Video Transcript Summarization in Marathi

Shivani Patil[1], Swati Yadav[2], Shreya Shinde[3], Darshani Waghmare[4], Rutuja Patil[5], Prof. S. A. Babar[6]

Students, Department of Computer Science & Engineering[1,2,3,4,5]

Assistant Professor, Department of Computer Science & Engineering[6]

Sanjeevan Engineering and Technology Institute Panhala, Maharashtra, India

Abstract: *This paper's purpose is a summarization of the video in Regional Languages. During the procedure, we used methodology NLP, LSA, and MoviePy. This paper aims to produce a Short video of long video without missing any point. The technique first short video of any downloaded video. A web application that takes an input of the video and accuracy of the video, then we get this summaries video into text and this text converted into any Regional language. This paper is going to represent an Extraordinary NLP application. This application benefits Students, and teachers by saving time.*

Keywords: LSA, POS Tagging, Text Recognition, Frame Generation, Tokenization

I. INTRODUCTION

We watch lots of Youtube videos and download videos on our mobile, TV any social media. In this pandemic situation, all peoples are busy with work so, they are not wasted their time watching the news, or any other video. So, We think making an application which is suitable for this person and some students are lazy in their studies. For that difficulties are going to be broken up by summarization. Summarization means to pick out a meaningful and important word from a given paragraph or chapter .this type Firstly, Video to text summary is very easy but some peoples are not interested in reading so, we are making a summary of a long video and creating a short video and audio clips. which can b easily access from it. After the video summary, this video was converted into text for documentation purposes. we can video as well as a document are collected from it. then this text is converted into any regional language for all people.

II. PROBLEM STATEMENT

To create Video transcript summarization for Marathi language

III. LITERATURE SURVEY

[3.1] Cuneyt M. Taskiran, Arnon Amir, Dulce Ponceleon and Edward J. Delp

Developing efficient representations for video browsing presents some unique algorithmic challenges, as well as new technical challenges.Video is a sequential and information-rich medium. It includes audio and motion, and it carries long temporal contextual relationships between shots and scenes. In contrast to images in an image database, manipulation of video is inherently more complex. For example, images can be represented as thumbnails and users can easily judge relevance of these images at a glance. The same task is very time consuming for video sequences, where one hour is composed of more than 100,000 frames, divided into hundreds of shots. Additionally, the audio, which often conveys much of the information (e.g., a video of a talking head accompanied by slides), is even harder to browse in an efficient manner.

[3.2] Kiran Agre, Ankur Chheda, Sairaj Gaonkar, Prof. Mahendra Patil

As the focus is shifting towards YouTube the paper has proposed system which makes it easy for user to access information contained by the text in these video in efficient and quicker way. The proposed system will convert the text in the video into editable form which is stored in a text file.These videos contain text which adds information to videos and makes it more meaningful. If the text from the videos is converted to editable form, it can be stored efficiently and it will be easier to access it next time. Once the user has watched the educational video, next time he may not want to go through the entire video as he has already watched it and reading the main points may be sufficient for him to revise the topic from that video. In such case the proposed system helps user to get access to the information by converting to text in video to editable form

[3.3] Abhay Adapanawar, Anita Garje, Paurnima Thakare, Prajakta Gundawaz, Priyanka Kulkarni

People of different linguistic background could not able tointeract with each other. This concept of translation will helppeople to communicate comfortably. Also it will help to fill communication gap between two linguistically different back-grounds. It will help to the people in the villages, who have taken education of English.Marathi is one of the richest languages among all the languages exist in the world and one of the largely spoken languages in the world. More than 72 million people speak in Marathi as their native language. It is ranked 19th, based on the number of speakers .Marathi is the mother language of India and also a large number of people in southern area of India(Maharashtra) speak and write in Marathi.

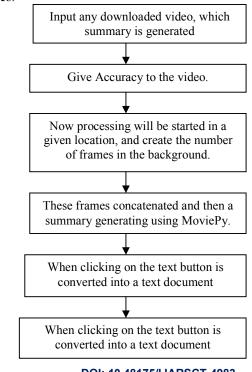**[3.4] Mr. S. B. Chaudhari**

Marathi is the one of popular language in India, Basically from Maharashtra i.e. Mother tongue of state Maharashtra. More than 80% peoples speak this language as their mother tongue. This Language is written from left to right, top to bottom of page. The Marathi words id akin to Sanskrit like mahina " as a maas" and „navin" as a nava". The different linguistic people could not able to interact with other language but they will not able to understand. This concept of translation will helps people to communicate. Also help to fill gap between communications of different linguistic people. It will also helpful who have taken education in English but poor knowledge of Marathi.

IV. PROPOSED SYSTEM

The main idea of the proposed system is that we are going to summarize the video into Marathi text. First step is, the video is converted into frames. At regular time interval the frames are generated so that text in the successive frame is not repeated very often. These frames can be saved in any image format. After that, Text Recognition and Text Extraction applied on every frame .This process gives a summarised video and Translation process is applied on this video and then it will produce the output.

V. METHODOLOGY

For the summarization of the video have strong technical knowledge of NLP. Summarization of text to text is easier because in this audio is not important. but video to video summarization audio is very important. we can't pause or neglect any word in the video.

Input any downloaded video, which summary is generated

↓

Give Accuracy to the video.

↓

Now processing will be started in a given location, and create the number of frames in the background.

↓

These frames concatenated and then a summary generating using MoviePy.

↓

When clicking on the text button is converted into a text document
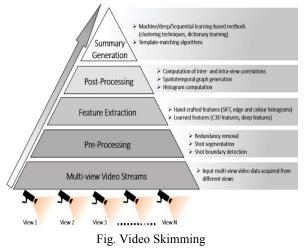
↓

When clicking on the text button is converted into a text document

5.1 LSA Algorithm

Latent Semantic Analysis is an algebraic statistical method that extracts the meaning of the words and the similarity of the sentence using the information about the usage of words in the context. find out the proper meaning of the word in sentences.
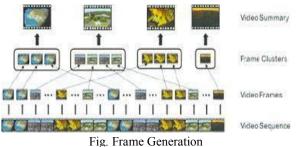
5.2 Video Skimming

Video skimming is used for segmentation. In the video find which word is repeated mostly that word will be important and find the unwanted words in the video.

Fig. Video Skimming

5.3 Frame Generation

In this step, the video is converted into frames. Frames are images of a particular time in avideo. At regular time intervals, the frames are generated so that text in the successive frame is not repeated very often. These frames can be saved in any image format.

Fig. Frame Generation

5.4 Text Recognition

This step is applied to every frame. In this step, the text region is detected using an algorithm described in the next part. The detected text regions are then refined to increase the efficiency of extracting text.
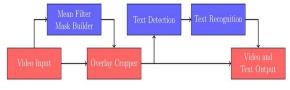
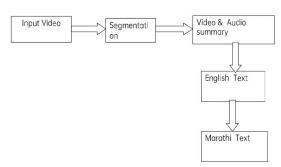

Fig. Text Recognition

Figure: System Architecture

VI. MODULE

6.1 Video to Video Summary

In this module, we create a short video summary of a long video. this is important for those people which are saving their time by watching a video. when we input the video processing will be started and the video is divided into frames and then we concatenate these frames into a single video.

6.2 Video to Text

For the students, a document file is important because always we can't access the video and write our notes for that purpose we are creating a text of that video. This text summary is saved in a word file. This text is in the English language which is understood by all people.

6.3 English to Regional Language

Some people did not understand the English language they only understood their regional language for that purpose we converted this English text into any regional language. For example, Marathi Hindi, Urdu, Bengali, Telugu

VII. CONCLUSION

In this paper, we purpose to convert long video summaries into video and audio. Then this video is converted into text and then regional language. For that, we used python different Libraries. Our purpose is to develop students by using our project and approach to the student on an online learning system. Because people conclude that mobile learning is harmful to the student. Saving the time for busy people and students .

REFERENCES

[1]. Automated Video Summarization Using Speech Transcripts Cuneyt M. Taskiran, Arnon Amir, Dulce Ponceleon, and Edward J. Delp

[2]. M. G. Christel, M. A. Smith, R. Taylor, and D. B. Winker, "Evolving video skims into useful multimedia abstractions," Proceedings of the ACM Computer-Human Interface Conference (CHI'98), April 18-23 1998, Los Angeles, CA, pp. 171–178

[3]. S. Srinivasan, D. Ponceleon, A. Amir, B. Blanchard, and D. Petkovic, "Engineering the web for multimedia," Proceeding of the Web Engineering Workshop (WEBE), WWW-9, May 15-19 2000, Amsterdam, Netherlands.

[4]. S. Archana AB, Sunitha. C ,"An Overview on Document Summarization Techniques" ,International Journal on Advanced Computer Theory and Engineering (IJACTE) ,ISSN (Print) : 2319 ″U 2526, Volume-1, Issue-2, 2013 .

[5]. Rafael Ferreira ,Luciano de Souza Cabral ,Rafael DueireLins ,Gabriel Pereira e Silva ,Fred Freitas ,George D.C. Cavalcanti ,Luciano Favaro , "Assessing sentence scoring techniques for extractive text summarization ",Expert Systems with Applications 40 (2013) 5755-5764 ,2013 Elsevier .

[6]. L. Suanmali , N. Salim and M.S. Binwahlan,"Fuzzy Logic Based Method for Improving Text Summarization" , International Journal of Computer Science and Information Security, 2009, Vol. 2, No.

1,pp. 4-10.

[7]. Mrs.A.R.Kulkarni , Dr.Mrs.S.S.Apte "A DOMAIN-SPECIFIC AUTOMATIC TEXT SUMMARIZATION USING FUZZY LOGIC ",International Journal of Computer Engineering and Technology (IJCET), ISSN 0976- 6367(Print), ISSN 0976 - 6375(Online) Volume 4, Issue 4, July-August (2013).

[8]. "An improvised Morphological Analyzer for Tamil:A case of implementing the open source platformApertium". Parameswari K. Unpublished M.Phil. Thesis. Hyderabad:University of Hyderabad. 2009.

[9]. " Design and Implementation of a Morphology-based Spellchecker forMarathi, an Indian Language". Dixit, Veena, Satish Dethe, and Rushikesh K. Joshi. 2006.

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