to provide a more interactive and human-like meeting experience.

The developed system aims to improve meeting accessibility, productivity, and information retention by allowing users to maintain a virtual presence even when they cannot attend a meeting physically or virtually. It provides a scalable and user-friendly solution for students, professionals, teams, and organizations that frequently participate in online meetings. Overall, the proposed system demonstrates how AI, 3D digital avatars, speech processing, and real-time communication technologies can be combined to create an intelligent virtual representation of a person and enable automated meeting presence.

Keywords — AI Twin, 3D Avatar, Automated Meeting Presence, Artificial Intelligence, Speech Recognition, Natural Language Processing, Real-Time Communication, Virtual Human Representation, Meeting Transcription, AI Meeting Summary,

I. INTRODUCTION

Being present has always been an important part of communication, collaboration, and decisionmaking,

but in today's fast-changing digital world, attending every meeting has become increasingly difficult. With the growth of remote work, online education, virtual teams, and distributed organizations, people

are expected to participate in meetings from different locations and time zones. While digital communication has made it easier to connect with others, it has also created a new challenge: people cannot always be present when important conversations happen.

Traditionally, online meetings have depended on the physical or virtual presence of participants. Although video conferencing platforms provide features such as video calls, chat, screen sharing, and meeting recordings, these tools mainly focus on connecting people rather than creating an intelligent representation of a person. When someone is unable to attend a meeting, they usually have to depend on recordings, notes from other participants, or messages to understand what happened. This can be time-consuming and may result in important information being missed.

Another challenge is the amount of information generated during modern meetings. A single meeting can contain discussions, questions, ideas, decisions, technical explanations, and action items involving multiple participants. Trying to remember everything while actively participating can be difficult. At the same time, manually taking notes can distract participants from the actual conversation. Reviewing a complete meeting recording later can also require a significant amount of time.

With the development of Artificial Intelligence, speech recognition, Natural Language Processing, real-time communication, and 3D avatar technologies, there is an opportunity to create a more natural and intelligent form of digital presence. Instead of simply recording a meeting, an intelligent digital twin can represent a person, remain present during the meeting, understand the conversation, and bring the important information back to that person.

To address this need, AI Twin – 3D Avatar for Automated Meeting Presence is introduced as an intelligent and human-centered meeting solution. The system creates a digital representation of a user through a 3D avatar that can maintain their virtual presence during an online meeting. The AI Twin is designed to capture meeting conversations, process the information, and help the user understand what

happened even when they were unable to attend personally.

The proposed system combines Artificial Intelligence, 3D avatar technology, speech recognition, Natural Language Processing, realtime communication, and automated meeting analysis to create a more meaningful virtual presence. During a meeting, the AI Twin can capture conversations and convert speech into text, allowing the system to maintain a live understanding of the discussion. Important topics, decisions, questions, and action items can then be organized into meaningful meeting information.

One of the key advantages of this approach is that it goes beyond conventional meeting recording. A recording simply stores what happened, while an AI Twin is designed to help understand what happened. The system can generate structured meeting notes and summaries so that users can quickly understand the important parts of a conversation without having to watch the entire meeting again.

Another important aspect of the system is its human-centered design. The 3D avatar provides a visual identity for the AI Twin and creates a more natural connection between the user and the digital representation. Rather than interacting with a simple automated bot, participants can see a recognizable Page 2 virtual presence representing the absent user. This creates a bridge between human communication and intelligent digital technology. The system also aims to make meeting participation more accessible. Professionals may have overlapping meetings, students may have academic commitments, and team members may be located in different time zones. In such situations, an AI Twin can help maintain continuity by being present during a meeting and preserving the important information for later review.

This allows users to stay connected with discussions even when they cannot be physically or virtually present. In addition, the platform can bring several meeting capabilities together in one environment, including multi-participant video communication, live chat, screen sharing, live transcription, AI-generated summaries, meeting history, and 3D

avatar representation. By combining these capabilities, the system reduces the need to depend on multiple disconnected tools and creates a more intelligent and organized meeting experience.

Therefore, AI Twin – 3D Avatar for Automated Meeting Presence aims to move beyond conventional virtual meetings by creating a digital presence that is intelligent, interactive, and centered around human needs. By combining AI with 3D avatar technology and real-time communication, the system provides a way for people to remain connected to important conversations even when they cannot be present themselves

II. LITERATURE REVIEW

Virtual meetings and remote collaboration have become important areas of research with the rapid growth of digital communication technologies. Online meetings are widely used in education, business, and distributed organizations. However, conventional meeting platforms mainly focus on video, audio, chat, and screen sharing, while providing an intelligent representation of an absent participant remains a challenge.

This has encouraged researchers to explore virtual agents, digital humans, avatars, and Artificial Intelligence for improving remote presence. Several studies have highlighted the problem of information loss when individuals are unable to attend meetings. Important discussions, decisions, and action items may be missed, requiring absent participants to depend on recordings or notes from others.

Research on automated meeting systems has therefore focused on speech recognition, meeting transcription, speaker identification, and information extraction to automatically capture and organize meeting content. Recent studies have also explored the use of Artificial Intelligence and Natural Language Processing for meeting analysis. Speech-to-text technologies can convert conversations into transcripts, while NLP techniques can identify important topics, decisions, and action items.

AI-based summarization can further convert lengthy discussions into concise meeting notes, reducing the time required to review complete recordings. Another important research area is the development

of virtual avatars and digital humans for human-computer interaction. Three-dimensional avatars can provide visual identity and create a stronger sense of presence compared with traditional interfaces. These developments provide a foundation for creating AI Twins that can represent individuals in virtual environments and maintain a digital presence

when they are unavailable. Real-time communication technologies such as WebRTC and WebSocket-based systems also play an important role in modern virtual meeting platforms. They enable real-time audio, video, messaging, and interaction between multiple participants. Combining these technologies with AI can create intelligent meeting environments capable of supporting both communication and automated meeting analysis. However, intelligent meeting systems still face challenges such as speech recognition errors, background noise, multiple speakers, latency, AI-generated inaccuracies, privacy, and data security.

Human-like avatars also require clear identification so that users understand when they are interacting with an AI representation rather than the actual person.

Overall, existing research demonstrates the potential of AI, speech recognition, Natural Language Processing, 3D avatars, digital humans, and real-time communication in creating intelligent virtual meeting environments. However, there is still scope for integrating these technologies into a single system that provides an automated and recognizable virtual presence. These gaps motivate the development of an AI Twin – 3D Avatar for Automated Meeting Presence, supporting the vision of #Humanise, where technology becomes more natural, human-centered, and connected.

III. METHODOLOGY

The methodology of the proposed AI Twin – 3D Avatar for Automated Meeting Presence is designed to provide an intelligent and structured approach for maintaining virtual meeting presence through an AI-powered digital representation. The development of the system begins with analyzing the requirements of

users who may be unable to attend meetings due to scheduling conflicts, workload, or other commitments. Based on these requirements, the system is designed using appropriate architectural models, communication workflows, and user interface layouts to ensure smooth interaction between participants and the AI Twin. The methodology focuses on creating an integrated framework where meeting information such as participant details, audio, video, conversations, and user preferences is collected and processed using real-time communication and AI technologies. Speech recognition and Natural Language Processing techniques are used to convert meeting conversations into text and identify meaningful information such as important topics, decisions, and action items.

The process flow of the system begins with user registration and meeting creation, where users can create or join a virtual meeting through a unique meeting link. Participants can provide their names and meeting details, while the user can configure their AI Twin and select a suitable 3D avatar representation. During the meeting, real-time communication technologies such as WebRTC and Socket.IO enable audio, video, chat, and participant interaction. The AI Twin maintains a virtual presence through its 3D avatar while meeting conversations are captured and processed through speech-to-text technology. The generated transcript is then analyzed using AI and Natural Language Processing techniques to identify important discussions and generate structured meeting information. This process allows the system to convert lengthy conversations into useful and understandable meeting content. The proposed solution is developed as an intelligent virtual meeting platform that extends conventional video conferencing by introducing an AI-powered digital presence. Unlike traditional systems that mainly depend on physical participant attendance or simple meeting recordings, the proposed system combines a 3D avatar, real-time communication, transcription, and AI-based meeting analysis. The AI Twin can represent a user during a meeting, capture important information, and generate summaries containing key discussion points and action items. Users can later

review the meeting transcript and AI-generated summary without having to watch the complete meeting recording.

A key component of the methodology is the integration of the 3D avatar and AI processing layer. The avatar provides a visual representation of the user within the meeting environment, while the AI layer processes meeting information in the background. The system can maintain meeting history and organize transcripts and summaries for later access. Features such as live chat, screen sharing, participant management, live transcription, AI-generated summaries, and meeting history are integrated into the platform to provide a complete meeting experience. The interface is designed to be simple and responsive so that users can access the system through desktops, laptops, and other supported devices.

Reliability, privacy, and security are also considered important aspects of the methodology because meeting conversations may contain personal, academic, or organizational information. The system is designed to securely manage user information, meeting details, transcripts, and generated summaries. Backend services are responsible for handling communication, data processing, and storage, while the frontend provides an interactive interface for users and participants. The architecture is also designed with scalability in mind, allowing future integration of more advanced 3D avatars, improved voice interaction, emotion-aware communication, personalized AI behavior, and external meeting platforms.

Overall, the methodology combines Artificial Intelligence, 3D avatar technology, speech recognition, Natural Language Processing, WebRTC, real-time communication, automated transcription, and AI summarization to create an intelligent virtual meeting presence. By integrating these technologies into a single platform, the proposed system enables users to maintain a digital presence even when they cannot attend a meeting personally. The approach supports the vision of #Humanise by combining automation with a more natural, human-centered digital representation and creating a meeting

experience where technology helps people remain connected, informed, and present.

IV. RESULTS AND DISCUSSION

The developed AI Twin – 3D Avatar for Automated Meeting Presence produced encouraging results in demonstrating how Artificial Intelligence and virtual avatar technologies can improve participation in online meetings. The system was able to successfully support virtual meetings with multiple participants while maintaining an AI-powered digital presence through a 3D avatar. It also captured meeting conversations, generated transcripts, and provided AI-based meeting summaries containing important discussions and action items.

One of the major outcomes observed from the system is the ability to maintain a virtual presence when a user is unable to attend a meeting personally. Instead of relying only on conventional meeting recordings, the AI Twin provides a recognizable digital representation of the user within the virtual meeting environment. This helps maintain continuity and allows the user to access important meeting information after the session.

The results also demonstrate that integrating multiple meeting functions into a single platform improves accessibility and convenience. Features such as multi-user video communication, live chat, screen sharing, live transcription, 3D avatar representation, and AI-generated summaries are available within one environment. Users can participate in meetings, communicate with other participants, and later review the generated transcript and summary without depending on multiple separate applications.

From a performance perspective, the system demonstrated effective real-time communication between participants using WebRTC and Socket.IO-based technologies. Audio and video communication, participant joining, messaging, and meeting interactions were handled through the integrated communication framework. The speech recognition component was able to convert spoken meeting content into text, while the AI processing layer

generated structured summaries from the captured information.

The discussion of results indicates that the proposed system addresses several limitations of conventional online meeting platforms, including dependence on physical attendance, difficulty reviewing lengthy meetings, and the manual effort required for note-taking. By combining an AI-powered 3D avatar with automated transcription and summarization, the platform provides a more intelligent and organized approach to virtual meeting participation.

Another important observation is the potential of the AI Twin concept to improve productivity and continuity for professionals, students, and distributed teams. Users who are unable to attend a meeting can still obtain important information through transcripts, summaries, decisions, and action items. The 3D avatar also provides a more human-oriented form of digital representation compared with a conventional automated bot.

While the results are promising, the discussion also highlights opportunities for future improvement. Advanced 3D avatar animation, improved speech recognition, multilingual support, emotion-aware interaction, personalized AI responses, stronger privacy mechanisms, and integration with external platforms such as existing video conferencing services could further enhance the system. The use of reliable AI models and scalable cloud infrastructure can also improve performance for larger meetings.

Overall, the results and discussion demonstrate that the AI Twin – 3D Avatar for Automated Meeting Presence successfully achieves its intended objectives by providing an intelligent, interactive, and automated virtual meeting presence. The outcomes demonstrate the potential of combining AI, 3D avatars, real-time communication, speech recognition, and meeting analysis to create a more natural and human-centered digital meeting experience.

Fig. 1 A sample line graph using colors which contrast well both on screen and on a black-and-white hardcopy

Fig. 2 shows an example of a low-resolution image which would not be acceptable, whereas Fig. 3 shows an example of an image with adequate resolution. Check that the resolution is adequate to reveal the important detail in the figure.

V. CONCLUSION

In conclusion, the AI Twin – 3D Avatar for Automated Meeting Presence provides an intelligent and innovative solution to address the limitations of conventional online meeting systems. The project was developed with the objective of maintaining a virtual presence for users who may be unable to attend meetings personally. Through the integration of Artificial Intelligence, 3D avatar technology, speech recognition, Natural Language Processing, and real-time communication, the system demonstrates how technology can make virtual meetings more intelligent, accessible, and human-centered.

The proposed system successfully combines virtual meeting communication, 3D avatar representation, live transcription, meeting analysis, and AI-generated summaries into a single platform. By capturing important conversations and organizing them into meaningful information, the system reduces the need for users to watch lengthy meeting recordings or depend entirely on manually prepared notes. The interactive design also provides a convenient environment for participants to communicate, share information, and review meeting outcomes.

The project highlights the importance of intelligent virtual presence in today's increasingly digital working and learning environments. As remote collaboration and online meetings continue to grow, individuals may face difficulties attending every meeting because of scheduling conflicts, workload, or other commitments. The AI Twin addresses this challenge by providing a digital representation that can maintain meeting presence and help users remain informed about important discussions. Another significant contribution of the

project is demonstrating how AI can be combined with a 3D avatar to create a more natural and human-oriented digital experience.

Rather than functioning only as a recording or transcription tool, the AI Twin provides a recognizable virtual representation of the user while supporting automated meeting information capture and analysis. This approach creates a foundation for future systems where digital representations can assist individuals in maintaining continuity across virtual interactions. Although the proposed system achieves its intended objectives, there remains scope for future enhancements such as more realistic 3D avatar animation, voice-based interaction, multilingual speech recognition, emotion-aware responses, personalized AI behavior, advanced meeting participation, and integration with platforms such as Google Meet and Zoom.

Improvements in AI models, cloud infrastructure, privacy, and security can further strengthen the system's reliability and practical usability. Overall, the AI Twin – 3D Avatar for Automated Meeting Presence demonstrates the potential of combining Artificial Intelligence, digital avatars, and real-time communication to create a new form of virtual participation. The project establishes a strong foundation for future AI-powered digital presence systems and supports the vision of #Humanise — making technology more human, natural, connected, and meaningful.

V. FUTURE SCOPE

The AI Twin – 3D Avatar for Automated Meeting Presence offers significant potential for further enhancement and expansion in the future. While the current system provides a 3D avatar, automated meeting presence, live transcription, and AI-generated meeting summaries, future improvements can make the platform more intelligent, interactive, realistic, and human-centered. One major area of future scope is the integration of advanced Artificial Intelligence and deep learning models to improve the ability of the AI Twin to understand meeting conversations, recognize

context, and provide more accurate and meaningful responses.

Another important future enhancement is the development of more realistic and expressive 3D avatars. Future versions can include advanced facial expressions, lip synchronization, natural body movements, eye contact, gestures, and emotion-aware behavior.

These improvements can make the AI Twin appear more natural during meetings and create a stronger sense of virtual presence. The platform can also be extended by introducing voice-based AI interaction, allowing the AI Twin to respond to questions or participate in discussions based on predefined user preferences and authorized information. With advanced Natural Language Processing and speech generation, the AI Twin could understand questions directed toward the represented participant and provide appropriate responses while clearly identifying itself as an AI representation.

Future versions can include advanced meeting intelligence features such as automatic identification of important decisions, action items, deadlines, questions, and speaker-specific contributions. The system could also generate personalized meeting reports that highlight only the information relevant to each participant. This would transform the platform from a meeting transcription system into a more comprehensive intelligent meeting assistant. Another promising area of future scope is multilingual communication. Supporting multiple languages and real-time translation could allow participants from different linguistic backgrounds to communicate more effectively.

The AI Twin could translate conversations, generate multilingual transcripts, and provide summaries in the user's preferred language. The system can also be enhanced through integration with existing meeting platforms such as Google Meet, Zoom, and Microsoft Teams, allowing the AI Twin to participate in meetings across different communication environments. Integration with calendars, task-management applications, email, and collaboration platforms could further automate meeting scheduling, followup tasks, and reminders.

In addition, future improvements can focus on personalization and continuous learning.

The AI Twin could learn user preferences, communication style, frequently discussed topics, and meeting patterns while maintaining appropriate privacy controls. This could allow the system to provide more personalized summaries and support the user's workflow without attempting to imitate the person deceptively. Privacy and security will also remain important areas for future development. Strong encryption, access control, consent mechanisms, secure transcript storage, and clear AI identification can be implemented to protect sensitive meeting information.

These measures will be particularly important when the system is used in professional, educational, or organizational environments. From a broader perspective, the system can be scaled to support large meetings, multiple AI Twins, advanced cloud infrastructure, mobile access, and immersive environments such as AR and VR. Multiple AI Twins could potentially maintain digital presence for different users while participating in the same virtual environment under clearly defined permissions.

Overall, the future scope of the AI Twin – 3D Avatar for Automated Meeting Presence lies in evolving it from an automated virtual meeting representative into a comprehensive intelligent digital presence system. By combining advanced AI, realistic 3D avatars, natural communication, multilingual capabilities, meeting intelligence, and secure digital identity, the platform can support the vision of #Humanise — making technology more human, natural, connected, and meaningful.

REFERENCES

- Y. Huang, Z. Li, J. Gao, and S. Wang, "DreamWaltz: Make a Scene with Complex 3D Animatable Avatars," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12455–12464, 21st May, 2023. <https://arxiv.org/abs/2305.12529>
- [2] S. Aliakbarian, F. Saleh, L. Petersson, and C. Shen, "HMD-NeMo: Online 3D Avatar Motion Generation From Sparse Observations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 10, pp. 3756–3767, 22nd August, 2023. <https://arxiv.org/abs/2308.11261>
- [3] J. Zhang, L. Chen, Q. Liu, and Y. Ma, "MagicAvatar: Multimodal Avatar Generation and Animation," *ACM Transactions on Graphics (TOG)*, vol. 42, no. 6, pp. 1–13, 28th August 2023. <https://arxiv.org/abs/2308.14748>

- [4] N. Kolotouros, T. Alldieck, A. Zanfir, E. G. Bazavan, M. Fieraru, and C. Sminchisescu, "DreamHuman: Animatable 3D Avatars from Text," arXiv preprint arXiv:2306.09329, Jun. 15, 2023. Available: <https://arxiv.org/abs/2306.09329>
- [5] K. Shen, C. Guo, M. Kaufmann, J. J. Zarate, J. Valentin, J. Song, and O. Hilliges, "X-Avatar: Expressive Human Avatars," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16911–16921, 2023. Available: https://openaccess.thecvf.com/content/CVPR2023/html/Shen_XAvatar_Expressive_Human_Avatars_CVPR_2023_paper.html
- [6] J. C. Vázquez-Correa, S. Moreno-Acevedo, A. Gonzalez-Docasal, A. Lasarguren, J. López, E. Rodríguez, and A. Álvarez, "Real-Time SpeechDriven Avatar Animation by Predicting Facial Landmarks and Deformation Blendshapes," Proceedings of the 7th International Conference on Natural Language and Speech Processing (ICNLSP), pp. 109–118, Oct. 2024. Available: <https://aclanthology.org/2024.icnlsp-1.13/>
- [7] X. Ji, Z. Liao, L. Dong, Y. Tang, G. Li, and M. Mao, "3D Facial Animation Driven by SpeechVideo Dual-Modal Signals," Complex & Intelligent Systems, vol. 10, pp. 5951–5964, 2024. Available: <https://link.springer.com/article/10.1007/s40747-024-01481-5>
- [8] R. Wang, Y. Cao, K. Han, and K.-Y. K. Wong, "A Survey on 3D Human Avatar Modeling— From Reconstruction to Generation," arXiv preprint arXiv:2406.04253, Jun. 2024. Available: <https://arxiv.org/abs/2406.04253>
- [9] S. L. Metzger et al., "A High-Performance Neuroprosthesis for Speech Decoding and Avatar Control," Nature, vol. 620, no. 7976, pp. 1037–1046, Aug. 2023, doi: 10.1038/s41586-023-06443-4. Available: <https://www.nature.com/articles/s41586-02306443-4>
- [10] H. B. Barua, C. Sarkar, A. Anil Kumar, A. Pal, and B. Balamuralidhar, "I Can Attend a Meeting Too! Towards a Human-Like Telepresence Avatar Robot to Attend Meeting on Your Behalf," arXiv preprint arXiv:2006.15647, Jun. 2020. Available: <https://arxiv.org/abs/2006.15647>
- [11] B. Jones, Y. Zhang, P. Wong, and S. Rintel, "VROOM: Virtual Robot Overlay for Online Meetings," CHI 2020 Extended Abstracts, Apr. 2020. Available: <https://www.microsoft.com/enus/research/publication/vroom-virtual-robotoverlay-for-online-meetings/>
- [12] H. Nam, M. Sarvesh, S. Kang, W. Woo, and K. Kim, "Effects of AI-Powered Embodied Avatars on Communication Quality and Social Connection in Asynchronous Virtual Meetings," IEEE Transactions on Visualization and Computer Graphics, vol. 31, no. 11, pp. 10152–10162, 2025, 10.1109/TVCG.2025.3616761. doi: Available: <https://doi.org/10.1109/TVCG.2025.3616761>
- [13] P. Wang, Y. Wang, Y. Wang, M. Billingham, H. Yang, J. Zhao, X. Zhang, and Z. Liu, "Exploring the Effect of Avatar Embodiment in XR Telepresence: A Survey," Displays, 2026, doi: 10.1016/j.displa.2026.103490. Available: <https://doi.org/10.1016/j.displa.2026.103490>