Cognitive AI-Driven Video Summarization with Context-Aware Chatbot Query Resolution

Vidya P
School of CSE
REVA University
Bengaluru,India

Sharon Chattopadhyay
School of CSE
REVA University
Bengaluru,India

Lavanya Latha P
School of CSE
REVA University
Bengaluru,India

Jeffrey Immanuel
School of CSE
REVA University
Bengaluru,India

Dr. S. Bhargavi Latha
Associate Professor, School of CSE
REVA University
Bengaluru,India

Abstract—**In the digital age, video content consumption has surged, making it essential to efficiently extract relevant information from lengthy videos. This project introduces a Flask-based web application that summarizes YouTube videos and uploaded MP4 files, addressing the time-consuming nature of video analysis. The application utilizes a structured architecture with Flask Blueprints to ensure modularity and maintainability. Key features include audio extraction and speech-to-text conversion via libraries such as Speech Recognition, followed by concise summarization using Natural Language Processing (NLP) models like those from OpenAI and Hugging Face. With a focus on processing everything in-memory to ensure user privacy and efficiency, the application allows users to swiftly obtain insights without needing to view entire videos. Future enhancements include user authentication, support for additional file formats, and multilingual capabilities, paving the way for an advanced emotion-aware video summarization tool tailored for diverse applications.**

Keywords—Flask, video summarization, audio processing, speech recognition, Natural Language Processing (NLP), modular architecture, web application.

I. INTRODUCTION

The rapid proliferation of digital content, particularly in the forms of video and text, has created an ever-growing demand for efficient and accurate summarization techniques. Vast amounts of data are generated across numerous platforms—ranging from educational and corporate content to social media and entertainment—which makes it increasingly critical to distill this information into concise. This growing need has motivated extensive research into video and text summarization methods that help users manage and leverage content in a time-efficient manner. Recent advancements in deep learning, particularly through models like hierarchical multimodal transformers, have significantly improved the quality of generated summaries by effectively capturing complex dependencies between visual and auditory elements. Furthermore, traditional techniques like TF-IDF and Latent Dirichlet Allocation (LDA) continue to be relevant, providing essential capabilities for keyword extraction and topic identification within textual data. Nonetheless, ongoing challenges such as scalability, the necessity for diverse datasets, and catering to user preferences present opportunities for future innovation in summarization techniques, across industries like healthcare, education, and entertainment is crucial. Previous research on neural summarization has focused on extractive, abstractive, and hybrid methods. However, these methods exhibit limitations when summarizing long documents the challenge of capturing diverse key information points [1].

In the realm of video summarization, various methodologies have been explored, including traditional techniques, recurrent neural network (RNN)-based methods, and attention-based approaches. Despite their contributions, existing methods often struggle to capture global dependencies and multi-hop relationships among video frames, leading to inadequate summarization results. transformer-based architectures as solution to these limitations, offering a more effective means of summarization [2].

The growing interest in video summarization research stems from its wide range of applications. Many existing methods treat the generation of video summaries and text summaries as independent tasks, neglecting the semantic correlation necessary for coherent cross-modal understanding. This work introduces the concept of cross-modal video summarization, which aims to condense lengthy videos into shorter clips while producing semantically aligned textual summaries [3].

Further exploration into deep learning for video summarization has revealed both supervised and unsupervised strategies. Notably, a novel approach incorporates an Actor-Critic model within a Generative Adversarial Network to learn optimal policies for key-fragment selection, all in an unsupervised context. This methodology addresses the limitations of prior models, particularly concerning the temporal dependencies of frames and the diversity of visual content [4].

As video content continues to surge, there is an increasing demand for effective summarization techniques that enhance content accessibility. Traditional methods such as template matching and end-to-end models have proven inadequate in accurately capturing and describing video content, necessitating more sophisticated approaches that can generate meaningful textual summaries [5][6]. Moreover, the evolution of sentiment analysis tools is gaining interest, with researchers emphasizing the benefits of integrating visual and auditory modalities to improve the accuracy of sentiment assessment within video content [7].

In the field of sentiment analysis, researchers have identified key areas of focus, including the evaluation of datasets, feature extraction methods, and fusion techniques that integrate multiple modalities for improved accuracy [7]. The evolving landscape of video summarization research has led to the categorization of approaches into three primary directions: domain-specific, generic, and query-focused. Despite advancements, existing methods struggle with accurately addressing user preferences and achieving semantic similarity in summaries. This has driven the exploration of more effective techniques capable of generating query-focused summaries that align closely with the semantic intent of user inquiries [8].

Furthermore, current video summarization methods often yield a single summary that may not meet the diverse interests of users. To address this limitation, researchers have proposed a topic-aware video summarization task, which aims to generate multiple summaries that correspond to different thematic elements within the video. However, this approach faces challenges, particularly concerning the availability of appropriate datasets and the necessity for models that can effectively integrate multimodal information to capture various topics [9].

The education sector has also been a focal point for video summarization research, where existing methods based on audio, visual, and textual elements often fall short. Limitations include the inadequacy of extractive techniques for summarizing educational subtitles, which calls for innovative approaches that can effectively encapsulate the essential content while maintaining fidelity to the original message [10].

Studies on video summarization and transfer learning highlight effective techniques. The Hugging Face Transformers framework achieves high accuracy in abstractive summarization, while the BERTSum model excels in summarizing instructional content. Additionally, podcast summarization integrates automated speech recognition (ASR) with text summarization, and models like BART enhance NLP performance through pre-trained language models [11]

The examination of conversational agents has also been a significant area of research. Studies have outlined three main methodologies for analyzing interactions with these agents, including the use of existing chatbots, the Wizard-of-Oz (WoZ) method, and the development of custom chatbots using APIs or toolkits like CART. Researchers note inherent limitations in these methods, such as issues with privacy and resource intensiveness, which may hinder the exploration of actual technological capabilities [12].

In the context of linguistic corpora, a review of bilingual and multilingual parallel corpora has revealed various limitations, including a dominance of high-density languages, restricted text types, and a lack of sentence alignment or part-of-speech tagging. Projects like WIT^3, which compiled TED talk subtitles, were noted for not providing necessary alignment features [13].

Concerning content summarization techniques, a classification of methods into extractive, compressive, and abstractive categories has been proposed, outlining limitations such as dependency on specific resources and the necessity of extensive training datasets for deep learning models. Moreover, various techniques for evaluating machine translation quality—ranging from word-level to sentence-level estimations—were discussed, emphasizing ongoing challenges in the field [14].

Recent advances in text summarization have highlighted traditional methods like TF-IDF and Bayesian models alongside contemporary deep learning techniques, including Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. Limitations addressed include the reliance on crucial phrase extraction techniques in traditional models and parallelization difficulties in RNNs and LSTMs. The significant role of attention mechanisms in enhancing the performance of sequence-to-sequence models for text summarization has also been underscored [15].

II. Literature Survey

The Gidiotis, A. and Tsoumakas, G. paper [1] reviews previous research on neural summarization, emphasizing methods including extractive, abstractive, and hybrid approaches. They address limitations faced when summarizing long documents, such as high computational complexity and difficulties in capturing diverse key information points, underscoring the need for more effective summarization techniques.

The authors Zhao, B., Gong, M., and Li, X. paper [2] reviews the literature on video summarization, encompassing traditional methods, RNN-based techniques, and attention-based approaches. They highlight the limitations of existing methods in capturing global dependencies and multi-hop relationships among video frames, proposing a transformer-based approach to address these challenges.

The authors J. Lin et al. paper [3], highlight the growing research interest in video summarization due to its wide range of applications. They note the limitations of existing methods, which often treat the generation of video and text summaries as independent tasks, thereby neglecting the semantic correlation between visual and textual summarization. This work introduces a new task called cross-modal video summarization, aiming to condense lengthy videos into shorter clips while producing semantically aligned textual summaries.

The authors E. Apostolidis, E. Adamantidou, A. I. Metsai, V. Mezaris, and I. Patras. paper [4], review the literature on deep-learning-based video summarization, encompassing both supervised and unsupervised approaches. They highlight the limitations of existing methods, particularly in capturing the temporal dependencies of video frames and the diversity of visual content. To address these issues, they propose a new method that embeds an Actor-Critic model into a Generative Adversarial Network, enabling the learning of a policy for key-fragment selection and summarization in a fully unsupervised manner.

The authors S. Ahmed et al. paper [5], highlight the rapid growth of video data and the increasing demand for effective video summarization techniques. They discuss the limitations of existing methods, such as template matching and end-to-end models, which often fail to accurately capture and describe the content of videos. The authors emphasize the necessity for a more sophisticated approach that can generate meaningful and accurate textual descriptions of video content.

The authors R. A. Albeer, H. F. Al-Shahad, H. J. Aleqabie, and N. D. Al-Shakarchy. paper [6], highlight the increasing volume of online videos and the corresponding need for effective summarization techniques to enhance accessibility. They discuss the limitations of existing methods, such as template matching and end-to-end models, which often struggle to accurately capture and describe video content. The authors emphasize the importance of developing a more sophisticated approach that can generate meaningful and accurate summaries of video transcripts.

The authors S. Al-Azani and E.-S. M. El-Alfy. paper [7] highlight the increasing volume of online videos and the need for effective sentiment analysis tools. They discuss the limitations of text-based sentiment analysis and emphasize the benefits of incorporating visual and auditory modalities. Their work reviews existing research on multimodal sentiment analysis, focusing on datasets, features, and fusion techniques.

The authors S. Xiao, Z. Zhao, Z. Zhang, Z. Guan, and D. Cai. paper [8] analyze three main research directions in video summarization: domain-specific, generic, and query-focused. They highlight the limitations of existing methods in addressing user preferences and capturing semantic similarity. Their work emphasizes the need for more effective approaches that generate query-focused summaries while preserving semantic meaning.

The authors Y. Zhu, W. Zhao, R. Hua, and X. Wu. paper [9] highlight the limitations of existing video summarization methods, which generate a single summary that may not address diverse user interests. They propose a topic-aware video summarization task that produces multiple summaries based on different topics within a video. Additionally, they emphasize the lack of suitable datasets for this task and the need for a model capable of effectively fusing multimodal information to capture topic variations.

The authors S. S. Alrumiah and A. A. Al-Shargabi. paper [10] examine previous work on educational video summarization, exploring methods that utilize audio, visual, and textual content. They highlight the limitations of existing approaches and emphasize the need for a more effective method for summarizing subtitles. Their work discusses extractive summarization techniques and the challenges associated with summarizing educational videos.

The authors V. Mehta, T. Deshpande, R. Pandey, T. Kandoi, and S. Nair. paper [11] explore various studies on video summarization and transfer learning, discussing different techniques and models. They highlight the effectiveness of Hugging Face Transformers in abstractive summarization, the BERTSum model for summarizing narrated instructional videos, and podcast audio summarization using automated speech recognition (ASR) and text summarization. Additionally, they examine the use of Latent Semantic Analysis for video summarization and emphasize the proficiency of pre-trained language models like BART in various NLP applications.

The author T. Araujo. paper [12] examines three main approaches for studying interactions with conversational agents in experimental settings: using existing chatbots, employing the Wizard-of-Oz (WoZ) method, or creating chatbots using APIs or toolkits like CART. The study highlights the limitations of existing chatbots and the WoZ method, including privacy concerns and resource intensity. Additionally, the author argues that simulating interactions instead of utilizing actual technology may hinder researchers from critically evaluating the computational methods and their potential limitations.

The authors I. Zeroual and A. Lakhouaja. paper [13] analyze related work on bilingual and multilingual parallel corpora, including CzEng, Scielo, FAPESP, OPUS, SwissAdmin, and AMARA. They discuss the limitations of existing corpora, such as their focus on high-density languages, restricted text types, and the absence of sentence alignment or POS tagging. Additionally, they mention WIT³, a project that collected TED talk subtitles but lacked sentence alignment and POS tagging.

The authors] E. L. Pontes, S. Huet, J.-M. Torres-Moreno, and A. C. Linhares . paper[14] categorize CLTS methods into extractive, compressive, and abstractive approaches. They discuss the limitations of existing methods, including dependency on specific resources, limited language support, and the requirement for large training datasets in deep learning models. Additionally, they examine various techniques for evaluating machine translation quality at the word, phrase, and sentence levels.

The authors B. N, D. Kumari, B. N, M. N, S. K. P and S. R. A. paper[15] examine various text summarization approaches, including traditional methods like TF-IDF and Bayesian models, alongside deep learning techniques such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. They discuss limitations, including reliance on key phrase extraction in traditional methods and parallelization challenges in RNNs and LSTMs. Additionally, they emphasize the role of attention mechanisms in enhancing the performance of sequence-to-sequence models for text summarization.

III. Methodology

A. *Proposed System*

The proposed system for video summarization integrates advanced NLP and machine learning techniques to generate concise and meaningful summaries from video content. The methodology follows a structured approach, beginning with data preprocessing, feature extraction, and summarization. The system is designed to process both YouTube videos and regular video files, even in the absence of transcripts.

Users can upload a YouTube video link via an upload button built into a Flask-based web interface. The YouTube Transcript API extracts the transcript, which is then processed using a transformer-based model such as Hugging Face Transformers to generate a summary. If a transcript is unavailable, the system extracts audio from the video using MoviePy and converts it into text using speech recognition modules. This extracted text is then summarized using the same transformer model. Additionally, the system includes an integrated chatbot that allows users to ask queries related to the summarized content and receive optimal responses. The chatbot enhances user interaction by leveraging NLP techniques to provide relevant answers based on the generated summaries.

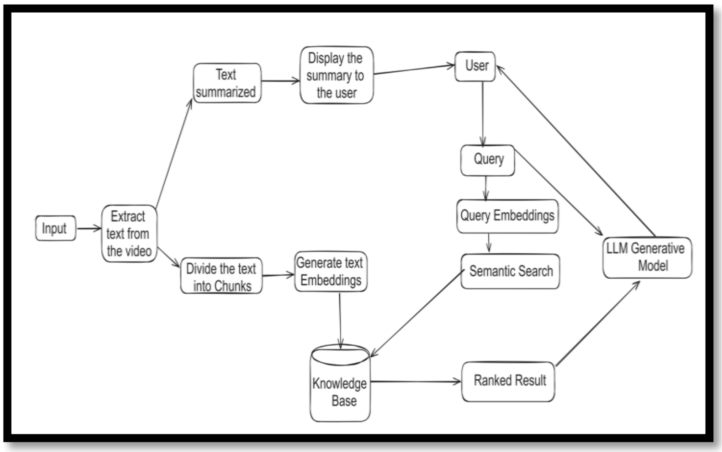

Fig.1. Internal working of the model .

The summarized text is further divided into chunks, and embeddings are generated to create a structured knowledge base. Users can interact with the system through a chatbot, where their queries are transformed into embeddings and matched against the knowledge base using semantic search. The system retrieves the most relevant ranked results, supplemented by responses from an LLM generative model to enhance accuracy and comprehensiveness.

The overall architecture of the proposed system is illustrated in Fig. 1, which depicts the flow from input video processing to user interaction via semantic search and chatbot integration.

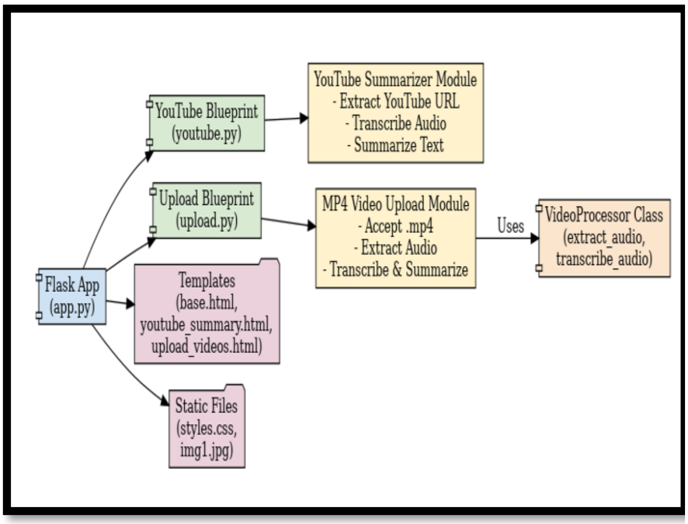

Fig.2.Architectural Design.

Fig. 2, illustrates the architecture of the Flask-based application, detailing the interaction between different components.At the core of the system is the app.py file, which serves as the main entry point, managing different blueprints and rendering templates. The application consists of two primary blueprints: the YouTube Blueprint (youtube.py), responsible for extracting YouTube video URLs, transcribing audio, and summarizing text, and the Upload Blueprint (upload.py), which enables users to upload .mp4 files, extract audio, and generate summaries.

The YouTube Summarizer Module processes video URLs, fetching transcripts or utilizing ASR if transcripts are unavailable. Similarly, the MP4 Video Upload Module handles locally uploaded videos, extracting audio and performing speech-to-text transcription before summarization. Both modules rely on the VideoProcessor Class, which includes methods for audio extraction and transcription. Additionally, the application uses template files (base.html, youtube_summary.html, upload_videos.html) to render dynamic web pages, while static files such as CSS and images enhance the UI. This structured modular approach ensures scalability and ease of integration with other NLP-based functionalities.

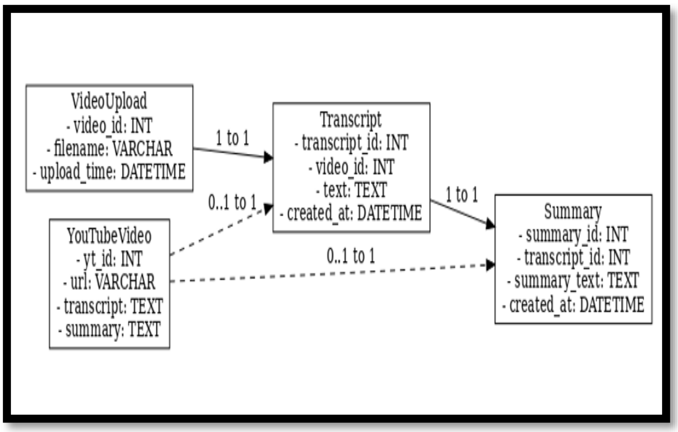

Fig.3.Entity Relationship Model.

Fig. 3, illustrates the Entity-Relationship (ER) model of the system, defining the data structure for handling video uploads, transcripts, and summaries. The VideoUpload entity stores uploaded video details, including video_id, filename, and upload_time, ensuring each uploaded video is uniquely tracked. Similarly, the YouTubeVideo entity maintains records of YouTube URLs, their transcripts, and summaries without overriding previous records.

The Transcript entity is associated with either a video_id (uploaded videos) or a yt_id (YouTube videos), ensuring that each transcript is uniquely stored. The Summary entity is linked to transcript_id, preventing duplication and ensuring that each generated summary is uniquely stored without collapsing previous video data. This structure enables efficient retrieval, preventing conflicts when multiple videos are processed.

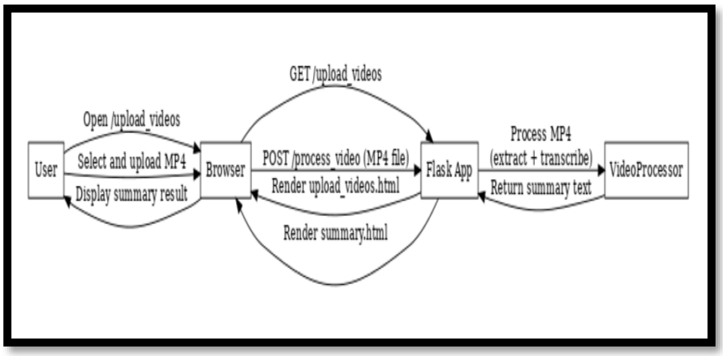

Fig.4. Sequence Diagram.

Fig. 4 represents the workflow of the MP4 video summarization process in the Flask-based application. The process begins when the User accesses the /upload_videos endpoint via a Browser, where they can select and upload an MP4 file. The Flask App handles this request by rendering the upload_videos.html page and processing the uploaded video via a POST request to process_video.Once received, the Flask App sends the video to the VideoProcessor module, which extracts audio, transcribes it, and generates a summary. The summarized text is then returned to the Flask App, which renders the summary.html page and displays the results to the User, completing the process.

B. Advantages of Proposed model

1. Works with All Languages: Speech recognition and summarization models support multiple languages, making the system highly adaptable.
2. No Dependence on Pre-Existing Transcripts: The model efficiently handles videos with and without transcripts by extracting audio and converting it into text.
3. Seamless Integration with Chatbot: Users can interact with the summarized content and obtain clarifications through an intelligent chatbot.
4. Scalability and Modularity: Flask Blueprints ensure modularity, while the transformer models offer scalability for enhanced summarization performance.
5. Future-Ready Design: The system is designed to extend its functionality to real-time summarization of live videos via webcam.

IV . EXPERIMENTAL ANALYSIS

The system was tested across various video genres, including educational content, news reports, and general discussions, to evaluate its performance. The summarization accuracy was assessed by comparing outputs with manually generated summaries. The chatbot's effectiveness was evaluated based on user interactions and response relevance.

Model Building –

The proposed system integrates multiple AI-driven techniques for extracting, transcribing, summarizing, and querying video content. The architecture primarily consists of Flask-based APIs, Hugging Face Transformer models, speech-to-text modules, and a semantic search-enabled chatbot. The model building process ensures robust handling of both YouTube videos (with or without transcripts) and locally uploaded videos.

The system follows a structured pipeline:

1. Video Processing & Transcript Extraction
    - If the input is a YouTube video, the YouTube Transcript API extracts subtitles.
    - If the video lacks subtitles or is a locally uploaded file (MP4 format), FFmpeg extracts the audio, which is then processed using Speech Recognition and whisper ASR models to convert speech to text.
2. Text Summarization
    - The extracted transcript is passed through a Hugging Face Transformer-based Summarization Model, such as BART, T5, or Pegasus, which condenses the content while preserving key information.
    - The summary is stored in a database for efficient retrieval.
3. Embedding Generation & Knowledge Base
    - The summarized text is divided into manageable chunks and vector embeddings are generated using models like SBERT .
    - These embeddings are stored in a knowledge base for semantic retrieval.
4. User Query Handling & Chatbot Interaction
    - Users can interact with the system via a chatbot to ask queries related to the video content.
    - The chatbot converts the query into an embedding and searches the knowledge base using semantic search.
    - Based on similarity ranking, the chatbot provides relevant responses, either retrieved from stored summaries or generated using a Large Language Model (LLM) for enhanced responses.
5. User Interface & Future Enhancements
    - The system provides a Flask-based web UI where users can upload videos, view summaries, and interact with the chatbot.
    - Future enhancements include real-time transcription and summarization from live video streams using webcam input.

Evaluation and Testing –

The system's performance was evaluated based on:

- Summarization Accuracy: Comparing AI-generated summaries with manually written summaries.
- Speech Recognition Efficiency: Testing the accuracy of transcriptions generated from video audio.
- Chatbot Response Quality: Measuring relevance and accuracy of chatbot-generated responses.

The model was found to be highly effective in summarizing both YouTube and local video files, handling various languages efficiently.

A . Result

The proposed system successfully generates accurate and concise summaries for both YouTube and normal video files, even when transcripts are unavailable. The chatbot integration enhances user experience by providing interactive query resolution. The system's adaptability to multiple languages makes it a versatile tool for diverse applications.

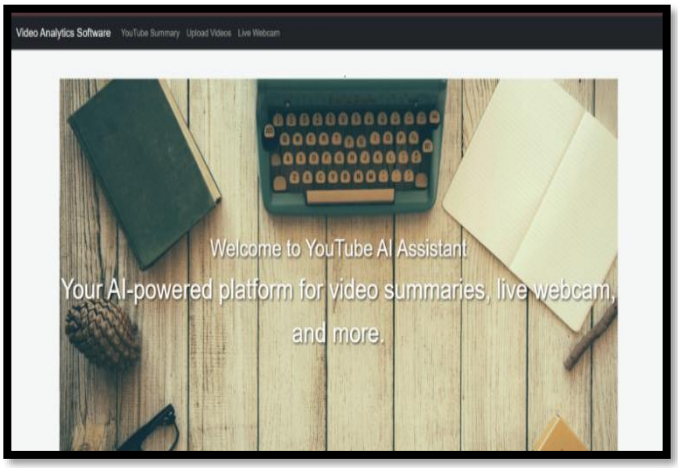

Fig.5.User Interface of Video Analytics Software.

The landing page of the Video Analytics Software provides access to YouTube summarization, MP4 video processing, and live webcam analytics. It ensures smooth navigation and an AI-driven approach to video analysis in Fig.5

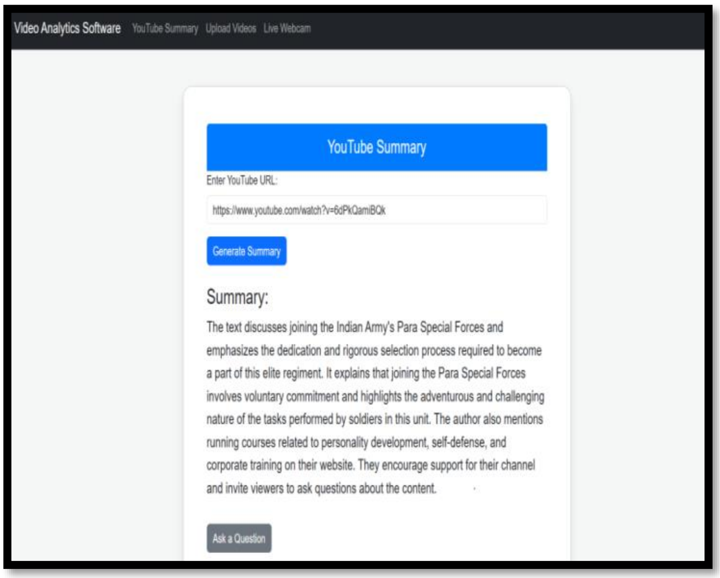

Fig.6. Video Summarization Process

The YouTube summarization module extracts transcripts from videos and generates concise summaries. Users input a YouTube link, and the system processes the video to deliver a brief textual overview in Fig.6.

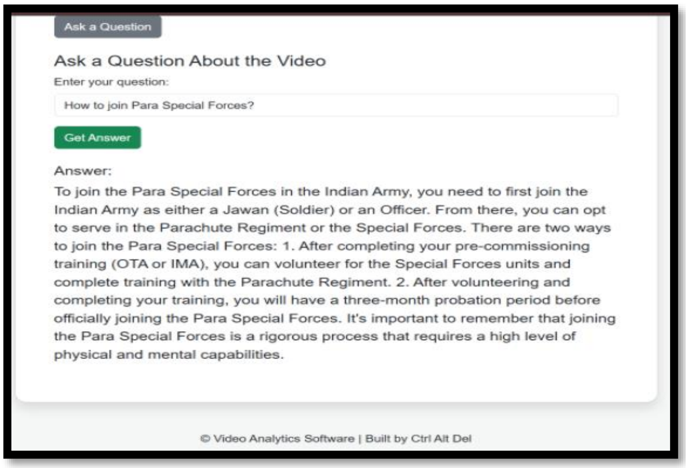

Fig.7. Question-Answering System for Summarized Videos

The question-answering feature enables users to ask specific questions related to the summarized video. Using NLP techniques, the system retrieves accurate responses based on the video transcript in Fig.7.

The Video Analytics Software effectively extracts and summarizes content from both YouTube and uploaded videos, ensuring accessibility across multiple languages. Additionally, the integrated chatbot enhances user interaction by providing accurate responses to queries based on the summarized content.

B. Conclusion

The proposed Video Analytics Software provides an efficient approach to video summarization and question-answering, making video content more accessible and understandable. By leveraging Hugging Face Transformers for text-based summarization and speech recognition for audio-based videos, the system ensures comprehensive content extraction, even from videos without transcripts. The integration of a chatbot further enhances user interaction, allowing seamless query resolution based on summarized content. The system supports multiple languages, making it adaptable for diverse users and applications.

For future enhancements, the model can be extended to support real-time video summarization using webcam feeds, enabling live content processing. Additionally, improving the chatbot's contextual understanding and integrating multimodal AI techniques could further refine user experience and result accuracy.

REFERENCES

[1] A. Gidiotis and G. Tsoumakas, “A Divide-and-Conquer approach to the summarization of long documents,” IEEE/ACM Transactions on Audio Speech and Language Processing, vol. 28, pp. 3029–3040, Jan. 2020, doi: 10.1109/taslp.2020.3037401.

[2] B. Zhao, M. Gong, and X. Li, “Hierarchical multimodal transformer to summarize videos,” Neurocomputing, vol. 468, pp. 360–369, Oct. 2021, doi: 10.1016/j.neucom.2021.10.039.

[3] J. Lin et al., “VideoXum: Cross-modal Visual and Textural Summarization of Videos,” IEEE Transactions on Multimedia, pp. 1– 13, Jan. 2024, doi: 10.1109/tmm.2023.3335875.

[4] 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 Transactions on Circuits and Systems for Video Technology, vol. 31, no. 8, pp. 3278–3292, Nov. 2020, doi: 10.1109/tcsvt.2020.3037883.

[5] S. Ahmed et al., “Att-BiL-SL: Attention-Based Bi-LSTM and Sequential LSTM for Describing Video in the Textual Formation,” Applied Sciences, vol. 12, no. 1, p. 317, Dec. 2021, doi: 10.3390/app12010317.

[6] R. A. Albeer, H. F. Al-Shahad, H. J. Aleqabie, and N. D. Al- Shakarchy, “Automatic summarization of YouTube video transcription text using term frequency-inverse document frequency,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 26, no. 3, p. 1512, Jun. 2022, doi: 10.11591/ijeecs.v26.i3.pp1512-1519.

[7] S. Al-Azani and E.-S. M. El-Alfy, “Enhanced Video Analytics for Sentiment Analysis Based on Fusing Textual, Auditory and Visual Information,” IEEE Access, vol. 8, pp. 136843–136857, Jan. 2020, doi: 10.1109/access.2020.3011977.

[8] S. Xiao, Z. Zhao, Z. Zhang, Z. Guan, and D. Cai, “Query-Biased Self- Attentive Network for Query-Focused Video Summarization,” IEEE Transactions on Image Processing, vol. 29, pp. 5889–5899, Jan. 2020, doi: 10.1109/tip.2020.2985868.B

[9] Y. Zhu, W. Zhao, R. Hua, and X. Wu, “Topic-aware video summarization using multimodal transformer,” Pattern Recognition, vol. 140, p. 109578, Mar. 2023, doi: 10.1016/j.patcog.2023.109578.

[10] S. S. Alrumiah and A. A. Al-Shargabi, “Educational Videos Subtitles’ Summarization Using Latent Dirichlet Allocation and Length Enhancement,” Computers, Materials & Continua/Computers, Materials & Continua (Print), vol. 70, no. 3, pp. 6205–6221, Oct. 2021, doi: 10.32604/cmc.2022.021780.

[11] V. Mehta, T. Deshpande, R. Pandey, T. Kandoi, and S. Nair, “Video Transcript Extraction and Summarization Using Transfer Learning,” in Lecture notes in networks and systems, 2023, pp. 625–636. doi: 10.1007/978-981-19-9638-2_54.

[12] T. Araujo, “Conversational Agent Research Toolkit,” Computational Communication Research, vol. 2, no. 1, pp. 35–51, Feb. 2020, doi: 10.5117/ccr2020.1.002.arau.

[13] I. Zeroual and A. Lakhouaja, “MulTed: a multilingual aligned and tagged parallel corpus,” Applied Computing and Informatics, vol. 18, no. 1/2, pp. 61–73, Dec. 2018, doi: 10.1016/j.aci.2018.12.003.

[14] E. L. Pontes, S. Huet, J.-M. Torres-Moreno, and A. C. Linhares, “Compressive approaches for cross-language multi-document summarization,” Data & Knowledge Engineering, vol. 125, p. 101763, Nov. 2019, doi: 10.1016/j.datak.2019.101763.

[15] B. N, D. Kumari, B. N, M. N, S. K. P and S. R. A, "Text Summarization using NLP Technique," 2022 International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Shivamogga, India, 2022, pp. 30-35, doi: 10.1109/DISCOVER55800.2022.9974823.