Developing an AI-Powered Robotic Assistant for Interactive Video-Based Learning: Engineering Innovations and System Design

Chen Giladi, Sami Shamoon College of Engineering, Israel

The European Conference on Education 2024
Official Conference Proceedings

Abstract
This study presents an AI-powered robotic assistant's concept and high-level design to revolutionize classroom video-based learning. The proposed system utilizes advanced natural language processing (NLP) and computer vision techniques to potentially generate interactive multiple-choice questions from educational YouTube videos automatically. The envisioned robotic assistant would transcribe video content, segment it, and use language models to create questions projected onto students' desks, potentially creating an immersive and interactive learning experience. The system concept includes a 3D-printed face with a human-like appearance and lip-sync capabilities to enhance communication. Student interaction is proposed through innovative 'flip-flop' devices with ArUco markers, potentially enabling real-time collection and analysis of responses. This paper introduces the system architecture and discusses its potential applications in enhancing video-based learning experiences and reducing teacher workload.

Keywords: AI in Education, Robotic Teaching Assistant, Video-Based Learning, Educational Technology Integration

1. Introduction

Video-based learning has become increasingly prevalent in modern education, offering flexibility and access to a wide range of content (Lin et al., 2018). However, it often lacks the interactivity and personalization of traditional classroom settings. This challenge presents an opportunity for innovative solutions to bridge the gap between passive video consumption and active learning experiences (Teresa et al., 2023). Artificial Intelligence (AI) has shown tremendous potential in various educational applications, from personalized learning paths to automated grading systems (Chen et al., 2020). Integrating AI in education promises to enhance learning outcomes, increase engagement, and provide valuable insights into student performance (Sajja et al., 2023).

This paper introduces the concept of an AI-powered robotic assistant designed to transform video-based learning in classroom settings. By combining advanced natural language processing, computer vision, and robotics, the proposed system aims to create an interactive and immersive learning environment that adapts to students' needs and enhances their engagement with educational video content.

Figure 1: AI-powered robotic teaching assistant engaging with students in a classroom setting.

2. Methods

2.1 System Design and Overview

The proposed AI-powered robotic assistant integrates various technologies to enhance video-based learning (Al-Shaikh et al., 2024). The system architecture includes an AI engine for processing educational video content, transcribing speech, and generating relevant questions (Forkan et al., 2023). A robotic interface with human-like features serves as a visual focal point, potentially increasing student engagement (Wrede et al., 2006). The design incorporates a projection system to display questions and information directly onto students' desks, creating an immersive learning environment (Yu, 2023). Student interaction is facilitated through a novel response mechanism using 'flip-flop' devices with ArUco markers, enabling real-time collection and analysis of responses (Mahalingam et al., 2024). This system adapts content dynamically based on individual student needs, aiming to improve learning outcomes compared to traditional teaching methods (Sajja et al., 2023).

2.2 Key Components

The AI engine functions as the system's core, using natural language processing and machine learning algorithms to analyze video content, generate questions, and process student responses (Sunitha et al., 2023). This component learns and improves based on student interactions and educator feedback (Sajja et al., 2023). The robotic interface, featuring a 3D-printed face, enhances the learning experience by visually representing the AI assistant (Kalyan Raj et al., 2023). This interface could incorporate lip-sync capabilities to match educational video audio, creating engaging student interactions. The projection system creates an immersive learning environment by displaying questions and information directly onto students' desks, blending digital content with the physical classroom space (Yu, 2023). Using 'flip-flop' devices with ArUco markers, the student response mechanism collects real-time feedback from students (Mahalingam et al., 2024). This component enables quick student responses, promotes active participation, and provides data for the AI engine to adapt its content delivery.

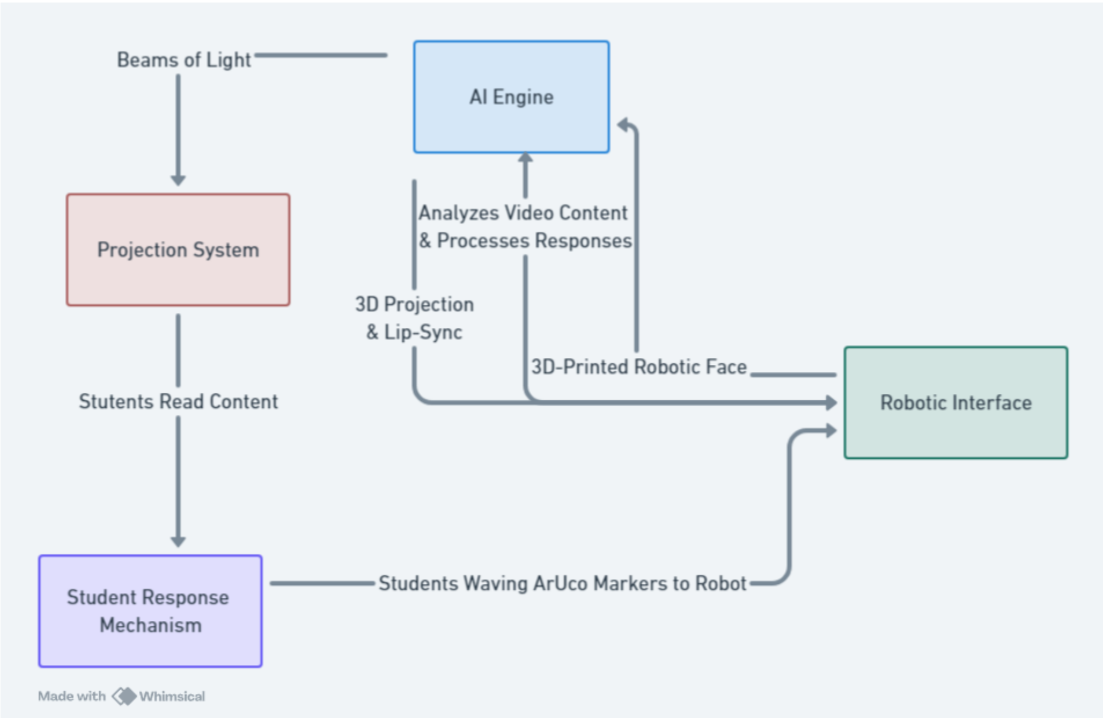

Figure 2: System Architecture. This diagram illustrates an advanced educational system where an AI Engine analyzes content and student responses, integrating with a 3D-printed robotic interface and a projection system that displays interactive content for student engagement through ArUco marker-equipped response devices.

3. Results and Discussion

3.1 Automated Question Generation

The proposed system's ability to automatically generate questions from educational videos represents a significant innovation in video-based learning. By leveraging natural language processing techniques, the AI engine analyzes video transcripts, identifies key concepts, and creates relevant multiple-choice, true/false, and short-answer questions (Kang et al., 2021; Bachiri & Mouncif, 2022). This feature saves educators significant time in preparing interactive elements for video lessons, with studies showing that AI-generated questions perform comparably to human-generated ones in assessing learner competency (Gala et al., 2021; Elshiny & Hamdy, 2023). The system enables the creation of a diverse range of questions, catering to different cognitive levels and learning objectives, with some approaches even incorporating pre-questions with images to improve learners' performance on comprehension tests (Skalban et al., 2012).

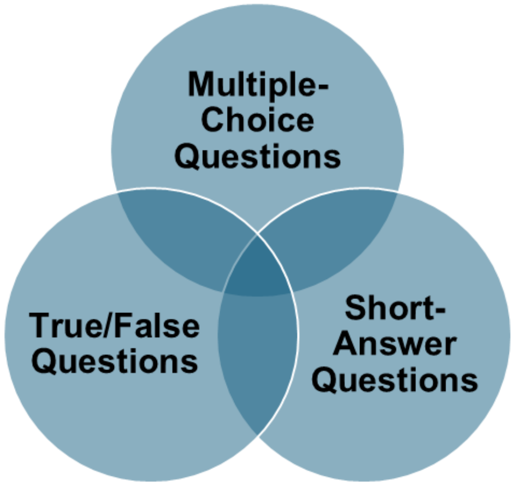

Figure 3: Types of Questions Generated: Multiple-choice (tests comprehension and recall, allows quick assessment), True/False (validates understanding, efficient for basic concepts), Short-Answer (encourages critical thinking, provides deeper insight into learner comprehension).

3.2 Immersive Learning Environment

Projecting questions and information directly onto students' desks creates an immersive learning environment, seamlessly integrating digital content into their physical space (Koike et al., 2000). This approach increases student engagement by enhancing interaction with learning materials and improving focus (Lin et al., 2014; Garrick et al., 2013). Presenting information in this novel way captures and maintains students' attention more effectively than traditional video playback methods, with studies showing improved task performance and user comfort (Wagner et al., 2018). The immersive nature of the projected content reduces distractions and fosters a more focused learning experience, particularly when information is displayed in the upper and nearby areas of the desk (Tokiwa & Fujinami, 2017).

Figure 4: Robotic assistant projecting content on the desk.

3.3 Real-Time Student Interaction

The proposed 'flip-flop' devices with ArUco markers represent an innovative approach to collecting student responses in real-time, similar to other student response systems that have shown promise in enhancing active learning and engagement (Mikic-Fonte et al., 2019; McLoone et al., 2015). This feature facilitates active participation from all students, including those hesitant to speak up in traditional settings, improving students' willingness to communicate and overall satisfaction with the learning experience (Hung, 2017). The real-time nature of this interaction enables the system to adapt its content delivery based on student responses, akin to real-time formative assessment tools like InkSurvey and Quiz It (Kowalski et al., 2013; Adam et al., 2014). Suppose many students struggle with a particular concept. In that case, the system can provide additional explanations or adjust the difficulty of subsequent questions, allowing instructors to gauge student understanding instantly and adapt their teaching accordingly (Mattei & Ennis, 2014).

Figure 5: Enhancing Classroom Engagement with Real-Time Response Devices. The image depicts students using 'flip-flop' devices with ArUco markers. This innovative approach improves engagement and learning by allowing real-time, active participation and immediate adaptation of teaching strategies based on student responses.

3.4 Potential Impact and Future Directions

The proposed AI-powered robotic assistant can significantly impact classroom-based learning by increasing interactivity and engagement, potentially enhancing student learning outcomes and making video lessons more effective (Suntharalingam, 2024; Kalyan Raj et al., 2023). The system could reduce the workload for educators by automating question generation and response analysis, freeing up time for more personalized instruction and support (Forkan et al., 2023; Kang et al., 2021). However, implementing such a system faces challenges, including ensuring the accuracy and relevance of generated questions, addressing privacy concerns related to data collection, and integrating the technology seamlessly into existing classroom environments (Kokku et al., 2018). Future research directions could explore the system's effectiveness in different subject areas and age groups and investigate its long-term impact on student learning outcomes and motivation (Seo et al., 2020; Lin et al., 2018).

4. Conclusion

The concept of an AI-powered robotic assistant for video-based learning represents an innovative approach to addressing the challenges of engagement and interactivity in digital education. By combining advanced AI technologies with a physical robotic presence and immersive projection systems, the proposed system has the potential to transform how students interact with educational video content.

While the implementation and effectiveness of such a system remain to be fully explored, the concept opens exciting possibilities for the future of classroom technology. Further research and development in this area could lead to significant advancements in educational technology, potentially revolutionizing how we approach video-based learning in classroom settings.

Acknowledgments

The author sincerely thanks the participating school, teachers, and students for their cooperation and invaluable feedback during this feasibility study—special thanks to the engineering team at Sami Shamoon College of Engineering for their technical support and expertise. The author acknowledges the Ministry of Education for facilitating and funding this research, which was instrumental in realizing this project.

References

Adam, D., Kioutsiouki, D., Karakostas, A., & Demetriadis, S. (2014). Do Your Students Get It? Quiz It! The Android Classroom Response System. In 2014 IEEE 14th International Conference on Advanced Learning Technologies (pp. 168-170). IEEE.

Al-Shaikh, R., Al-Malki, N., & Almasre, M. (2024). Implementing the cognitive theory of multimedia learning in designing and evaluating an AI educational video assistant utilizing large language models. Heliyon, 10(1), e25361.

Bachiri, Y., & Mouncif, H. (2022). Increasing Student Engagement in Lessons and Assessing MOOC Participants Through Artificial Intelligence. In Business Information Systems (pp. 128-142). Springer, Cham.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. IEEE Access, 8, 75264-75278.

Elshiny, R. M., & Hamdy, A. (2023). Automatic Question Generation Using Natural Language Processing and Transformers. In 2023 International Conference on Computing Advancements (ICCA) (pp. 1–6). IEEE.

Forkan, A., Kang, Y., Jayaraman, P., Du, H., Thomson, S., Kollias, E., & Wieland, N. (2023). VideoDL: Video-Based Digital Learning Framework Using AI Question Generation and Answer Assessment. International Journal of Advanced Corporate Learning, 16(1), 35–47.

Gala, R., Vijayaraghavan, R., Nikam, V., & Kiwelekar, A. (2021). Real-Time Cognitive Evaluation of Online Learners through Automatically Generated Questions. In 2021 IEEE 20th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC) (pp. 166-173). IEEE.

Garrick, R., Villasmil, L., Dell, E., & Hart, R. (2013). Creating Technology-Rich Learning Environments for the Classroom. In Technological Advancement in Developed and Developing Countries: Discoveries in Global Information Management (pp. 302–316). IGI Global.

Hung, H. T. (2017). Language Teaching and Technology Forum: Integrating a Student Response System in Flipped Classrooms. TESOL Quarterly, 51(4), 988–998.

Kalyan Raj, L., Bijapur, C. S., & Gururaja, H. S. (2023). Digital Assistant Robot for Academic Fraternity Using Deep Learning. In 2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS) (pp. 1-6). IEEE.

Kang, Y. B., Forkan, A., Jayaraman, P., Wieland, N., Kollias, E., Du, H., Thomson, S., & Li, Y. F. (2021). An AI-based Solution for Enhancing Delivery of Digital Learning for Future Teachers. arXiv preprint arXiv:2106.04719.

Koike, H., Sato, Y., Kobayashi, Y., Tobita, H., & Kobayashi, M. (2000). Interactive textbook and interactive Venn diagram: natural and intuitive interfaces on augmented desk system. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 121–128). ACM.

Kokku, R., Sundararajan, S., Dey, P., Sindhgatta, R., Nitta, S. V., & Sengupta, B. (2018). Augmenting Classrooms with AI for Personalized Education. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6976-6980). IEEE.

Kowalski, F. V., Kowalski, S. E., & Gardner, T. Q. (2013). Using Mixed Mobile Computing Devices for Real-Time Formative Assessment. In 2013 ASEE Annual Conference & Exposition (pp. 23–1186). ASEE.

Lin, P. H., Wooders, A., Wang, J. T. Y., & Yuan, W. M. (2018). Artificial Intelligence, the Missing Piece of Online Education? IEEE Engineering Management Review, 46(3), 25–28.

Lin, T. Y., Chen, C. F., Huang, D., Huang, C. W., & Chen, G. D. (2014). Using Resource of Classroom and Content of Textbook to Build Immersive Interactive Learning Playground. In 2014 IEEE 14th International Conference on Advanced Learning Technologies (pp. 669–670). IEEE.

Mahalingam, T., Nishanth, S., Sanjay, T., & Sathya, K. (2024). AI-Assisted Studying Partner for Students. International Journal of Advanced Research, 12(1), 670–675.

Mattei, M. D., & Ennis, E. (2014). Continuous, Real-Time Assessment of Every Student's Progress in the Flipped Higher Education Classroom Using Nearpod. Journal of Learning in Higher Education, 10(1), 1-7.

McLoone, S. C., Villing, R., & O'Keeffe, S. (2015). A Novel Smart Device Student Response System for Supporting High-Quality Active Learning in the Engineering and Science Disciplines. AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education, 7(2), 2391–23919.

Mikic-Fonte, F. A., Liz-Domínguez, M., Caeiro-Rodríguez, M., & Llamas-Nistal, M. (2019). Work-in-Progress: QMSFC: A System for Encouraging the Active Participation of Students in Flipped Classroom. In 2019 IEEE Global Engineering Education Conference (EDUCON) (pp. 913-917). IEEE.

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2023). Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education. arXiv preprint arXiv:2309.10892.

Seo, K., Fels, S., Yoon, D., Roll, I., Dodson, S., & Fong, M. (2020). Artificial Intelligence for Video-based Learning at Scale. In Proceedings of the Seventh ACM Conference on Learning @ Scale (pp. 215–218). ACM.

Skalban, Y., Ha, L., Specia, L., & Mitkov, R. (2012). Automatic Question Generation in Multimedia-Based Learning. In Proceedings of COLING 2012: Posters (pp. 1151-1160).

Sunitha, M., Vijitha, B., & Gunavardhan, E. (2023). Artificial Intelligence-based Smart Education System. In 2023 International Conference on Electronic Systems, Signal Processing and Computing Technologies (ICESC) (pp. 1–6). IEEE.

Suntharalingam, H. (2024). Enhancing Digital Learning Outcomes Through the Application of Artificial Intelligence: A Comprehensive Review. International Journal of Innovative Science and Research Technology, 9(4), 530-539.

Teresa, L. A., Sunil, N. M., Andrews, S. R., Thengumpallil, T. T., Thomas, S., & Binson, V. A. (2023). Enhancing Children's Learning Experience: Interactive and Personalized Video Learning with AI Technology. In 2023 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE) (pp. 1–6). IEEE.

Tokiwa, M., & Fujinami, K. (2017). Understanding Where to Project Information on the Desk for Supporting Work with Paper and Pen. In Proceedings of the XVIII International Conference on Human-Computer Interaction (pp. 1-4).

Wagner, J., Freitas, C., & Nedel, L. (2018). VirtualDesk: A Comfortable and Efficient Immersive Information Visualization Approach. Computer Graphics Forum, 37(3), 415-426.

Wrede, B., Kleinehagenbrock, M., & Fritsch, J. (2006). Towards an Integrated Robotic System for Interactive Learning in a Social Context. In 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 5295-5300). IEEE.

Yu, D. (2023). AI-Empowered Metaverse Learning Simulation Technology Application. In 2023 International Conference on Metaverse, Extended Reality and Artificial Intelligence (iMeta) (pp. 178–181). IEEE.

Contact email: chengi1@sce.ac.il