



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 10, October 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.625



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com



Analysis and Summarization of YouTube Video Using Natural Language Processing

Ajita Jane Antony Vincent¹, Sruthi Suresh¹, Thrishal L.¹, Dr. Akila K.²

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India¹

Assistant Professor (S.G.), Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India²

ABSTRACT: In this study, we aimed at presenting a system for automating the summarization of YouTube transcripts to reduce the time required for content consumption. To obtain relevant information from YouTube videos, viewers often need to spend excessive time in watching entire videos. This system uses Natural Language Processing (NLP) techniques to analyze video transcripts and generates concise summaries with the key points. The system effectively reduces transcript length by using the YouTube Transcript API, transformer-based models, and summarization pipelines without affecting the essential details. This tool offers enhanced video accessibility through efficient transcript summarization for both viewers and content creators.

KEYWORDS: Transcript, YouTube Transcript API, Text Summarizer, Natural Language Processing (NLP), Latent Semantic Analysis (LSA), Graphical User Interface (GUI)

I. INTRODUCTION

In the era of social media, video content has emerged as one of the most effective means of disseminating information. YouTube, in particular, hosts billions of videos that cover a variety of topics. However, the time-consuming aspect of watching lengthy videos, where the viewers need to access only the key information, has resulted in greater interest in text summarization technologies. Automatic summarization of YouTube contents becomes feasible due to recent developments in Natural Language Processing (NLP). This study discusses an automated system that is designed to summarize YouTube transcripts quickly and very effectively using different NLP techniques. This system provides users a concise yet informative overview of video contents.

The system uses the YouTube Transcript API to fetch video transcripts and applies summarization algorithms, converting them to other languages when required. The main purpose of our study is to help individuals with time constraints or language barriers who require YouTube content information. This tool allows viewers to obtain the main ideas of a video without watching it in full by simply summarizing transcripts. This tool it is useful for both content consumers and creators.

II. LITERATURE SURVEY

A review of recent literature highlights the advancements in the field of YouTube video transcription and summarization. The following are key contributions that offer insights into the technologies and methodologies applied in this domain.

Video Transcript Summarizer

This article by Atluri et al. (2022) proposes a system that summarizes YouTube video transcripts by using natural language processing (NLP) and machine learning without affecting key information [1]. With the growing number of online educational videos, the system addresses the challenges of users who have to watch full-length videos to understand the context. The model fetches video transcripts from user-provided video links and uses Hugging Face Transformers and pipelining techniques to generate summaries. The system allows users to input a required summary



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

duration, delivering an optimized transcript without omitting important key information. Further test results on this method revealed reduced processing time compared to other techniques. It is worth noting that these obtained summaries successfully preserved the central message of the videos. This method is applicable for educational purposes across platforms such as YouTube, Facebook, and Google, enhancing both accessibility and content retrieval.

YouTube Transcript Summarizer

This article by Kumar et al. (2023) presents a web application developed to summarize YouTube video transcripts using NLP and the Flask framework [2]. The tool extracts essential information from lengthy video transcripts by employing text processing techniques such as lemmatization, named entity recognition, and portion labelling. The Flask-based system is hosted on a web server, where users can input YouTube video links and get concise summaries. This tool only works with videos that have properly formatted closed captions, making it easier for viewers to quickly access important information without watching the entire video. The authors highlight the ease of use, allowing quick API queries to generate transcript summaries online.

YouTube Transcript Summarizer

The article by Panthagani et al. (2024) introduces a video transcript summarizer system that is aimed at improving the extraction of crucial information from YouTube videos [3]. The system, based on NLP and machine learning, summarizes video transcripts without losing important content and helps users to avoid the less relevant parts of videos. As the amount of video content on platforms like YouTube continues to grow, the summarizer helps in quick retrieval of important data by generating a concise output video. The article highlights its effectiveness for educational and technical applications where video summaries reduce storage needs and allow users to locate essential sections faster. In this study, the authors use modern NLP techniques and provide an opportunity for advanced text summarization implementation.

This review demonstrates the continuous development of transcription and summarization tools for video content, especially for platforms like YouTube. While these methods offer substantial improvements in terms of accuracy and usability, they also highlight challenges such as setup complexity and the need for well-prepared input data.

III. METHODOLOGY AND ARCHITECTURE DIAGRAM

The proposed system uses several key components to deliver a summarized version of YouTube video transcripts using NLP. Figure 1 illustrates the core steps from the URL input to saving the final output summary.

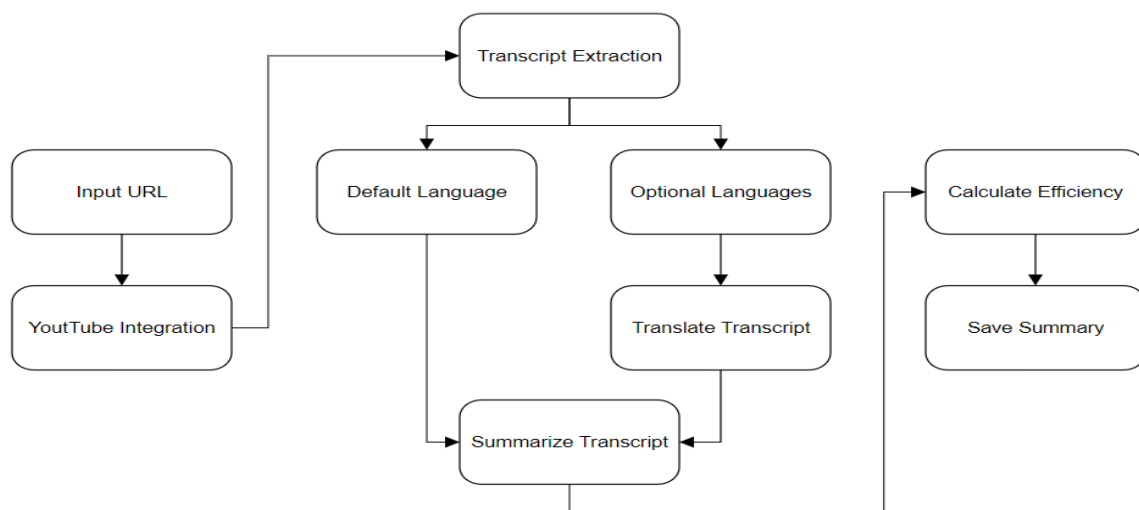


Figure 1. Architecture diagram: the flow of summarization.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Transcript Retrieval

The system uses the YouTube Transcript API [4] to retrieve the video's transcript based on the provided video ID or URL. The transcripts contain the spoken words from the video, which allow the system to work with raw textual data.

Translation

For non-English videos, the transcripts are translated using Google's Translator API [5]. Users can select from a variety of languages, and the system will provide the summarized text in their preferred language.

Text Summarization

The summarization process is performed using a Latent Semantic Analysis (LSA) [6], a model that effectively reduces the transcript while retaining key points. LSA identifies important topics by analysing relationships between terms in the transcript, extracting the sentences that best represent the core content. In this system, the user can specify the desired length of the summary.

PDF Export

The final transcript and summary can be saved as a PDF document, making it easy to share and reference later. The PDF includes both the original transcript and the generated summary, ensuring the user has access to the complete information if needed.

Efficiency Metrics

An additional feature calculates and displays the summarization efficiency, i.e., the ratio of the word count in the summary compared to the original transcript. This allows users to have a quantitative view of how much the transcript was condensed.

Graphical User Interface (GUI)

A user-friendly GUI, built using the Tkinter library [7], allows users to interact with the system, input video URLs, select languages, view transcripts, and download summaries. The interface displays real-time updates on word count and summarization efficiency, ensuring complete transparency in the summarization process.

IV. EXPERIMENTAL ANALYSIS

Input

To execute an example of the YouTube summarization module, the user must provide a URL as input. For example, let us consider the following YouTube Video URL: https://www.youtube.com/watch?v=MS5UjNKw_1M

Output

Figures 2–4 describe the transcript summary along with the efficiency levels for the input URL in English, Spanish, and French, respectively.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The screenshot shows a web browser window titled "YouTube Transcript Summarizer". At the top center is a large red play button icon. Below it, the text "Enter YouTube video URL or ID:" is followed by a text input field containing the video ID "M55UjNKw_1M". Underneath is a "Select a Language:" label with a dropdown menu showing "english". A "Submit" button is located below the language selection. The main content area is divided into two sections: "Original Transcript:" and "Summary:". The original transcript text reads: "BARACK OBAMA: I'm so in love with you. This is your victory. I know you didn't do this just to win an election. And I know you didn't do it for me. You did it because you understand the enormity of the task that lies ahead. But my main message is to the parents of Trayvon Martin. If I had a son, he'd look like Trayvon. They had their entire lives ahead of them, birthdays, graduations,". The summary text reads: "I know you didn't do this just to win an election. But my main message is to the parents of Trayvon Martin. It's the idea held by generations of citizens, who believe that America is a constant work in progress, who believe that loving this country requires more than singing its praises or avoiding uncomfortable truths. Facts, evidence, reason, logic, an understanding of science, these are good things. Black, white, Latino, Asian, Native American, young, old, gay, straight, men, women, folks with disabilities all pledging allegiance under the same proud flag to this big,". At the bottom, it displays "Transcript Words: 238, Summary Words: 99" and "Efficiency: 41.60%".

Figure 2. Transcript and summary of the video in English with efficiency.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The screenshot shows a web browser window titled "YouTube Transcript Summarizer". At the top center is a large red play button icon. Below it, the text "Enter YouTube video URL or ID:" is followed by a text input field containing "MSSUjNKw_1M". Underneath is a "Select a Language:" label with a dropdown menu showing "french". A "Submit" button is located below the language selection. The "Original Transcript:" section contains the following text: "BARACK OBAMA: I'm so in love with you. This is your victory. I know you didn't do this just to win an election. And I know you didn't do it for me. You did it because you understand the enormity of the task that lies ahead. But my main message is to the parents of Trayvon Martin. If I had a son, he'd look like Trayvon. They had their entire lives ahead of them, birthdays, graduations,". Below this is a "Summary:" section with the French translation: "Barack Obama: Je suis tellement amoureux de toi. Ceci est votre victoire. Je sais que tu n'as pas fait ça Juste pour gagner une élection. Et je te connais Je ne l'ai pas fait pour moi. Tu l'as fait parce que tu Comprenez l'énormité de la tâche qui nous attend. Mais mon message principal est au parents de Trayvon Martin. Si j'avais un fils, il ressemble à Trayvon. Ils avaient leur des vies entières devant eux, anniversaires, diplômes, Mariages, enfants à eux. Parmi les morts étaient aussi enseignants, hommes et femmes qui ont consacré leur vie pour aider nos enfants à réaliser leurs". At the bottom, it displays "Transcript Words: 238, Summary Words: 232" and "Efficiency: 97.48%".

Figure 3. Summary of the video in French with efficiency.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

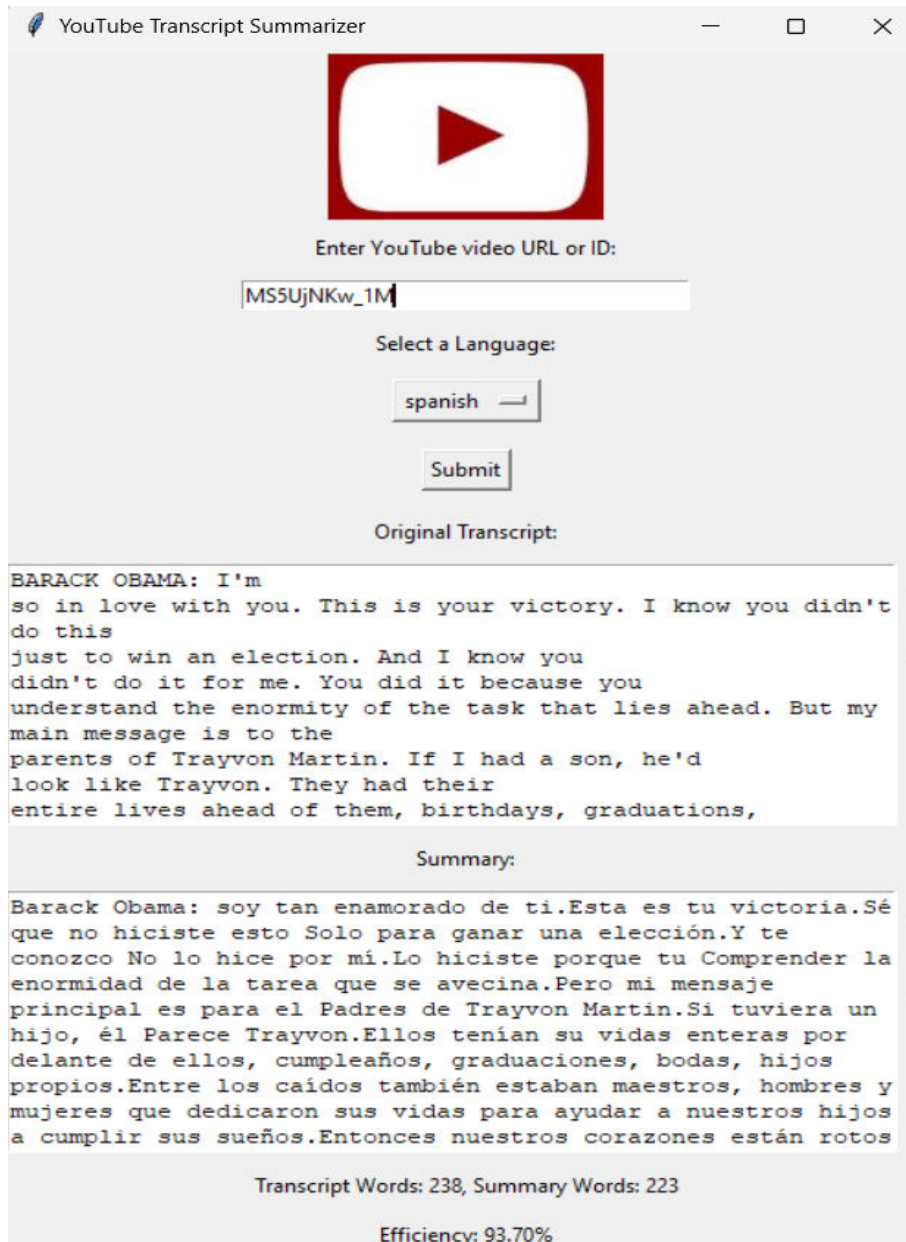


Figure 4. Summary of the video in Spanish with efficiency.

Efficiency

The efficiency of the summarization process was evaluated using a defined input transcript and its corresponding summarized output. Efficiency can be calculated as

$$\text{Efficiency \%} = \frac{\text{Output word count}}{\text{Input word count}} \times 100.$$

This formula provides a measure of how effectively the system condenses the input text while retaining key information. By comparing the word counts of the input and output, the system’s efficiency can be measured in terms of its ability to reduce the overall length of the transcript while maintaining the essential content.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

For example, consider an input transcript with 238 words and the generated summary with 99 words in English. Substituting these values into the formula

$$\text{Efficiency \%} = \frac{99}{238} \times 100 = 41.6\%.$$

This calculation shows an efficiency of 41.6%, meaning that 41.6% of the original transcript content was retained in the summary. This experimental result demonstrates the system's capability to balance content compression with the retention of meaningful information. A higher efficiency reflects less compression, indicating that more of the original content is retained, while a lower efficiency points to a more condensed summary. In Figures 3 and 4, the efficiency levels are not up to the mark for the non-English languages, which require further fine-tuning focusing on the respective languages.

This approach allows for a quantitative evaluation of the summarization technique, providing insights into its effectiveness in various use cases where content reduction and information preservation are equally important.

V. CONCLUSION

This article presents an efficient and user-friendly system for summarizing YouTube video transcripts. By employing state-of-the-art NLP techniques and focusing on accessibility, the system enhances video content consumption for viewers who prefer shorter versions of long videos. The system's ability to translate transcripts into different languages adds to its utility, making it a versatile tool for global users. Future work could explore more advanced models, such as Transformer-based summarization algorithms, and incorporate multimedia elements like video metadata and captions for further content enrichment.

The proposed system successfully reduces the time and effort required to understand video content, providing an accessible and efficient solution for summarizing spoken language from videos. It holds promise for broader applications, including educational platforms, content creation, and media analysis.

ACKNOWLEDGMENTS

We are extremely grateful to Dr. Neelam Sanjeev Kumar, Assistant Professor (S.G.), Department of CSE, SRM Institute of Science and Technology, Vadapalani, for his support and encouragement in leading and helping us to complete our project. We express our humble gratitude to Dr. Akila, Assistant Professor (S.G.), Department of CSE, SRM Institute of Science and Technology, Vadapalani, for providing this opportunity to pursue our project, her valuable guidance, and suggestions.

REFERENCES

1. Atluri, N. S. S. V., Laggiseti, V. S., Jahnvi, D., & Jayanag, B. (2022). Video Transcript Summarizer. 2022 International Mobile and Embedded Technology Conference (MECON). IEEE. <https://doi.org/10.1109/MECON53876.2022.9751991>
2. Kumar, J., Vashistha, R., Lal, R., & Somanir, D. (2023). YouTube Transcript Summarizer. 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE. <https://doi.org/10.1109/ICCCNT56998.2023.10308325>
3. Panthagani, V. B., Duggempudi, V. D. R., Kommera, N. L., & Vakkalagadda, N. (2024). YouTube Transcript Summarizer. 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI). IEEE. <https://doi.org/10.1109/ICMCSI61536.2024.00141>
4. <https://pypi.org/project/youtube-transcript-api/>
5. <https://pypi.org/project/googletrans/>
6. <https://www.datacamp.com/tutorial/discovering-hidden-topics-python>
7. <https://realpython.com/python-gui-tkinter/>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details