

AI-Based Video Chaptering and Learning Support System

Atharva A Digambar¹, Jasshan S Bhalgat², Aditya N Adak³, Riyaz Jamadar⁴

Students, Artificial Intelligence & Data Science¹⁻³

HOD & Guide, Artificial Intelligence & Data Science⁴

AISSMS Institute of Information Technology, Pune, India

Abstract: *With the explosion of educational videos on the Internet, ranging from lectures to webinars, it has become practically impossible and very inefficient to review video content manually. Artificial intelligence, therefore, has a lot to promise in automating this process via advanced multimodal understanding of video content. The work surveys some current AI-driven approaches which include Automatic Speech Recognition, Computer Vision, and Large Language Models for transforming long videos into structured chapter-wise summaries, interactive quizzes, and searchable transcripts. It also highlights the limitations of unimodal systems operating only on speech data and offers exciting possibilities due to multimodal learning on improving educational accessibility, efficiency, and engagement. This paper reviews recent AI-driven techniques for converting lengthy educational videos into chapter-level summaries, interactive assessments, and transcripts that are easily searchable, thus making educational content more accessible and engaging. Whereas traditional unimodal approaches operating on speech data alone allow limited insight, richer context is captured when multiple streams of information are integrated through multimodal learning. Key challenges such as data quality, model interpretability, and scalability are discussed; at the same time, the paper underlines the transformative potential of AI for improving the efficiency, inclusivity, and personalization of digital education.*

Keywords: Automatic Speech Recognition, Video Summarization, Image Captioning, Multimodal AI, Educational Technology, Quiz Generation, Large Language Models

