



Archives available at journals.mriindia.com

International Journal on Advanced Computer Theory and Engineering

ISSN: 2319-2526

Volume 14 Issue 01, 2025

Video to Text Summarizer with Highlights and Quiz Generator

¹Mr. Jalindar N. Ekatpure, ²Jagadale Samiksha, ³Kargal Snehal, ⁴Lavand Sakshi, ⁵Yadav Sanika

^{1 2 3 4 5}Department of Computer Engineering

S. B. Patil College of Engineering Indapur, Pune, India

Email: j.ekatpure@gmail.com¹, jagadalesamiksha0605@gmail.com², snehalkargal0@gmail.com³, sakshilawand98@gmail.com⁴, sanikayadav926@gmail.com⁵

Peer Review Information	Abstract
<p><i>Submission: 11 Sept 2025</i></p> <p><i>Revision: 10 Oct 2025</i></p> <p><i>Acceptance: 22 Oct 2025</i></p> <p>Keywords</p> <p><i>Video Summarization, Natural Language Processing (NLP), Machine Learning (ML), Automatic Speech Recognition, Text Summarization, Quiz Generation, Educational Technology (EdTech), Multimodal Fusion, Cloud Integration, Deep Learning.</i></p>	<p>The exponential growth of educational, training, and professional video content has created challenges in effectively extracting and retaining critical information. Manually watching long videos is often time-consuming and inefficient, leading to reduced learner engagement and knowledge retention. This project proposes the development of an intelligent system that leverages Natural Language Processing (NLP) and Machine Learning (ML) techniques to automatically generate concise text summaries and context-based quizzes from video content. The system integrates video processing for audio-visual extraction, speech-to-text conversion for transcription, and advanced NLP models for summarization and quiz generation. Additionally, cloud integration ensures scalability, real-time processing, and accessibility across devices. The proposed solution aims to save time, enhance comprehension, and improve learning outcomes for students, educators, and professionals.</p>

Introduction

In recent years, digital learning and online education platforms have witnessed exponential growth, resulting in an enormous increase in video-based content. Lectures, tutorials, and training sessions are widely recorded and shared, but the challenge lies in extracting meaningful insights efficiently from lengthy videos.

Traditional manual methods, where learners watch entire videos, often lead to wasted time and reduced engagement. Learners may find it difficult to navigate to the most relevant sections or recall important information effectively. Hence, intelligent solutions that summarize content automatically are increasingly important.

Advances in Artificial Intelligence (AI) and Machine Learning (ML), particularly in Natural Language Processing (NLP) and Computer Vision, provide a foundation for developing such intelligent systems. Techniques like speech-to-text conversion, semantic analysis, and deep learning models enable automated summarization and highlight detection.

The integration of automatic summarization with quiz generation adds significant value to learning. Quizzes enhance comprehension, reinforce retention, and encourage active learning. Generating context-aware, non-trivial questions from summaries ensures that learners engage critically with the content.

This project, Video to Text Summarizer with Highlights and Quiz Generator, aims to address

these challenges by creating an intelligent, AI-driven system. It combines video and audio processing, NLP-based summarization, highlight extraction, and automated question generation, ultimately improving learning efficiency, accessibility, and scalability across platforms.

Literature Survey

1. Video Summarization Techniques: A Comprehensive Review – Toqa Alaa et al. (2024): Explores deep learning-based multimodal approaches; future scope in scalable, real-time summarization.

2. Video-to-Text Summarization using NLP – Prerna Mishra et al. (2023): Uses ASR and SpaCy entity extraction; emphasizes multilingual summarization challenges.

3. Video Transcript Summarization Using BERT – Iswarya M. et al. (2023): Applies BERT for large transcript summarization; reinforcement learning suggested for improvement.

4. Video Transcript Summarizer – Ilampiray P. et al. (2023): Uses YouTube Transcript API + BERT for multilingual summarization with Flask integration.

5. Video Summarization using NLP – Raghav Malu et al. (2023): Combines transcript extraction, transformer models, and ROUGE evaluation for education/surveillance.

6. Video to Text Summarisation and Timestamp Generation – Dhiraj Shah et al. (2022): Uses CNN + GRU with attention for event detection; useful in long video navigation.

7. Automatic Generation of MCQs from Educational Texts in Hindi – Shweta Yadav et al. (2021): Employs BERT embeddings for question generation in Hindi; suggests expansion to other Indian languages.

8. Video Summarization Using Highlight Detection and Deep Learning – Sarvesh Kolhe et al. (2020): Uses CNN + LSTM for highlight detection; scope in real-time summarization.

9. Video Summarization using Deep Semantic Features – Mayu Otani et al. (2016): Employs DNN to capture semantic meaning in videos; suggests improvements in segmentation and retrieval.

10. Other related NLP/EdTech works (from survey list): Focus on multimodal learning, abstraction accuracy, and e-learning integration.

Research Gap

- Existing methods struggle with real-time summarization of long videos.
- Multilingual support is limited in current summarization systems.
- Quiz generation research is underdeveloped, especially for non-English languages.

- Many approaches lack cloud integration and scalability.
- Summarization models often face hallucination and coherence issues.
- Limited focus on user engagement and personalization in generated outputs.

Problem Statement

With the increasing volume of video content in education, training, and professional fields, users face challenges in efficiently extracting key information. Manually watching long videos is time-consuming and ineffective. There is a need for an automated system that converts video content into concise text summaries, highlights important parts, and generates quizzes to enhance learning outcomes.

Conclusion

The project addresses the growing challenge of extracting and retaining knowledge from video content in education and training. By integrating speech-to-text conversion, NLP-based summarization, highlight detection, and automatic quiz generation, the system provides a holistic solution that enhances comprehension and engagement. Unlike existing models, this approach emphasizes scalability, real-time accessibility, and the inclusion of quizzes for active learning. Thus, it fills critical research gaps and contributes to the field of AI-powered EdTech systems.

References

- Alaa, T., Mongy, A., Bakr, A., Diab, M., & Gomaa, W. (2024). *Video Summarization Techniques: A Comprehensive Review*.
- Mishra, P., Garg, K., & Rathi, N. (2023). *Video-to-Text Summarization using NLP*.
- Iswarya, M., Krishna, P. S., Naveen, K., Ganesh, M., & Yasin, M. (2023). *Video Transcript Summarization Using BERT*.
- Ilampiray, P., Raju, N. D., Thilagavathy, A., et al. (2023). *Video Transcript Summarizer*.
- Malu, R., Andhale, S., Potdar, V., & Khatavkar, T. S. (2023). *Video Summarization using NLP*.
- Shah, D., Namdev, U., Dedhia, M., Kanani, P., & Desai, R. (2022). *Video to Text Summarisation and Timestamp Generation*.
- Yadav, S., Ekbal, A., & Bhattacharyya, P. (2021). *Automatic Generation of MCQs from Educational Texts in Hindi*.

Kolhe, S., Sharma, R., & Pise, S. (2020). *Video Summarization Using Highlight Detection and Deep Learning*.

Otani, M., Nakashima, Y., Rahtu, E., Heikkilä, J., & Yokoya, N. (2016). *Video Summarization using Deep Semantic Features*.

Lin, C. Y. (2004). *ROUGE: A Package for Automatic Evaluation of Summaries*.

Vaswani, A., et al. (2017). *Attention Is All You Need*. NeurIPS.

Devlin, J., et al. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. NAACL.

Zhang, J., Zhao, Y., & LeCun, Y. (2020). *Multimodal Fusion Techniques for Video Understanding*. IEEE TPAMI.

Gong, Y., & Liu, X. (2001). *Generic Text Summarization Using Relevance Measures*. SIGIR.

See, A., Liu, P. J., & Manning, C. D. (2017). *Get to the Point: Abstractive Summarization with Pointer-Generator Networks*. ACL.

Ekatpure, J. N., Tavate, C. S., Malshikare, S. S., Khomane, A. B., & Tamboli, M. J. L. (2025). Artificial intelligence based virtual keyboard and mouse for computer. *International Journal on Advanced Computer Theory and Engineering*, 14(1), 449–456.

Ekatpure, J. N., Mohite, S. D., Shinde, A. A., Shirkande, N. B., & Upase, V. V. (2025). Campus recruitment system using machine learning. *International Journal on Advanced Computer Theory and Engineering*, 14(1), 427–432

Aware, D. B., Sayyad, S. R., Shaikh, A. H., Thombare, S. B., & Ekatpure, J. N. (2025). Translation Assistant for Converting Sign Language to Text and Audio. *International Journal on Advanced Computer Engineering and Communication Technology*, 14(1), 445–449.

Ekatpure, J. N., Aware, D. B., Shaikh, A. H., Sayyad, S. R., & Thombare, S. B. (2024). A comprehensive survey on sign language translation systems: Bridging gestures, text, and audio for enhanced communication. *International Journal of Recent Advances in Engineering and Technology*, 13(2), 15–21.

Ekatpure, J. N., Tavate, C., Malshikare, S.,

Khomane, A., & Tamboli, M. J. (2024). Advancements in AI-powered virtual keyboards and mice: A survey of cutting-edge technologies for modern computing. *International Journal on Advanced Computer Theory and Engineering*, 13(2), 52–57.