IJIRCCE

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

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 7, July 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007006|

Video Summarization using NLP

Sharannya, Selvi S, Divyashree C M, Sowmya M R, Prof. Yadhukrishna M R

Student, Dept. of IS Engineering, The Oxford College of Engineering, Bangalore, Karnataka, India

Student, Dept. of IS Engineering, The Oxford College of Engineering, Bangalore, Karnataka, India

Student, Dept. of IS Engineering, The Oxford College of Engineering, Bangalore, Karnataka, India

Student, Dept. of IS Engineering, The Oxford College of Engineering, Bangalore, Karnataka, India

Dept. of IS Engineering, The Oxford College of Engineering, Bangalore, Karnataka, India

ABSTRACT -This paper proposes an automatic subtitle generation and semantic video summarization technique using NLP-based algorithms. Automatic video summarization is critical in big data. Video summarization facilitates the efficient storage and rapid browsing of large amounts of video content without losing the important ones. This paper aims to produce a short video summary that summarizes various YouTube videos. A web application that takes input as a YouTube video link and the required summary duration from the user developed. After successful processing, the output displayed on the web page

KEYWORDS: video summarization, NLP-based algorithms, LSA, time duration

I. INTRODUCTION

The number of YouTube users was estimated to be over 2.3 billion in 2020 and has been growing every year, with 300 hours of YouTube videos posted per minute. According to a Google survey [11], nearly one-third of YouTube visitors in India watch videos on their phones and spend more than 48 hours per month on the site. Searching for videos that contain the information we are seeking in, is frustrating and time-consuming. For example, we have many Ted Talk videos available online in which the speaker speaks for an extended time about a specific topic, so finding the content the speaker is most focused on requires watching the entire video. Many machines learning-based videos summarising approaches exist, but they necessitate machines with high processing power, as each movie contains hundreds of frames, and processing them takes a long time. We propose in this study to employ the LSA Natural Language Computing algorithm, which requires less processing resources and no training data.

II. RELATED WORK

In [2] Authors present VideoSET, a method for Video Summary Evaluation through Text that can evaluate how well a video summary is able to retain the semantic information contained in its original video. It is observed that semantics is most easily expressed in words, and develop a text- based approach for the evaluation. Given a video summary, a text representation of the video summary is first generated, and an NLP-based metric is then used to measure its semantic distance to ground-truth text summaries written by humans. It shows that this technique has higher agreement with human judgment than pixel-based distance metrics. They also released text annotations and ground-truth text summaries for a number of publicly available video datasets, for use by the computer vision community[3].authors are proposing an automatic subtitle generation and semantic video summarization technique. The importance of automatic video summarization is vast in the present era of big data. Video summarization helps in efficient storage and also quick surfing of large collection of videos without losing the important ones. The summarization of the videos is done with the help of subtitles which is obtained using several text summarization algorithms. The proposed technique generates the subtitle for videos with/without subtitles using speech recognition and then applies NLP based Text summarization algorithms on the subtitles. The performance of subtitle generation and video summarization is boosted through Ensemble method with two approaches such as Intersection method and Weight based learning method. Experimental results reported show the satisfactory performance of the proposed method.[4] the authors propose a novel method for supervised, keyshots based video summarization by applying a conceptually simple and computationally efficient soft, self-attention mechanism. Current state of the art methods leverage bi-directional recurrent networks such as BiLSTM combined with attention. It proposed a simple, self-attention based network for video summarization which performs the entire sequence to sequence transformation in a single feed forward pass and

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007006|

single backward pass during training.[5]They formulate video summarization as a sequential decision making process and develop a deep summarization network (DSN) to summarize videos. This paper formulates video summarization as a sequential decision making process.. In [6] authors Today there is a huge amount of information from a lot of various resources such as World Wide Web, news articles, e-books and emails. On the one hand, human beings face a shortage of time, and on the other hand, due to the social and occupational needs, they need to obtain the most important information from various resources. Automatic text summarization enables us to access the most important content in the shortest possible time. In this paper a query-oriented text summarization technique is proposed by extracting the most informative sentences.

III. PROPOSED ALGORITHM

The algorithm LSA consists of three major steps:

1. Input matrix creation
2. SVD
3. Sentence selection

Latent Semantic Analysis is a type of semantic analysis that looks for hidden meanings.

In natural language processing, LSA is an unsupervised approach technique. It's a statistical-algebraic method for extracting aspects of sentences that cannot describe directly. These characteristics are necessary for data to function, but they are not unique to the dataset.

Working of LSA:

1. Term Co-occurrence Matrix:
   The dimension of this matrix is (Vocabulary size) * (Vocabulary size). It reflects the frequency with which the terms in the dataset appear together. The matrix aids in the perception of words that belong together. In cosine similarity between the summary, a matrix determines the similarity between two different summaries.
2. The SVD algorithm divides the matrix into three parts: an orthogonal column matrix, an orthogonal row matrix, and a singular matrix.
3. SVD results pick relevant sentences, which are then processed using various methods.

IV. PSEUDO CODE

Step 1: Input YouTube video link
Step 2: Ask user “How long do you want the video summary duration to be?” in minutes
Step 3: Id of this video is used for further processes and video will now be downloaded to the current location
Step 4: Get automatically generated subtitles directly from YouTube using YouTube transcript API from the given link. The transcript is a machine learning based automatically generated using voice recognition.
Step 5: Receive subtitles in Json format and store in a file, it also consists of time stamps of each subtitle.
Step 6: Compute the number of subtitles required for the output summary based on total summary time asked by user. (ex: 5 min video: 200 subtitles)
Step 7: Extract only the required number of subtitles from the json file and use Natural language processing to generate the required number of subtitles.
Step 8: Now use these subtitles to generate summarized video by adding all clips having these important subtitles.
Step 9: The summarized video is now present in the current directory.

V. SIMULATION RESULTS

1. *Subtitles are subjected to Natural Language Processing.*
   A video with subtitles is available. To generate the summary, we used an Automatic NLP-based LSA summarization algorithm on the subtitle. We just turned the video's subtitles into a text document and then ran the summarization algorithm on it. Sumy is a Python package that generates a summary for a text document based on the amount of sentences you provide as an argument. With the help of this library, we may use a variety of summarization algorithms. We have, however, used the LSA algorithm.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007006|

2. *Fitting the duration specified by the user.*

Each sentence can be ranked using the sumy python module (or subtitles in our case). In the video, each subtitle has a specific duration. We determined the average time of each subtitle by dividing the total duration of the video by the number of subtitles in order to fit the user duration. We determined the approximate number of sentences required to create the summarised video using this average time. The highest ranked subtitles are considered for the final summarised video in this summary technique. If the overall time of the summary subtitles is longer, the least ranked one can be cut, and vice versa. It is possible to match the video to the time specified by the user in this way.

3. *Finalizing the summary video*.

So now that we have the subtitles summary, we must create the summarised video. Moviepy is a Python module that we used. We separated the film into various segments and then blended them to create the final summarised video using the time stamps in the summary subtitles. As an outcome, we were able to construct the video summarization for the given video by following the above procedures.

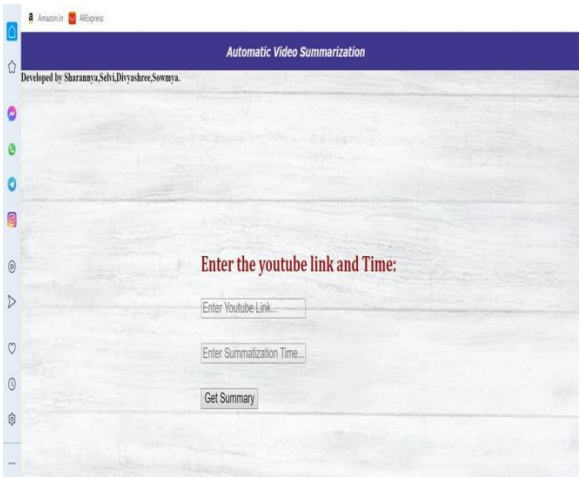

Fig.1. Ad Hoc Network of 5 Nodes

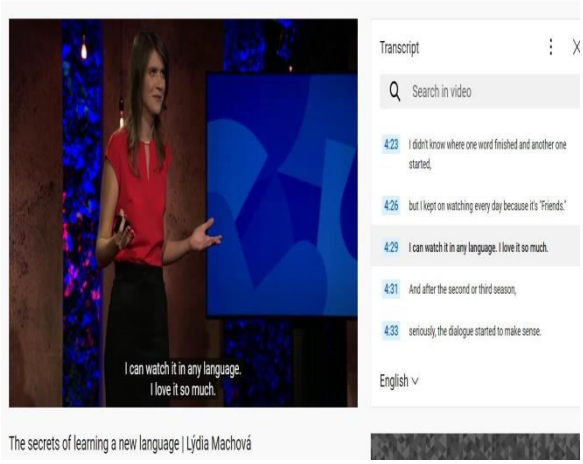

Fig. 2. Energy Consumption by Each Node

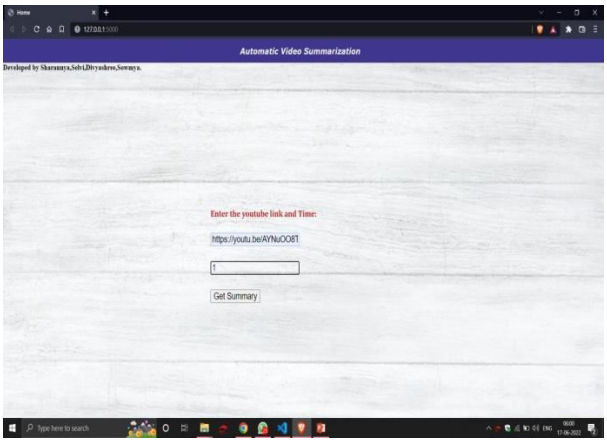

Fig. 3. Ad Hoc Network of 5 Nodes

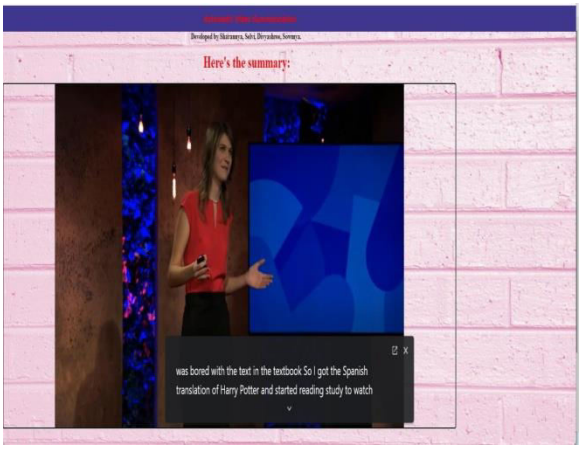

Fig 4. Energy Consumption by Each Node

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007006|

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=BFZiNN6eNvQ	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

TABLE1: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 displays the outcome of the suggested technique for obtaining video summary from the video's subtitle. For the specific input, the algorithm generates a less than 5-second error video as an output.

Table 2 displays the metric table for the outcome, which includes memory use and processing time. The memory use and processing time are determined based on the total duration of the input video and the user-specific summary time.
The growing prevalence of video content on the internet necessitates a more efficient method of expressing or managing it. This can be done by portraying the videos based on their summary.

VI. CONCLUSION AND FUTURE WORK

Since increasing popularity of Internet videos has accelerated the demand for efficient video retrieval, it is very necessary to build a smart preview feature to improve our search results and help our users more efficiently find videos the web. So, it is very important to have video summarization technique for an internet video to provide a quick way to overview its content.

This project is used to automate the creation of video summaries and also to create framework which is used to alter videos to the time duration of user choice.
These techniques are just the start of a new era in deep learning technology when it comes to video summarization. Many advances will be made in the near future to create and optimize the best summaries based on the audience, delivery medium, and intent of summarization. Together, with efforts across the industry, this project aims to make video summarization highly-scalable, reliable, and incredibly efficient.

The precision of the model can be further improved by improving the loss function of the pairwise ranking model to consider a triplet of segments consisting of highlight segments from a video belonging to a certain category, non-highlight segments from the same video, and segments from a video belonging to some other category.

Further this idea can also developed as a performance measurer based on a simulated experiment whose results are easily interpretable (Predicting the success of video on different platforms like Instagram, Facebook etc.)

REFERENCES

1. 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.
2. 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.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007006|

3. 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.
4. 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.
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. 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.
8. 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).
9. Mrigank Rochan, Linwei Ye, and Yang Wang: Video Summarization Using Fully Convolutional Sequence Networks In: International Conference on Learning Representations (2018).
10. YouTube link: https://www.youtube.com/watch?v=8p-kVjmQ0v8 Adobe Summit Sneak,Video Ad Al 2018.
11. Jiri Fajtl, Hajar Sadeghi Sokeh, Vasileios Argyriou , Dorothy Monekosso , and Paolo Remagnino: Summarizing Videos with Attention In: Proceedings of the ICLR. Vol. 5 (2019).
12. YouTube.com VideoAdAI: https://www.youtube.com/watch?time_continue=417&v=Z8zqOpChQnU&feature=emb_logo
13. Gao, X., Tang, X.: Unsupervised video-shot segmentation and model-free anchorperson detection fornews video story parsing. IEEE Transactions on Circuits and Systems for Video Technology 12(9), 765{776 (Sep 2002)
14. Janowski, L., Papir, Z.: Modeling subjective tests of quality of experience with a generalized linear model. In: Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on. pp. 35{40(July 2009)
15. Leszczuk, M., Papir, Z.: Protocols and Systems for Interactive Distributed Multimedia: Joint International Workshops on Interactive Distributed Multimedia Systems and Protocols for Multimedia Systems, IDMS/PROMS 2002 Coimbra, Portugal, November 26{29, 2002 Proceedings, chap. Accuracyvs. Speed Trade-O_ in Detecting of Shots in Video Content for Abstracting Digital Video Libraries, pp. 176{189. Springer Berlin Heidelberg, Berlin, Heidelberg (2002), http://dx.doi.org/10.1007/3-540-36166-9_16
16. Leszczuk, M.I., Duplaga, M.: Algorithm for video summarization of bronchoscopy procedures. BioMedical Engineering OnLine 10(1), 110 (2011), http://dx.doi.org/10.1186/1475-925X-10-110
17. Maybury, M.T., Merlino, A.E.: Multimedia summaries of broadcast news. In: Intelligent InformationSystems, 1997. IIS '97. Proceedings. pp. 442{449 (Dec 1997)\
18. Mu, M., Romaniak, P., Mauthe, A., Leszczuk, M., Janowski, L., Cerqueira, E.: Framework for the integrated video quality assessment. Multimedia Tools and Applications 61(3), 787{817 (2012), http://dx.doi.org/10.1007/s11042-011-0946-3
19. Nawa la, J., Leszczuk, M., Zajdel, M., Baran, R.: Software package for measurement of quality indicators working in no-reference model. Multimedia Tools and Applications pp. 1{17 (2016), http://dx.doi.org/10.1007/s11042-016-4195-3
20. Romaniak, P., Janowski, L., Leszczuk, M., Papir, Z.: Perceptual quality assessment for h.264/avc compression. In: 2012 IEEE Consumer Communications and Networking Conference (CCNC). pp. 597{602 (Jan 2012)
21. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. vol. 1, pp. I{511{I{518 vol.1 (2001)

BIOGRAPHY

Anjum Asma Mohammed is a Research Assistant in the Information Technology Department, College of Computer and Information Sciences, King Saud University. She received Master of Computer Application (MCA) degree in 2005 from BAMU, Aurangabad, MS, India. Her research interests are Computer Networks (wireless Networks), HCI, Algorithms, web 2.0 etc.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

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