International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

YouTube Summarization

Dr.V. Indu Nair, Guide, SNS College of Engineering, Coimbatore, India
Dr.P. Sumathi, Head of Department, SNS College of Engineering, Coimbatore, India
Krithika M, Student, SNS College of Engineering, Coimbatore, India
SriVignesh A, Student, SNS College of Engineering, Coimbatore, India
Sudharsan S, Student, SNS College of Engineering, Coimbatore,India
Vijith Kumar V, Student, SNS College of Engineering, Coimbatore, India

Abstract The YouTube Summarizer is an AI-powered web application designed to enhance digital content consumption by providing concise summaries of YouTube videos. Developed using the Next.js framework, the platform integrates state-of-the-art language models such as GPT, Gemini, and LLaMA to generate context-aware summaries from extracted video transcripts. It supports multilingual outputs and offers summary customization—like video-style or podcast-style formats—tailored to user preferences. The application features a sleek, responsive UI with session history tracking, making it accessible and efficient for students, researchers, and content creators aiming to quickly grasp video content without full viewing. Keywords — YouTube Summarizer, Natural Language Processing, GPT, Gemini, LLaMA, Next.js, AI-based Tool, Multilingual Summarization. This research presents an AI-powered web application that addresses the growing challenge of processing lengthy YouTube videos by generating customizable, context-aware summaries. The tool leverages cutting-edge transformer-based language models (GPT, Gemini, and LLaMA), capable of multilingual processing and summary style adjustments. Key parameters such as transcript quality, model selection, and summarization goals are evaluated and optimized. This approach enhances accessibility, efficiency, and content comprehension for diverse user groups including students, researchers, and educators.

I. INTRODUCTION

Video content has become a primary source of learning and entertainment, but watching long-form videos can be time-consuming. Many students, educators, and professionals face challenges in extracting relevant information from these videos efficiently. The YouTube Summarizer addresses this by offering an intelligent platform that automates the summarization process using modern AI technologies. By integrating transcript extraction and AI-powered summarization models, the tool delivers quick, high-quality insights in seconds, thereby saving time and improving user productivity.

II. EXISTING SYSTEM

Existing video summarization tools are often limited in functionality, typically offering only timestamp-based clipping or keyword extraction. Most platforms do not integrate powerful AI language models or provide customization options for summary formats. Furthermore, they lack multilingual support, real-time processing, and user-friendly interfaces. Some tools rely solely on static text output without adapting to user intent or style. These limitations restrict the effectiveness and scalability of traditional summarizers in educational and professional settings.

III. IDEATION

The idea for the YouTube Summarizer emerged from the growing demand to consume educational and informational video content more efficiently. With the rise in remote learning, online courses, and content-heavy tutorials, users often spend hours viewing full-length videos just to extract key points. The project aimed to address this issue by:

- Enabling automatic transcript extraction from YouTube videos.
- Using state-of-the-art AI models to generate context-aware summaries.
- Ensure Providing user-selectable summary styles (e.g., video or podcast format).
- Supporting multilingual summaries (English, German) to cater to a wider audience.
- Creating an intuitive interface with history tracking and model selection options.

Innovative Features Include:

- **Model Flexibility –** Choose between Gemini, GPT, or LLaMA for summarization. NLP Translation to map text to ISL gestures.
- **Transcript Automation** – Fetches and cleans transcripts without manual steps. Google Image Integration to support visual understanding.
- **Format Customization** – Users can switch between summary styles based on need.
- **Multilingual Output –** Generates summaries in English or German.

- **Session History –** Remembers previous summarizations for future reference.

Technological Brainstorming:

1. **Natural Language Processing (NLP):**

NLP is used to understand and convert raw transcripts into coherent summaries. Tokenization, context mapping, and summarization algorithms are applied through AI models like GPT and Gemini to extract meaningful content from lengthy scripts.

2. **Model Switching Logic:**

The platform allows dynamic selection among **International Journal for Research in Engineering Application & Management (IJREAM)** 2**Volume 1, Issue 1, 04 2025.**

multiple AI models. Each model interprets transcripts differently, enabling users to choose based on tone or summary length preferences.

3. **Transcript Preprocessing:**

Before summarization, transcripts are cleaned by removing timestamps, noise words, and irrelevant tokens to ensure quality input for the AI engine.

4. **Multilingual Processing:**

The system incorporates translation mechanisms to allow summaries in multiple languages. Currently, English and German are supported with future scope for expanding to other global/regional languages.

5. **User Interface & Accessibility:**

The application is crafted with a strong emphasis on usability and responsive design. It features a clean user interface where users can easily paste YouTube video links and select their preferred AI model and language using intuitive toggles. Once initiated, the system processes the video and displays the summary in a scrollable, copy-friendly area. A responsive layout ensures seamless access across devices, and dark mode enhances visual comfort. Additionally, a sidebar provides quick access to previously generated summaries, improving user convenience and workflow continuity.

6. **Scalability and Inclusivity:**

The model can be extended to support regional languages and dialects in the future. It can also integrate with educational platforms and public information kiosks for maximum impact. The current system is designed with future scalability in mind, enabling expansion into several key areas. Planned enhancements include support for additional regional and international languages to broaden accessibility, and the development of a browser extension for real-time summarization while watching videos. A mobile app version is also envisioned, allowing users to summarize on the go and access saved content offline. Further improvements include integrating speech-to-text for videos lacking transcripts, customizable summary lengths, and user account systems with cloud storage to sync and securely manage summaries across multiple devices.

7. **Ethical Design and Privacy:**

The system ensures ethical design by prioritizing user privacy, using encrypted API calls, avoiding data storage, and maintaining transparency, consent, and control throughout user interactions.

IV. CONCLUSION

The YouTube Summarizer project provides a powerful yet user-friendly solution for quick understanding of video content using cutting-edge AI. Its model flexibility, multilingual support, and clean interface make it applicable across various domains including education, content creation, and research. Future improvements may include support for more languages, personalized summarization tones, and integration with browser extensions or mobile apps to increase reach and usability.

V. ACKNOWLEDGEMENT

We extend our sincere gratitude to the faculty and administration of SNS college of engineering for their
continued encouragement and support. We especially thank our project guide Dr.V.Indu Nair for her/his valuable insights and mentorship. We also acknowledge our peers and team members for their collaboration and constant motivation throughout this endeavor.

REFERENCES

[1] Huang, J.-H. (2024).“Personalized Video Summarization using Text-Based Queries and Conditional Modeling” This thesis explores enhancing video summarization by integrating text-based queries and conditional modeling, aiming to tailor summaries to user needs and improve semantic understanding.
[2] Bandabe, S., Zambre, J., Gosavi, P., Gupta, R., & Gaikwad, J. A. (2023)” YouTube Transcript Summarizer Using Flask “This

project proposes a YouTube Transcript summarizer that accesses video transcripts via Python API and summarizes them using the transformers package, enhancing user comprehension without watching entire videos.
[3] N. Alok, K. Krishan, and P. Chauhan, "Deep learning-based image classifier for malaria cell detection," in Machine learning for healthcare applications, pp. 187–197, 2021.
[4] M. Z. Alom et al., "A state-of-the-art survey on Deep Learning theory and architectures," Electronics (Basel), vol. 8, no. 3, p. 292, 2019.
[5] L. Alzubaidi et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," J. Big Data, vol. 8, no. 1, p. 53, 2021.
[6] E. Apostolidis, E. Adamantidou, A. I. Metsai, V. Mezaris, and I. Patras, "Unsupervised video summarization via at-ten-tion-driven adversarial learning," in international conference on multimedia modeling, Springer, pp. 492–504, 2020.
[7] E. Apostolidis, E. Adamantidou, A. I. Metsai, V. Mezaris, and I. Patras, "AC-SUM-GAN: Connecting actor-critic and generative adversarial networks for unsupervised video summarization," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 8, pp. 3278–3292, 2021.
[8] N. Archana and N. Malmurugan, "RETRACTED ARTICLE: Multi-edge optimized LSTM RNN for video summarization," J. Ambient Intel. Humanize. Compute., vol. 12, no. 5, pp. 5381–5395, 2021.
[9] M. Basavarajaiah and P. Sharma, "Survey of compressed domain video summarization techniques," ACM compute. Surv., vol. 52, no. 6, pp. 1–29, 2020.