International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 08 | Aug 2021 www.irjet.net p-ISSN: 2395-0072

Video Summarization using NLP

Sanjana R[1], Sai Gagana V[2], Vedhavathi K R[3], Kiran K N[4]

[1]Sanjana R, Student, Dept. of EC Engineering, BNMIT, Karnataka, India
[2]Sai Gagana V, Student, Dept. of EC Engineering, BNMIT, Karnataka, India
[3]Vedhavathi K R, Student, Dept. of EC Engineering, BNMIT, Karnataka, India
[4]Kiran K N, Assistant Professor, Dept. of ECE, BNMIT, Karnataka, India

---

Abstract - *This paper proposes an automatic video summarization algorithm using NLP based algorithms. With an increase in internet videos on the video repository platforms like YouTube, Instagram etc. there is an increase in demand for good summarization algorithms to summarize various videos. This paper aims to produce short and concise video summary that summarizes various YouTube videos. The proposed technique first summarizes the YouTube video transcripts based on which summarized video is generated. A web application that takes input as a YouTube video link, and the required summary duration from the user is also developed. Summarized video output is generated and displayed on the web page, after successful processing.*

1. INTRODUCTION

The number of YouTube users in 2020 was approximately 2.3 billion, and has been increasing every year. Every minute, 300 hours of YouTube videos are uploaded. Almost one-third of the YouTube viewers in India access videos on their mobiles and spend over 48 hours a month on the website, a Google study said [11]. It is frustrating and time consuming to search for the videos that contains the information we are actually looking for. For instance, there are many Ted Talk videos available online in which the speaker talks for a long time on a given topic, but it is hard to find the content the speaker is mainly focusing on unless we watch the entire video.

Many machine learning based video summarization techniques are present but they require devices with large processing powers, this is because each video contains thousands of frames and processing all frames takes a very long time. In this paper we propose to use the LSA Natural Language Processing algorithm, which requires less processing power and no training data required to train the algorithm.

2. PROPOSED SYSTEM

The proposed system takes input of a YouTube video link and the time duration to which video has to be summarized as shown in Figure 1. After generating the summarized video using NLP based LSA algorithm, video output is displayed on the web page as shown in Figure 2. For the purpose of illustration, we have considered a Ted talk available on YouTube, and the summary for the Ted talk is displayed.

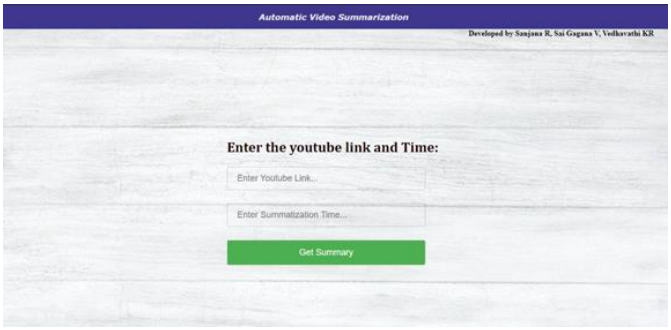

Figure 1: User Input Page

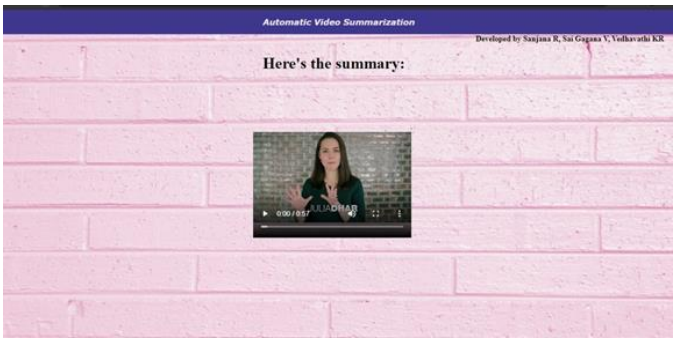

Figure 2: Final Output Page

3. METHODOLOGY

Existing video summarization systems require strong prior technical knowledge. Machine learning based algorithms require high processing power. Summarizing video based on it's subtitle is the fastest way of generating video summary, because dealing with text is easier and faster compared to training various videos using machine learning models.

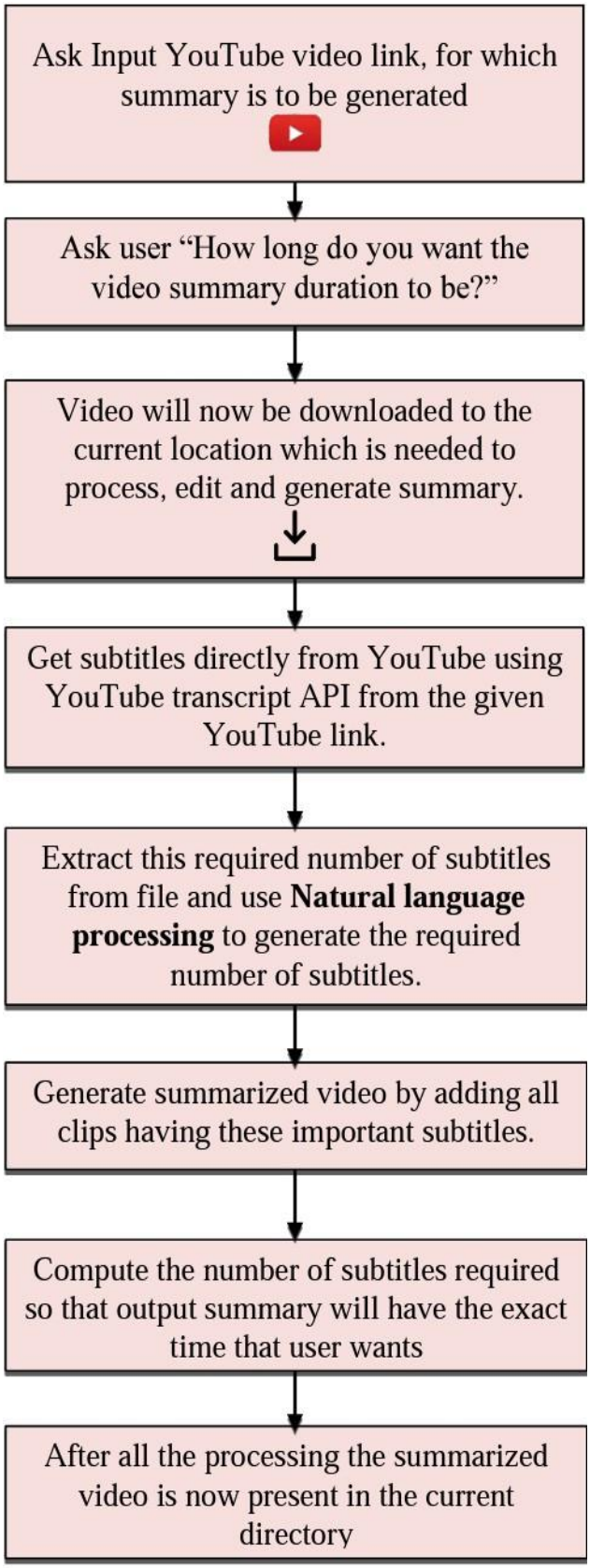

Chart -1: Methodology Chart

3.1. YouTube video

YouTube is the second most visited website worldwide. The range of videos on YouTube includes short films, music videos, feature films, documentaries, audio recordings, corporate sponsored movie trailers, live streams, vlogs, and many other contents from popular YouTubers. YouTube users watch more than one billion hours of video every day. Hence, we have considered YouTube videos as the data for our proposed video summarization algorithm. Using the link, YouTube transcript API will extract subtitles from that particular video.

Downloading videos from YouTube is difficult. To do so first we have to copy the link of the video we want to download then paste the link in the YouTube video downloader website. This method of downloading is time consuming.

Pytube is a lightweight, dependency-free Python library which is used to download YouTube videos easily. This can be achieved with just one or two lines of code.

Pytube library creates the object of the YouTube module by passing a YouTube link of the video as the parameter. Then, it gets the appropriate extension and resolution of the video. Name of the file can be kept based on user convenience. After that, download the file using the download function of pytube library. This download function takes only one parameter: the location where downloaded files need to be saved.

In python URLs are handled using the urllib method, which calls a particular url and handles results after visiting the url. We are using urllib to get the title of a video using the YouTube link.

3.2. LSA Algorithm

Latent Semantic Analysis LSA is an unsupervised approach technique in Natural Language Processing. It is an Algebraic-Statistical method which extracts the features of the sentences that cannot be directly mentioned. These features are essential to data, but are not original features of the dataset.

Working of LSA

1. Term Co-occurrence Matrix:

This matrix is of dimension (vocabulary size) * (vocabulary size). It represents the frequency of the words coming together in the dataset. The matrix helps to understand the words which belong together. Similarity between two different summaries is found out using cosine similarity between the summary matrix.

2. Singular Value Decomposition:

SVD decomposes the matrix into three different matrices: orthogonal column matrix, orthogonal row matrix and one singular matrix.

3. Sentence Selection:

Using the results of SVD different algorithms are used to select important sentences.

3.3. Moviepy python video editor

After the required number of subtitles are generated using the LSA algorithm, these subtitles are used for further processing in the formation of the summarized video. After shortlisting the important subtitles, these subtitles are used to get the time duration during which they were spoken. Once these timings are obtained, we look for that portion of video clip duration in the video. Similarly, we look for all the required subtitles that were shortlisted for the summary and simultaneously concatenate each of the corresponding video clips.

MoviePy library is used for concatenation of required video clips based on subtitles selected for getting the video summary.

RESULTS

Video Link	Total time of video (minutes)	Summary time requested by the user (minutes)	Summary time formed by our algorithm (minutes)
https://www.youtube.com/watch?v=BFZtNN6eNvQ	11	3	2 min 55 sec
https://www.youtube.com/watch?v=2xSKCAtWIyo	17	6	5 min 56 sec
https://www.youtube.com/watch?v=LN06tzw7mb0	15	4	3 min 55 sec

TABLE 1: Metric table of results

Video Title	Total time of video (minutes)	Summary time requested by the user (minutes)	Processing time (minutes)	Memory usage (MB)
How to have constructive conversations | Julia Dhar	11	3	3	190.23
How I Built 5 Income Sources That Make $42,407 Per Month	17	6	10	170.03
Self-Taught Programmer vs Coding Bootcamp vs Computer Science Degree	17	5	4	150.63

TABLE 2: Metric table of Processing time and Memory usage

Table 1 shows the result of the proposed algorithm used to obtain the video summarization using the subtitles of the video. The algorithm gives less than 5 seconds error video as output for the input given.

Table 2 shows the metric table for results with memory usage and processing time. Here the memory usage and processing time results depend on the total time of the input video and the summary time requested by the user.

4. CONCLUSIONS

The increase in popularity of video content on the internet requires an efficient way of representing or managing the video. This can be done by representing the videos on the basis of their summary.

1. Applying Natural Language processing on the subtitles

We have a video with subtitles. We applied an Automatic NLP based LSA summarization algorithm on the subtitle to generate the summary. Basically, we converted the subtitles of the video into a text document and then applied the summarization algorithm. Python library sumy provides the summary for a text document to the number of sentences you specify as argument. There are many summarization algorithms that we can use with the help of this library. But we have used the LSA algorithm.

2. Fitting the duration which user provides

Using the python library sumy, it is possible to rank each sentence (or subtitles in our case). Each subtitle has a certain duration in the video. In order to fit the user duration, we found the average duration of each subtitle by dividing the Total duration of the video with the Number of subtitles.

Using this average duration we have found the approximate number of sentences which we need to produce the summarized video. This summarization technique works in such a way that the top most ranked subtitles are taken into consideration for the final summarized video. If the total duration of the summarized subtitles are more, then it is possible to reduce the one that is least ranked and vice versa. In this way, it is possible to fit the video to the time provided by user.

3. Creating the final summarized video

So now we got the summary of the subtitles and now we have to generate the summarized video. We have used the python module called Moviepy. Using the time stamps in the summarized subtitles we divided the video into several segments and finally merged to create the final summarized video. Hence by following the above steps we were able to generate the video summarization for the given video.

REFERENCES

[1] Jiri Fajtl, Hajar Sadeghi Sokeh, Vasileios Argyriou , Dorothy Monekosso , and Paolo Remagnino: Summarizing Videos with Attention In: Proceedings of the ICLR. Vol. 5 (2019).

[2] Kaiyang Zhou, Yu Qiao, Tao Xiang : Deep Reinforcement Learning for Unsupervised Video Summarization In: Diversity-Representativeness Reward 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org).

[3] Mrigank Rochan, Linwei Ye, and Yang Wang: Video Summarization Using Fully Convolutional Sequence Networks In: International Conference on Learning Representations (2018).

[4] Mayu Otani, Yuta Nakashima , Esa Rahtu , Janne Heikkil¨a , and Naokazu Yokoya : Video Summarization using Deep Semantic Features In: Proc. Advances in Neural Information Processing Systems (NIPS) 2016.

[5] Wang F. and Ngo C.W. Rushes video summarization by object and event understanding. In TRECVID Workshop on Rushes Summarization in ACM Multimedia Conference September 2007.

[6] You J., Liu G., Sun L., and Li H. A multiple visual models based perceptive analysis framework for multilevel video summarization. IEEE Trans. Circuits Syst. Video Tech., 17(3), 2007.

[7] Ngo C.W., Ma Y.F., and Zhang H.J. Video summarization and scene detection by graph modeling. IEEE Trans. Circuits Syst. Video Tech., 15(2):296–305, 2005.

[8] Xu C., Shao X., Maddage N.C., and Kankanhalli M.S. Automatic music video summarization based on audio-visual-text analysis and alignment. In Proc. 31st Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, 2005.

[9] Duan L.Y., Xu M., Chua T.S., Tian Q., and Xu C. A Mid-Level Representation Framework for Semantic Sports Video Analysis. In Proc. 11th ACM Int. Conf. on Multimedia, 2003.

[10] Ferman A.M. and Tekalp A.M. Two-stage hierarchical video summary extraction to match low-level user browsing preerences. IEEE Trans. Multimedia, 5(2):244–256, 2003.

[11] P.Sushma, Dr.S.Nagaprasad, Dr. V. Ajantha Devi. Youtube: Bigdata Analytics using Hadoop and Map Reduce in International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), vol 5, Issue 4, April 2018.