Unaided Video Search Engine

Simmar Kalsi[1], Harsh Kanzariya[2], Mandar Ganjapurkar[3]

[1,2]*Excelssior Education Society's, KC College of Engineering and Management Studies and Research, Kopri, Thane (East), Mumbai, Maharashtra, India*

[3]*Assistant Professor, KC College of Engineering and Management Studies and Research, Kopri, Thane (East), Mumbai, Maharashtra, India*

***Abstract*- There has been a tremendous growth in digital media, because of the focus shift to visual content and video broadcasting. The increase in the bandwidth and abundant availability of mobile machines has resulted in lag-free billions of hours of videos being watched on various video search platforms. Publishing a video on the internet is the easiest way of conveying one's thoughts, emotions and company products to the world. Internet being cheap and quick, many commercial organisations are emerging with their presence on video search platforms where they publish and earn money though the traffic fetched by the video, also it is easy for them to link it with organisation's or individual's main website & other social media.**

***Keywords*- keyword-based search, Py Scene Detect, FFmpeg, LSTM, Image captioning, image vector**

I. INTRODUCTION

Video currently being closest to the actual representation of the real world, as it provides both visual and auditory information in the most fathomable way possible. We humans have started to use it as a means of communication language via the help of the Internet. Video streaming platforms such as YouTube, Twitch, any more where the widespread population is allowed to Showcase their views and earn money have grown exponentially in popularity. Organizations from one corner of the planet are now reaching out to the consumer on a totally different side of the world. Having access to the internet opens up the land of opportunity to users to the fullest extent. Which includes making use of it for mankind's good as well as personal gains. The problematic part being there is no universal system for restriction of making bad use of the available resources. The available resources being other individuals, organizations their time and money.

Platforms for video streaming have made it their business to provide users with content related to every possible stream (Science, Technology, Engineering, Math) to name a few. All the streaming platforms use advertisements for revenue generation. Money is on the stake users try to manipulate the system for personal gains. Video is generally a knowledge base of audio and visual content. Thus, it is impossible to show all the content at a glance. To tackle this problematic part, Video title, Thumbnail, and description are used alongside to convey brief information about the video.

The title of video and thumbnail is the most useful part alongside description to reach maximum users. Using these added information malpractices are performed to exploit and increase the misinformation. Adding a thumbnail, and writing a title for the uploaded video is done by the consumer/user only. A description is written so that video consumer is directed towards the organization's main website. Which in turn might lead to a scam in the worst-case scenario.

The video title is used by every single search engine to provide the results to search query fired by the user. Description and thumbnail are supportive systems to get a more accurate result. As human gets attracted to a catchy thumbnail over a title. No matter how much you improve the search result, most of the time catchy thumbnail wins. Now There have been multiple different efforts taken by platforms to provide more accurate information to end-user. One of the major problems is how much video content owner has control over its influences even after dozens of algorithms try to minimize the spread of inappropriate.

The readily available technology is the main driving force behind this growing trend. Deep fake is an upcoming technology which if not taken care of properly would be disastrous. Stopping this practice is not an easy task, but it is achievable. Having an unbiased system or a person who verifies or in the best case generating supportive information is the way forward. Thus, to tackle this rapidly increasing problem, available technology should be used to the fullest. The method proposed in the paper helps in minimizing the spread of eye-catchy and misleading videos that are there to help the creator of the video to gain the fortune. This is done by extracting the actual information from the video using scene detection and frame extraction from video by scene detection and video processing using python and FFmpeg. Also, audio is also extracted and used to keep information synced to help users with better search results.

II. METHODOLOGY

A. Video processing

A scene is generally thought of as the action in a single location and continuous time. Generally, a video is made up of many different scenes, containing different characteristics and objects. A scene is used to display and highlight a particular entity in video e.g. A kid playing with a ball, and the scene is changed to divert viewers' attention to different things. Thus, a scene is easy enough to detect by humans. But the machine requires a special ability to carry out such tasks.

There are two main detection methods Py Scene Detect uses: detect-threshold (comparing each frame to a set black level, useful for detecting cuts and fades to/from black), and detect-content (compares each frame sequentially looking for changes in content, useful for detecting fast cuts between video scenes, although slower to process. The threshold-based mode is what most traditional scene detection programs use, which looks at the average intensity of the current frame, triggering a scene break when the intensity falls below the threshold (or crosses back upwards). The content-aware scene detector finds areas where the difference between two subsequent frames exceeds the threshold value that is set. The threshold-based scene detector (detect-threshold) is how most traditional scene detection methods work (e.g. the FFmpeg black frame filter), by comparing the intensity/brightness of the current frame with a set threshold, and triggering a scene cut/break when this value crosses the threshold. In PyScene Detect, this value is computed by averaging the R, G, and B values for every pixel in the frame, yielding a single floating-point number representing the average pixel value (from 0.0 to 255.0).

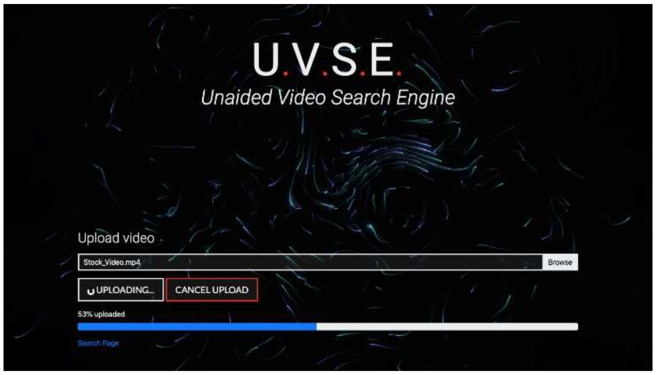
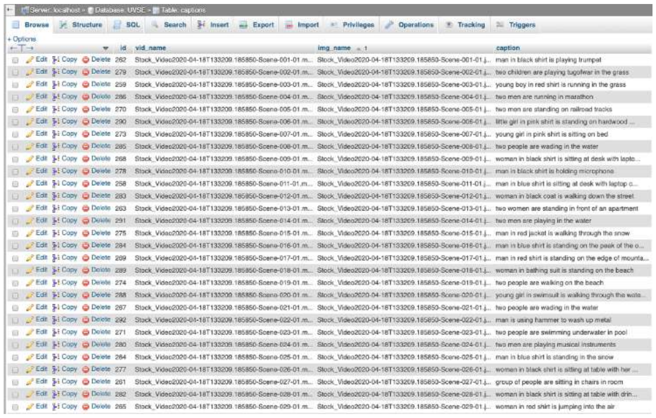


Fig. 1Home page for uploading Video

B. *Caption Generation and Database generation*

Once we receive an image from the scene it is up next for caption generation, We have implemented a model with two inputs i.e. image feature vector and partial caption and trained on 30,000 images obtained from the Flickr website that do not contain any famous person or place so that the entire image can be learned based on all the different objects in the image. To generate the image vector, we opt for transfer learning using the InceptionV3 model. 2048 length vector is extracted using automatic feature engineering. Every word in the training caption set is indexed and mapped to a 200-long vector using the GloVe model. The output generated is an appropriate word that transpires the sequence of the partial caption provided. This is done recursively till the end tag or max length limit is reached. Once all the captions are generated,they are saved in SQL DB.

```
for location in images_paths:
    image_name = location[25:]
    print(image_name)
    vid_name_temp = image_name[:-7]
    vid_name = vid_name_temp + ".mp4"
    print(vid_name)
    img_caption = disp_caption(location)
    print(img_caption)
    sql = "INSERT INTO captions (vid_name,
        img_name, caption) VALUES (%s, %s, %s)"
    val = (vid_name, image_name, img_caption)
    mycursor.execute(sql, val)
    mydb.commit()
```



	id	vid_name	img_name	caption
Edit Copy Delete	262	Stock_Video2020-04-18T133209.185850-Scene-001-01.m...	Stock_Video2020-04-18T133209.185850-Scene-001-01.j...	man in black shirt is playing trumpet
Edit Copy Delete	279	Stock_Video2020-04-18T133209.185850-Scene-002-01.m...	Stock_Video2020-04-18T133209.185850-Scene-002-01.j...	two children are playing tugofwar in the grass
Edit Copy Delete	259	Stock_Video2020-04-18T133209.185850-Scene-003-01.m...	Stock_Video2020-04-18T133209.185850-Scene-003-01.j...	young boy in red shirt is running in the grass
Edit Copy Delete	286	Stock_Video2020-04-18T133209.185850-Scene-004-01.m...	Stock_Video2020-04-18T133209.185850-Scene-004-01.j...	two men are running in marathon
Edit Copy Delete	270	Stock_Video2020-04-18T133209.185850-Scene-005-01.m...	Stock_Video2020-04-18T133209.185850-Scene-005-01.j...	two men are standing on railroad tracks
Edit Copy Delete	290	Stock_Video2020-04-18T133209.185850-Scene-006-01.m...	Stock_Video2020-04-18T133209.185850-Scene-006-01.j...	little girl in pink shirt is standing on hardwood ...
Edit Copy Delete	273	Stock_Video2020-04-18T133209.185850-Scene-007-01.m...	Stock_Video2020-04-18T133209.185850-Scene-007-01.j...	young girl in pink shirt is sitting on bed
Edit Copy Delete	285	Stock_Video2020-04-18T133209.185850-Scene-008-01.m...	Stock_Video2020-04-18T133209.185850-Scene-008-01.j...	two people are wading in the water
Edit Copy Delete	268	Stock_Video2020-04-18T133209.185850-Scene-009-01.m...	Stock_Video2020-04-18T133209.185850-Scene-009-01.j...	woman in black shirt is sitting at desk with lapto...
Edit Copy Delete	278	Stock_Video2020-04-18T133209.185850-Scene-010-01.m...	Stock_Video2020-04-18T133209.185850-Scene-010-01.j...	man in black shirt is holding microphone
Edit Copy Delete	258	Stock_Video2020-04-18T133209.185850-Scene-011-01.m...	Stock_Video2020-04-18T133209.185850-Scene-011-01.j...	man in blue shirt is sitting at desk with laptop c...
Edit Copy Delete	283	Stock_Video2020-04-18T133209.185850-Scene-012-01.m...	Stock_Video2020-04-18T133209.185850-Scene-012-01.j...	woman in black coat is walking down the street
Edit Copy Delete	263	Stock_Video2020-04-18T133209.185850-Scene-013-01.m...	Stock_Video2020-04-18T133209.185850-Scene-013-01.j...	two women are standing in front of an apartment
Edit Copy Delete	291	Stock_Video2020-04-18T133209.185850-Scene-014-01.m...	Stock_Video2020-04-18T133209.185850-Scene-014-01.j...	two men are playing in the water
Edit Copy Delete	275	Stock_Video2020-04-18T133209.185850-Scene-015-01.m...	Stock_Video2020-04-18T133209.185850-Scene-015-01.j...	man in red jacket is walking through the snow
Edit Copy Delete	284	Stock_Video2020-04-18T133209.185850-Scene-016-01.m...	Stock_Video2020-04-18T133209.185850-Scene-016-01.j...	man in blue shirt is standing on the peak of the o...
Edit Copy Delete	269	Stock_Video2020-04-18T133209.185850-Scene-017-01.m...	Stock_Video2020-04-18T133209.185850-Scene-017-01.j...	man in red shirt is standing on the edge of mounta...
Edit Copy Delete	289	Stock_Video2020-04-18T133209.185850-Scene-018-01.m...	Stock_Video2020-04-18T133209.185850-Scene-018-01.j...	woman in bathing suit is standing on the beach
Edit Copy Delete	274	Stock_Video2020-04-18T133209.185850-Scene-019-01.m...	Stock_Video2020-04-18T133209.185850-Scene-019-01.j...	two people are walking on the beach
Edit Copy Delete	288	Stock_Video2020-04-18T133209.185850-Scene-020-01.m...	Stock_Video2020-04-18T133209.185850-Scene-020-01.j...	young girl in swimsuit is walking through the wate...
Edit Copy Delete	267	Stock_Video2020-04-18T133209.185850-Scene-021-01.m...	Stock_Video2020-04-18T133209.185850-Scene-021-01.j...	two people are wading in the water
Edit Copy Delete	292	Stock_Video2020-04-18T133209.185850-Scene-022-01.m...	Stock_Video2020-04-18T133209.185850-Scene-022-01.j...	man is using hammer to wash up metal
Edit Copy Delete	271	Stock_Video2020-04-18T133209.185850-Scene-023-01.m...	Stock_Video2020-04-18T133209.185850-Scene-023-01.j...	two people are swimming underwater in pool
Edit Copy Delete	280	Stock_Video2020-04-18T133209.185850-Scene-024-01.m...	Stock_Video2020-04-18T133209.185850-Scene-024-01.j...	two men are playing musical instruments
Edit Copy Delete	264	Stock_Video2020-04-18T133209.185850-Scene-025-01.m...	Stock_Video2020-04-18T133209.185850-Scene-025-01.j...	man in blue shirt is standing in the snow
Edit Copy Delete	277	Stock_Video2020-04-18T133209.185850-Scene-026-01.m...	Stock_Video2020-04-18T133209.185850-Scene-026-01.j...	woman in black shirt is sitting at table with her ...
Edit Copy Delete	261	Stock_Video2020-04-18T133209.185850-Scene-027-01.m...	Stock_Video2020-04-18T133209.185850-Scene-027-01.j...	group of people are sitting in chairs in room
Edit Copy Delete	282	Stock_Video2020-04-18T133209.185850-Scene-028-01.m...	Stock_Video2020-04-18T133209.185850-Scene-028-01.j...	woman in black shirt is sitting at table with drin...
Edit Copy Delete	265	Stock_Video2020-04-18T133209.185850-Scene-029-01.m...	Stock_Video2020-04-18T133209.185850-Scene-029-01.j...	woman in red shirt is jumping into the air

Fig. 2 Database

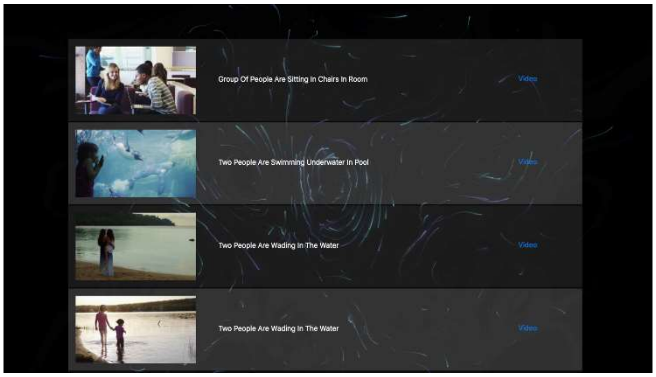
C. *Search and Result Display*

All the processing on video is done by the system even before any content consumer user checks in. Now, whenever any user searches for particular video content,

```
sql_query = f"SELECT id, vid_name, img_name, caption, MATCH(caption) AGAINST('{search_query}' IN NATURAL LANGUAGE MODE) AS score FROM captions ORDER BY `score` DESC LIMIT 10"
```

The above SQL query looks for a caption saved for every clip and returns the video result with an image as a thumbnail. The system also shows relative searches with a range from 0 to 100 percent.

Fig. 3 Search page



Group Of People Are Sitting In Chairs In Room

Two People Are Swimming Underwater In Pool

Two People Are Wading In The Water

Two People Are Wading In The Water

Fig. 4 Output page

III. CONCLUSION

The proposed system takes video as an input, split the scene, and further extracts an informative frame which is further given for caption generation. Once all captions are generated, they are used for searching the video content and scene rather than just using a video title to assume the content of the video. This system can be used as a service along with any other currently present system to provide even more accurate results to the user.

ACKNOWLEDGMENT

We would like to thank our guide Asst. Prof. (Mr) Assistant Professor Mandar Ganjapurkar, K. C College of Engineering & Management Studies & Research, Thane, for initiating us into this field of research and for providing me with the necessary guidance, great encouragement throughout the preparation of this paper. We take this opportunity to express my gratitude to the Staff members, Non-Teaching Staff members and Research Scholars of the Department of Computer Engineering, K. C College of Engineering & Management Studies & Research, Thane, for their timely help and encouragement.

REFERENCES

[1]. P.Geetha, Vasumathi Narayanan. An Effective Video Search Re-Ranking for Content Based Video Retrieval. IEEE 2011; 55-60.

[2]. Takahiro Yoshida, Kazuki tada, Seiichiro Hagai. A Keyword Accessible Lecture Video Player and Its Evaluation. IEEE 2003; 610-614.

[3]. Wen-Hsuan Chang, Jie-Chi Yang, Yu-Chieh Wu. A Keyword-based Video Summarization Learning Platform with Multimodal Surrogates

[4]. Chang, W.-H., Yang, J.-C., & Wu, Y.-C. (2011). A Keyword-based Video Summarization Learning Platform with Multimodal Surrogates. 2011 IEEE 11th International Conference on Advanced Learning Technologies. doi:10.1109/icalt.2011.19

[5]. TakahiroYanhida, Kazuki Tada, Takayuki Hamamoto, Seiichiro Hangai, "A Keyword Accessible AV-Stream Player", Prw. of ISAS SCDOOZ IVRCIA, Val. 16, pp.100-104, Orlando, USA July. 2002.

[6]. Jason A. Brotherton, Janak R. Bhalcdia, GregoIy D. Abowd, "Automated Capture, Integration, and Visualiratlon OfMultiple Media Slreamr", Pm. OfIEEEMultimedia, July 1998.

AUTHOR BIOGRAPHIES

Simmar Kalsi

Born in Thane, Maharashtra, India on 17/12/1998.

The author is currently pursuing Bachelors of Engineering in the stream of Computer Science from Excelssior Education Society's KC College of Engineering and Management Studies and Research and will earn his UG degree in 2020.

Harsh kanzariya

Born in Wadhvan, Gujarat, India on 21/05/1999.

The author is currently pursuing Bachelors of Engineering in the stream of Computer Science from Excelssior Education Society's KC College of Engineering and Management Studies and Research and will earn his UG degree in 2020.

Mandar Gankapurkar

Born in Kalyan, Maharashtra, India. On 24/02/1982

Assistant Professor

M.E. Computer Engineering

Experience: 10 YEARS

Area of Interest: Data Mining, Software Engineering, Cloud Computing