2026 Volume: 02 Issue: 05 www.ijrpa.com ISSN 2456-9995 Review Article

International Journal Research Publication Analysis

Page: 01-09

AI-BASED INTELLIGENT LEARNING FRAMEWORK FOR INTERACTIVE EXPLANATION AND ADAPTIVE CONTENT DELIVERY

***Varsha Ravi**

India.

Article Received: 08 April 2026
Article Revised: 28 April 2026
Published on: 18 May 2026

***Corresponding Author: Varsha Ravi**
India.

DOI: https://doi-doi.org/101555/ijrpa.4325

ABSTRACT

The advancement of digital education has led to the emergence of AI-based tutoring systems, e-learning platforms, and virtual assistants. However, most existing solutions lack integration of interactive, multimodal, and personalized learning approaches. This survey reviews recent developments in AI-powered audio and video tutoring systems that enable interactive explanations and intelligent content delivery.

The paper analyzes techniques based on Natural Language Processing (NLP), speech recognition, and machine learning, highlighting their role in real-time interaction and adaptive learning. It also identifies key challenges such as limited personalization, lack of contextual understanding, and insufficient integration of audio-visual learning methods. Finally, the survey emphasizes the need for unified systems that combine voice interaction, video explanation, and AI-driven personalization to improve learning effectiveness.

KEYWORDS: Artificial Intelligence; Survey; Interactive Learning; NLP; Audio-Video Tutoring.

1. INTRODUCTION

The rapid advancement of digital technologies has significantly transformed the education sector, leading to the widespread adoption of e-learning platforms, virtual classrooms, and online tutoring systems. These systems provide easy access to educational content; however, most of them rely on static resources such as recorded videos, textual materials, and predefined assessments. As a result, learning often becomes a one-way process with limited interaction, reduced engagement, and lack of personalization.

In recent years, Artificial Intelligence (AI) has emerged as a promising solution to address these limitations. AI-based systems, including intelligent tutoring systems and virtual assistants, enable automated responses, adaptive learning paths, and performance tracking. Despite these advancements, many existing solutions focus primarily on either text-based interaction or standalone video content, without effectively integrating multiple modes of learning such as audio, video, and real-time interaction.

Interactive learning plays a crucial role in improving knowledge retention and student engagement. A system that can understand user queries, provide real-time explanations, and adapt content based on individual learning needs can significantly enhance the learning experience. In this context, the integration of Natural Language Processing (NLP), speech recognition, and machine learning enables the development of intelligent tutoring systems capable of simulating human-like teaching behavior.

This survey focuses on AI-powered audio and video tutoring systems designed for interactive explanation and intelligent content delivery.

It reviews existing approaches, analyzes their strengths and limitations, and identifies key challenges such as lack of contextual understanding, insufficient personalization, and limited multimodal integration. Furthermore, the survey highlights the need for a unified framework that combines voice interaction, video-based explanation, and AI-driven adaptability to create a more effective and engaging learning environment.

2. Literature Survey

Recent advancements in digital education have led to significant research in video-based learning, AI tutoring systems, virtual agents, and multimodal AI technologies. This section reviews key contributions relevant to AI-powered audio-video tutoring systems

2.1 Paper 1: Video-Based Learning Systems

Author: Evelyn Navarrete, Anett Hoppe, Ralph Ewerth

This paper presents a comprehensive review of video-based learning (VBL) technologies and their impact on modern education. It highlights the increasing use of online video platforms for learning and analyzes video features, interaction methods, and technologies used in educational videos.

Advantages:

- Enhances understanding through visual representation of concepts
- Improves learner engagement and motivation
- Provides wide accessibility to educational content

Limitations:

- Mostly passive learning with limited interaction
- Lack of personalization and adaptive learning
- Difficulty in selecting relevant content due volume
- No real-time doubt clarification

2.2 Paper 2: State of Video-Based Learning

Author: Ahmed Mohamed Fahmy Yousef, Mohamed Amine Chatti, Ulrik Schroeder

This paper reviews the evolution and effectiveness of video-based learning systems, including MOOCs and flipped classroom models. It analyzes the role of video in improving learning outcomes and teaching methodologies.

Advantages:

- Improves learning outcomes through interactive video content
- Supports blended and online learning environments
- Enhances student engagement and collaboration
- Enables flexible learning anytime and anywhere

Limitations:

- Learners remain passive without interaction tools
- Limited adaptability to individual learning needs
- Lack of intelligent feedback mechanisms
- Does not support personalized tutoring

2.3 Paper 3: Voice-Based Intelligent Virtual Agents

Author: Eirene Katsarou et al.

This paper reviews the role of AI-powered voice-based intelligent virtual agents in education. It focuses on conversational interaction, adaptive learning, and the use of NLP in educational environments.

Advantages:

- Enables natural human-computer interaction
- Supports real-time communication and feedback
- Provides personalized and adaptive learning
- Enhances engagement through conversational learning

Limitations:

- Limited contextual understanding in complex queries
- Lack of visual explanation support
- Requires high-quality speech recognition systems
- Challenges in handling diverse user inputs

2.4 Paper 4: AI-Based Intelligent Tutoring Systems

Author: Meriem Zerkouk, Miloud Mihoubi, Belkacem Chikhaoui

This paper provides a comprehensive review of AI-based Intelligent Tutoring Systems (ITS), focusing on personalization, adaptive learning, and machine learning integration in education.

Advantages:

- Provides personalized learning experience
- Offers real-time feedback and performance tracking
- Adapts content based on student behavior
- Improves learning outcomes significantly

Limitations:

- High complexity in implementation
- Limited multimodal interaction (mostly text-based)
- Requires large datasets for training
- Challenges in real-world deployment

2.5 Paper 5: Multimodal and Generative AI in Education Author: Ville Heilala, Roberto Araya, Raija Hämäläinen

This paper explores the role of multimodal and generative AI technologies in education, including text-to-speech, text-to-video, and content generation techniques.

Advantages:

- Supports multiple learning modalities (audio, video, text)
- Enhances interactive and engaging learning
- Enables automated content generation
- Improves accessibility for diverse learners

Limitations:

- Limited research on multimodal integration
- Most systems focus only on text-based AI

2.6 Paper 6: AI Conversational Tutors Author: Nikolaos Avouris

This paper evaluates AI conversational tutors used in language learning, focusing on interaction quality, personalization, and real-time feedback.

Advantages:

- Provides interactive conversational learning
- Enables real-time feedback and correction
- Simulates human-like tutoring experience
- Improves communication and language skills

3. Proposed System

The proposed system is an AI-powered audio and video tutor system designed to provide interactive explanations and intelligent content delivery. The system combines voice interaction, video-based learning, and artificial intelligence to create a personalized and engaging learning experience. Unlike traditional e-learning platforms, the system enables real-time interaction, allowing users to ask questions and receive adaptive explanations in both audio and visual formats. The system acts as a virtual tutor, capable of understanding user queries and delivering content based on individual learning needs.

3.1 System Architecture

Each module is interconnected, forming a complete workflow for intelligent tutoring and content delivery. The architecture of the proposed system consists of the following main modules:

1. User Interaction Module
2. AI Processing Engine

3. Content Generation Module
4. Personalization Module
5. Output and Visualization Interface

3.3 Module Description:

- **User Interaction Module:**

This module allows users to interact with the system Voice input (speech-based queries)
Text input (typed questions)
Speech input is converted into text using speech recognition techniques, enabling the system to process user queries efficiently.

- **AI Processing Engine**

This module is responsible for understanding and analyzing user queries using:
Natural Language Processing (NLP) Machine Learning algorithms
It identifies:

- User intent
- Topic of the query
- Difficulty level

Based on this analysis, the system generates.

- **Content Generation Module**

This module generates learning content in multiple formats:

- Audio explanations using Text-to-Speech (TTS)
- Video explanations using animations, slides, or visual aids The system ensures that explanations are clear, structured, and easy to understand.

- **Personalization Module:**

This module tracks user behavior and learning patterns, including:

- Previous queries
- Performance level
- Learning speed

Based on this data, the system adapts:

- Explanation complexity
- Content format

- Learning recommendations

- **Output and Visualization Interface:**

The final output is delivered through an interactive interface that includes:

- Audio explanation (voice output)
- Video/visual content
- Text summary

Users can interact further by asking follow-up questions, making the system fully interactive.

3.4 System Working Flow:

The working of the system follows these steps:

1 User provides input (voice or text)

2 Speech is converted to text (if voice input)

3 AI engine processes the query

4 Relevant content is generated

5 Audio and video explanations are delivered

6 User receives output and can interact further

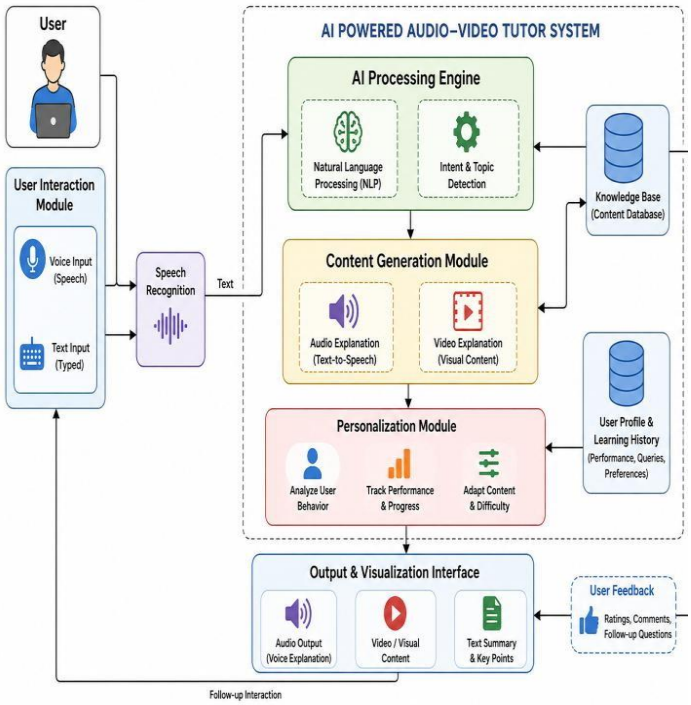

Fig. 1. Architecture of the Proposed AI-Powered Audio–Video Tutor System

6.3 Advantages of the Proposed System:

- Provides real-time interactive learning

- Combines audio, video, and artificial intelligence technologies
- Adapts to individual learning needs
- Enhances user engagement and understanding
- Reduces dependency on human tutors

4. CONCLUSION

The rapid advancement of artificial intelligence and digital technologies has significantly transformed modern education systems. This paper reviewed various approaches, including video-based learning, intelligent tutoring systems, voice-based virtual agents, conversational tutors, and multimodal AI technologies. While these systems offer valuable features such as accessibility, personalization, and interactive learning, they often operate independently and lack full integration of multimodal capabilities.

The proposed AI-powered audio-video tutor system addresses these limitations by combining audio interaction, video-based explanation, and intelligent content delivery into a unified framework. The system enables real-time interaction, adapts to individual learning needs, and enhances user engagement through personalized explanations. By integrating multiple technologies, it provides a more effective and interactive learning experience compared to traditional methods.

Overall, the proposed approach contributes to the development of next-generation educational systems that simulate human-like teaching behavior. It offers a scalable and efficient solution for modern digital learning environments, with the potential to improve knowledge retention, accessibility, and learning outcomes.

REFERENCES

1 E. Navarrete, A. Hoppe, and R. Ewerth, “A Review on Recent Advances in Video-based Learning Research: Video Features, Interaction, Tools, and Technologies,” Proceedings of CIKM Workshops, 2021.

2 A. M. F. Yousef, M. A. Chatti, and U. Schroeder, “The State of Video-Based Learning: A Review and Future Perspectives,” International Journal on Advances in Life Sciences, 2014.

3 E. Katsarou et al., “A Systematic Review of Voice-based Intelligent Virtual Agents in EFL Education,” iJET, vol. 18, no. 10, 2023.

4 V. Heilala, R. Araya, and R. Hämäläinen, “Beyond Text-to-Text: An Overview of Multimodal and Generative Artificial Intelligence for Education Using Topic Modeling,”

ACM SAC, 2025.

5 N. Avouris, “AI Conversational Tutors in Foreign Language Learning: A Mixed-Methods Evaluation Study,” Proceedings of iCT in Education, 2025.

6 G. Kestin et al., “AI Tutoring Outperforms In-Class Active Learning,” *Nature Scientific Reports*, 2025.

7 A. M. Vieriu, “The Impact of Artificial Intelligence on Students’ Learning Outcomes and Engagement,” *Education Sciences*, 2025.

8 T. Kabudi, I. Pappas, and D. Olsen, “AI-Enabled Adaptive Learning Systems: A Systematic Review,” *Computers and Education: Artificial Intelligence*, 2021.

9 J. Du et al., “A Systematic Review of AI-Powered Chatbots for Language Learning,” *Computers and Education: Artificial Intelligence*, 2024.

10 B. Klimova et al., “Exploring the Effects of Artificial Intelligence on Higher Education Learning,” *Education and Information Technologies*, 2025.