Analysis of a video to find its key elements

Shaurya Aggarwal[1], Jitesh[2], Shivani Verma[3], Vaibhav Saxena[4], Shailendra Gaur[5]

[1]*Bhagwan Pershuram Institute of Technology*
[2]*Bhagwan Pershuram Institute of Technology*
[3]*Bhagwan Pershuram Institute of Technology*
[4]*Bhagwan Pershuram Institute of Technology*
[5]*Bhagwan Pershuram Institute of Technology*

(shauryaa06@gmail.com[1],one.jitesh@gmail.com[2],vermashivani052@gmail.com[3],vbhv.300@gmail.com[4] shailendra.gaur08@gmail.com[5])

Abstract-- During the previous couple of years, we have seen an uncommon development of User Generated Videos (UGV).Being ready to efficiently and viably process unedited recordings has turned out to be progressively essential.
In this project, we expect to improve video understanding by summarizing the central and most fascinating parts out of recordings. An extensive sum exertion is spent in separating such features from the respective video requiring a great deal of speculations. The experts of our domain decide that which frames should be selected which is a time taking and costly process, but out project do the same in the most efficient and viable manner.

***Keywords*:** Navigation, Google chart, Xampp, CMS.

I. INTRODUCTION

With an explosion in multimedia content production due to ever increasing reach of internet, content in the form of videos is becoming increasingly common place, becoming the preferred method for the purpose of delivering content. Advent of social media and growth of video sharing websites, in particular YouTube, has only contributed to the increasing importance of video graphic content. More content is uploaded to YouTube a day, than a person is capable of watching in his/her whole lifetime. With the emergence of video content as an effective mode of information propagation, the increased automation of the video summarization process has become paramount. In recent times, video summarization has emerged as a challenging issue in the field of machine learning, which intends to automatically evaluate a video's content and help generate a summary with the video's most quality content. Video brief summary finds applications to generate event highlights like sports, trailers for movies and in general shortening video to the most relevant subsequence that Enabling people to efficiently browse large video repositories.

Video quick summary is a difficult and challenging issue today in several ways. There is no natural ordering of video summaries. Among a given set of video summaries, the best representation of the original video is highly subjective. The general objective in modern literature for video summarization aims at producing a summary, typically 5%-15% of the whole video, consisting of the most informative content from the original video. The content is usually represented in the form of key frames or more appropriately as video skims. A good video summary depicts the synopsis of the original video, in a compact way depicting all important and relevant scenes/shots. Through this project, we review the major techniques in video summarization, and look at their performance on some recent datasets.

II. RELATED WORK

A. Input Video

The format of the video can be AVI (Audio Video Interleaved), MP4, MPEG (Motion Picture Experts Group), MKV (Mittelschüler-Kartellverband) etc. These are the video formats that can be used while making the highlights of any video. After the input of the video, frame extraction takes place[2].

B. Frame Extraction

Since the original video (used for input) consists of 800 frames in 30 seconds and the output video (highlighted) consists of 100 frames in 3 seconds. Our project can make the highlight video of only one minute, that is, 2000 frames to be exact. Now the original video is extracted or divided into the frames which will further be used in highlighted video[2].

C. Feature Extraction

The first part of our model creates include vectors utilizing highlights extricated from recordings. One popular feature learning algorithm is Sumy Algorithm to extract text summarization. It is developed by Miso-Belica which stops the individual to use the algorithms by yourself, but the

legendary Luhn to Edmundson who provides this library to perform extractive summarization in an easy way.
This Sumy technology has further provided the wide range of algorithms named: Lex Rank, Luhn, LSA, TextRank.

D. Key Frame Extraction

Key frame extraction from video data is an active research issue in the recognition and retrieval of video objects. Key frame refers to the video sequence image frame that is representative and capable of reflecting a video content summary.
By using the key frame, the main content of video data can always be clearly specified and the amountof memory required for video data process ing and complexity significantly reduced[1].

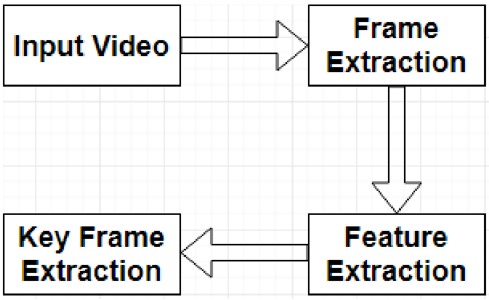

Figure 1. Flow chart for the video summarization process [1]

E. Highlight Classification

Numerous early feature discovery works center around specific classifications of entertainment recordings including stand up comedies, T.V series, conference recording including Ted Talks, Titu Talks. These models for the most part depend on low level visual highlights. As of late, a couple of techniques have been proposed to find features in non-exclusive individual recordings.
As per this feature, the video fragments that are relied upon to energize the users the most. In this paper we endeavor to take care of this issue via naturally learning visual highlights which used Sumy technology.

III. RELEVANT WORK

Video wrap-up has picked up a great deal of consideration lately. The assignment of video rundown can be practiced in fundamentally two different ways. A ton of the underlying work in video wrap-up has been done in the area of unsupervised learning, yet the ongoing pattern has moved towards gaining from how people create video synopsis, driving it into the supervised domain. A short audit of the work done in the video outline is done as per the general classification above.

IV. DATA SET

We use the same user-generated video dataset as Zeng el al used. In the paper "Title Generation for User Generated Videos," they train a highlight detector for their video captioning task as an intermediate step. Each video in the dataset comes with a hand marked interval of the highlight start and end frame in this video. A video with 1000 frames, for example[3].
We pre-process the dataset by filtering out unusually long or short videos and videos with corrupted data. Our final dataset contains 3687 videos (around 3900000 frames). By and large, the middle video length is around 800 casings, or 30 seconds, and the middle feature length is around 100 edges, or 3 seconds. Next, we break every video into windows of 64 outlines. A window is viewed as a feature if over 75% of the casings in this window is inside the feature interim, i.e., in excess of 48 edges of the window is inside the feature interim. At last, we have around 180000 windows altogether. Preferably, we might want to store every one of the windows of a video as a solitary unit and arrange the windows at the same time incorporate the sitemap and a stream of different procedure. Since our project can make the highlight video of 1 minute, that is, 2000 frames in total In the paper "Title Generation for User Generated Videos," they train a highlight detector for their video captioning task as an intermediate step. Each video in the dataset comes with a hand marked interval of the highlight start and end frame in this video. A video with 1000 frames, for example [3].

V. METHODS

Our entire model is based on Sumy based technology which is further classified into 4 algorithms:

- Lex Rank
- Luhn
- Latent Semantic Analysis (LSA)
- Text Rank

A. Lex Rank

This is a graphical based text summarizer. It is a Simple library and command line utility for extracting summary from HTML pages or plain texts. The package also includes a simple framework for evaluating text summaries.

B. Luhn

It is one of the earliest advised algorithm via the famous IBM researcher it became named after. It scores sentences based totally on frequency of the most essential phrases.
The LUHN formulation became created within the past due Sixties or in the late 1960s by means of a collection of mathematicians.
Since the algorithm is in the open area, it tends to be utilized by anybody. Most Master cards and numerous administration and ID numbers utilize the calculation as a straightforward strategy for recognizing legitimate numbers from mistyped or generally erroneous numbers.

C. Latent Semantic Analysis (LSA)

Latent semantic analysis is an unmanaged summary method. It is also called as LSI (Latent Semantic Indexing) which is a retrieval approach that makes use of a mathematical method referred to as Singular Value Decomposition (SVD) to pick out the patterns within the relationships among the terms and the principles contained in an unstructure text collection. It is developed by Jean-Paul Benzecri in early 1970s.

LSA can correlate semantically related terms that are latent in a collection of text, uncover the underlying latent semantic structure in the usage of words in a body of text and how it can extract the meaning of the textual content in response to the person's queries, typically called conceptual searches.

- Singular Value Decomposition

$$M = U \sum V* \qquad (1)$$

where:

- M is mxm matrix.
- U is mxn left singular matrix.
- $\sum$ is nxn diagnol matrix with non−negative real numbers.
- V is mxn right singular matrix
- V* is nxm matrix, which is a transpose of V.

D. Text Rank

Text rank is a summary technique based on a graph with extractions of keywords from a document. Text Rank works as follows:

- Pre-process the content: evacuate stop words and stem the rest of the words.
- Make a diagram where vertices are sentences.
- Associate each sentence to each other sentence by an edge. The heaviness of the edge is the way comparable the two sentences are.
- Run the PageRank calculation on the diagram.
- Pick the vertices (sentences) with the most noteworthy PageRank score

VI. WORKING

In this project we are working on summarizing a video or making the highlights of a video. So in this one can make the highlight video of maximum one minute. So there are some screen-shots of the highlights of a video which will make one understand that how this project is working.

[5]

[5]

[5]

[5]

VII. FIGURES

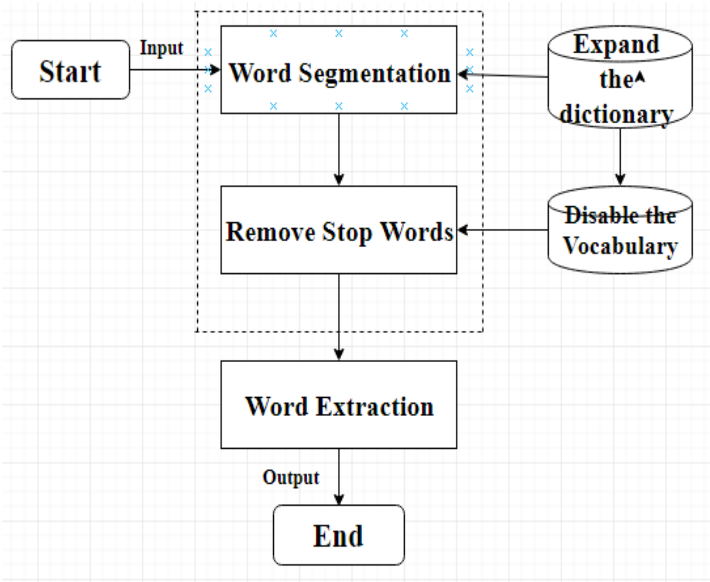

Figure 2: Example of a Graph Based Lex Rank.

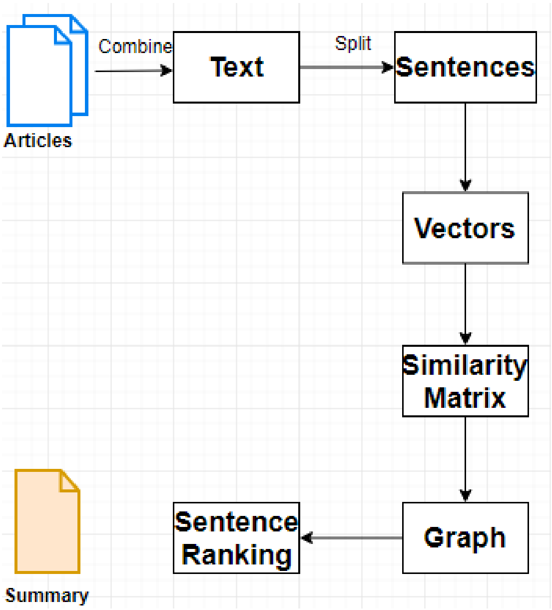

Figure 3: Example of a Text Summarization by Text Rank.

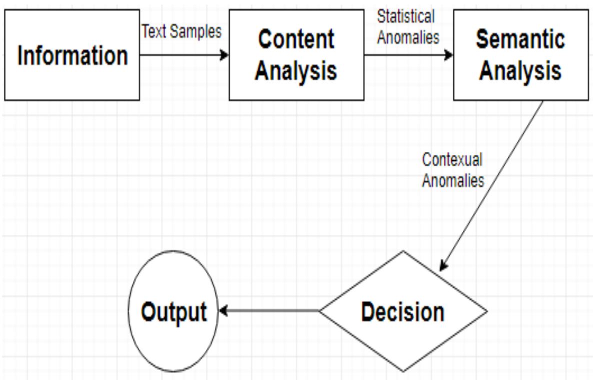

Figure 4: Process of Latent Semantic Analysis.

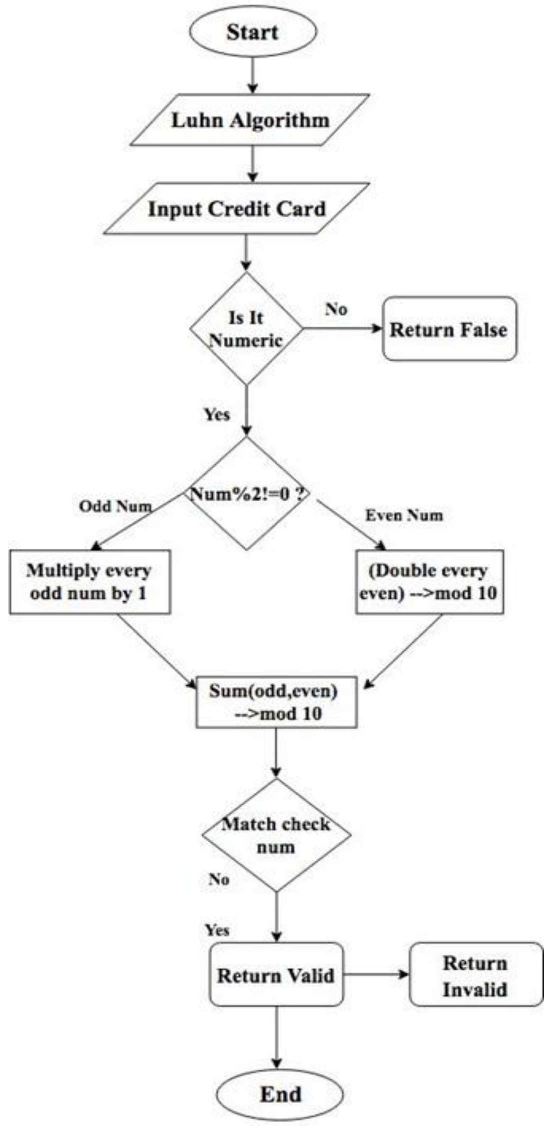

Figure 5. Luhn algorithm

VIII. IMPLEMENTATION

A. Lex Rank

```
from sumy.summarizers.lex_rank
import LexRankSummarizer
summarizer = LexRankSummarizer()

summary = summarizer (parser.documnent.document,2)

for line in summary:
 print (line)[4]
```

B. Luhn

```
from sumy.summarizers.luhn
import LuhnSummarizer
summarizer_1=LuhnSummarizer()
summary_1 = summarizer_1(parser.document,2)
```

```
for line in summary_1:
 print(line)[4]
```

C. Latent Semantic Analysis (LSA)

```
from sumy.summarizers.lsa
import LsaSummarizer
summarizer_2=LsaSummarizer()
summary_2 = summarizer_2(parser.document,2)

for line in summary_2:
 print(line)[4]
```

D. Text Rank

```
from sumy.summarizer.text_rank
import TextRankSummarizer
summarizer_3=TextRankSummarizer()
summary_3 = summarizer_3(parser.document,2)
for line in summary_3: print(line)[4]
```

IX. RESULT

As a result of this project, we extract features for video summarization, we delved into the basics of computer vision and finding features of interest using popular techniques including Sumy algorithm. These techniques included Sumy and its further classified algorithm for feature extraction. As a result,this project is utilizing a typical vault of code and working in a group by sharing and utilizing secluded code pieces all through the task.

Libraries/Tools Used:

- Pycharm
- Imageio
- Moviepy
- Pytube
- Sumy
- Pysrt

X. FUTURE PLANS

We intend to complete an increasingly point by point examination of Sumy calculation on the SumMe benchmark, assessing the execution of the calculations dependent on their classes. We likewise plan to actualize different calculations, especially those dependent on learning "rundown" utilizing a blend of shallow and profound highlights in the regulated setup. Additionally, we intend to devise a calculation, conceivably a group or a variation of the calculations we have investigated, and assess its execution on the SumMe benchmark.

Later on, a blend of specialists model could be prepared to broaden the scope of recordings that the model can manage. We likewise need to try different things with consideration display both crosswise over picture pixel measurement and crosswise over time measurement. Propelled by Yeung el al, we would love to join support learning and non-uniform and henceforth effective video preparing in our model too.

XI. CONCLUSION

In this project, we executed a strategy to transiently find features in a different recordings occasion by breaking down the gathering of people conduct. Through broad hyper parameter seek, we accomplish fair execution on our video feature arrangement assignment utilizing the Sumy calculation. Considering the assorted variety of points and styles of client created recordings and the subjectivity of video features, we trust our model has accomplished acceptable outcomes. This demonstrates understanding the connection between successive windows in a video boosts the execution, as is normal. In addition, we attempt a few regularization strategies to diminish over-fitting.

We imagine that this venture has a ton of real use cases. We get a part of positive input on the model yields, which appears that it has potential to help individuals better comprehend recordings.

REFERENCES

[1] Key-frame extraction by analysis of histograms of video frames using statistical methods Sheena C Va , N. K. Narayanan.

[2] A Survey of Automatic Video Summarization Techniques Sachan Priyamvada Rajendra, Dr. Keshaveni N.

[3] Spotlight: A Smart Video Highlight Generator Jun-Ting (Tim) Hsieh, Chengshu (Eric) Li, Wendi Liu , Kuo-Hao Zeng.

[4] https://medium.com/@ondenyi.eric/extractive-text-summarization-techniques-with-sumy-3d3b127a0a32

[5] https://sites.google.com/site/vsummsite/