A Web Service for Video Summarization

Chrysa Collyda
ckol@iti.gr
CERTH-ITI
Thessaloniki, Greece

Konstantinos Apostolidis
kapost@iti.gr
CERTH-ITI
Thessaloniki, Greece

Evlampios Apostolidis
apostolid@iti.gr
CERTH-ITI
Thessaloniki, Greece &
Queen Mary Univ. of London, UK

Eleni Adamantidou
adamelen@iti.gr
CERTH-ITI
Thessaloniki, Greece

Alexandros I. Metsai
alexmetsai@iti.gr
CERTH-ITI
Thessaloniki, Greece

Vasileios Mezaris
bmezaris@iti.gr
CERTH-ITI
Thessaloniki, Greece

ABSTRACT

This paper presents a Web service that supports the automatic generation of video summaries for user-submitted videos. The developed Web application decomposes the video into segments, evaluates the fitness of each segment to be included in the video summary and selects appropriate segments until a pre-defined time budget is filled. The integrated deep-learning-based video analysis and summarization technologies exhibit state-of-the-art performance and, by exploiting the processing capabilities of modern GPUs, offer faster than real-time processing. Configurations for generating video summaries that fulfill the specifications for posting on the most common video sharing platforms and social networks are available in the user interface of this application, enabling the one-click generation of distribution-channel-specific summaries.

CCS CONCEPTS

• **Information systems → Summarization**; **Multimedia information systems**; • **Human-centered computing →** *User interface design*; • **Computing methodologies → Machine learning**.

KEYWORDS

Video summarization, Deep learning, Generative adversarial networks, Web service, Social networks, Video sharing platforms

ACM Reference Format:
Chrysa Collyda, Konstantinos Apostolidis, Evlampios Apostolidis, Eleni Adamantidou, Alexandros I. Metsai, and Vasileios Mezaris. 2020. A Web Service for Video Summarization. In *ACM International Conference on Interactive Media Experiences (IMX '20), June 17–19, 2020, Cornella, Barcelona, Spain.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3391614.3399391

1 INTRODUCTION

Journalists, broadcasters, as well as simple users regularly create video content to be published in established (e.g., TV) and emerging (e.g., Twitter, Facebook, YouTube, Instagram) channels. Whilst content produced for TV (e.g., an episode of a TV show) or for an amateur's own use (e.g., vacation footage) can be quite lengthy, this is not optimal for content distribution on social media. We live in an era where social media users' demands dictate the production of short and attractive audio-visual content that attracts the users' attention and can be ingested quickly; studies show that audience engagement drops significantly during the playback of long videos [37]. Therefore, for sharing on social platforms, video creators often need a trimmed-down version of their original full-length video. Also, different social platforms impose different restrictions on the duration and format of the video that they accept, e.g., on TikTok up to 60-second videos can be shared, whereas on Instagram stories the upper limit is 15 seconds. This makes the generation of tailored versions of video content for publication in multiple platforms a tedious task.

In this paper, we introduce a Web service that harnesses the power of artificial intelligence to automatically generate video summaries. It takes as input a video and produces a video summary that encapsulates the flow of the story and the essential parts of the full-length video, adapting the length and format of the produced summary for publication on social media platforms, thus easing the creation of engaging video stories for on-line audiences.

2 RELATED WORK

Several approaches have been proposed for addressing the task of video summarization over the last couple of decades. For a long time, the relevant research area was dominated by methods that select the key segments of the video based on the extraction and processing of low-level visual features from the video frames. These methods include: algorithms that assess the visual similarity over sequences of frames (e.g., [8, 39]); clustering-based techniques that group frames according to their visual similarity (e.g., [5, 14]); and approaches for visual attention modeling that imitate the human attention mechanism in order to spot the most important parts of the video for generating the summary (e.g., [7, 21]). Early supervised machine learning methods aimed to capture the underlying frame selection criterion from summaries created by humans to produce video summaries that meet human expectations (e.g., [15]); exploit auxiliary information, such as the video title or metadata, to extract the semantically-related parts of the video (e.g., [27]); and directly optimize multiple objectives for video summarization, such as representativeness, importance, and actionness (e.g., [9]). More recently, a number of deep learning video summarization approaches were introduced, with the majority of them being trained in a supervised

IMX '20, June 17–19, 2020, Cornella, Barcelona, Spain
© 2020 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7976-2/20/06.
https://doi.org/10.1145/3391614.3399391

manner, i.e., using ground-truth summaries. The learning efficiency of Convolutional Neural Networks (CNN) was exploited to extract the semantics of the video content and use this information for supervised video summarization [12, 26]. Other supervised techniques utilize advanced variations of Recurrent Neural Networks and perform video summarization by capturing the temporal dependency over the frame sequence [13, 20]. The latter together with some very recently proposed unsupervised learning methods that rely on Generative Adversarial Networks (GAN) for assessing the representativeness of the video summary, shape the current state of the art in general-purpose video summarization [1, 2, 38, 40]. For an extensive analysis of related video summarization works, the interested reader is referred to [1].

Regarding free web-based summarization tools, there are quite a few text summarization technologies [23, 24, 28, 32, 33] but none for video summarization, to the best of our knowledge. Of course, there exists a plethora of scientific papers; some of them provide source code that can be used for video summarization (e.g., [1, 2, 11, 25, 30, 41]). However, this requires that a user is well-informed on machine learning and proficient in computer programming, in order to use such code. Motivated by the lack of web-based video summarization tools, we built a freely accessible Web application that enables users to submit locally-stored or on-line available videos and automatically generate shorter versions of the submitted full-length video.

3 PROPOSED VIDEO SUMMARIZATION SERVICE AND USER INTERFACE IMPLEMENTATION

We designed a framework which consists of: a) a REST service that hosts the developed technologies for video summarization (backend) and b) an interactive user interface (UI) that allows the user to exploit the functionality of our Web service. In particular through the UI of this tool, the user is able to: a) submit a video for analysis (either available on-line or locally stored in the user's device), b) select the specifications of the generated summary from a list of predefined configurations that were properly adjusted for the most common social media and video sharing platforms, and c) get the created summary in a way that enables both immediate on-line inspection through the UI of our tool and the downloading of the video file in the user's device.

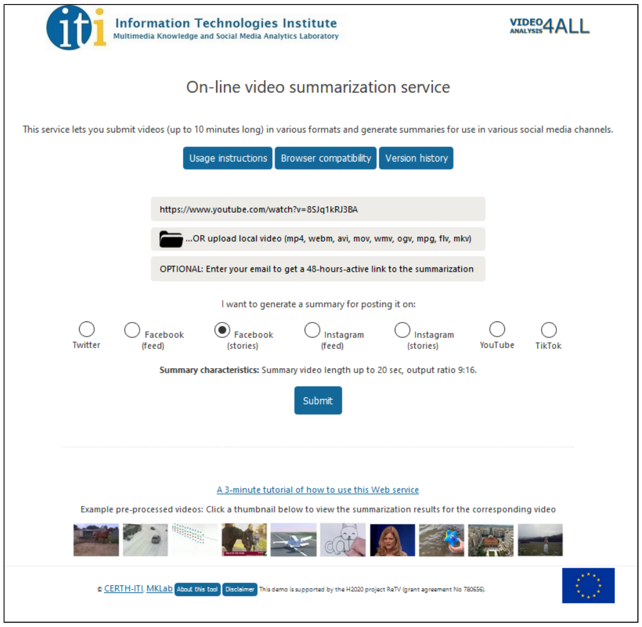
3.1 Description of backend

We employ a variation of the summarization method of [1] (which in turn is an extension of [2]). Combining the effectiveness of attention mechanisms in spotting the most important parts of the video with the learning efficiency of GANs for unsupervised training, this method ([1]) achieves state-of-the-art results. The algorithm works by calculating an importance score for each frame of the original full-length video. Given a video segmentation, fragment-level scores are calculated by averaging the scores of each fragment's frames. The summary is created by selecting the fragments that maximize its total importance score, under the constraint that the summary length does not exceed 15% of video duration, by solving the Knapsack problem [29].

We introduced two necessary modifications to [1]. First, instead of using the Knapsack algorithm for the selection of key segments, we are performing this selection through a shot-ranking method. In this way, we are able to combine the shot ranking according to the computed importance scores of [1] with different rankings that can be based on additional domain-based rules, a future perspective which we are interested in further investigating. Examples of such domain-based rules include the exclusion of video segments with, e.g., "talking heads" shots for video summaries generated by journalists, or the exclusion of shots that exhibit motion blur due to intense camera shaking. Second, instead of imposing a fixed limit regarding the duration of the generated video summary, we enable the selection of the target duration - a particularly important feature, since a different video summary of particular length has to be generated for each individual distribution channel where the video will be shared through.

The applied process for video analysis and summarization starts by segmenting the video into shots. A shot is an elementary structural unit of the video, that is composed of a set of consecutive frames captured by a single camera without interruption [3]. For the video shot segmentation we trained a 3D CNN with a similar architecture to [16] using the BBC Planet Earth dataset [4], which contains ground truth annotations for training a shot segmentation method. Each detected shot is ranked according to a value denoting its suitability to be part of the video summary, calculated as the mean of the estimated importance scores of the shot's frames (using the method of [1]), thus producing a ranked list of shots. The shot with the highest rank is selected as a candidate to be included in the summary and is removed from the ranked list. Then, two empirically-set thresholds, *min_segment_duration* and *max_segment_duration* are utilized, with the intent to impose bounds on the duration of the selected part from each shot, i.e., to avoid the inclusion of very short or very long segments in the summary. Specifically, if the selected shot's duration is greater than *min_segment_duration* and lower than *max_segment_duration* then the whole shot is included in the summary. If this duration is greater than *max_segment_duration* then we select a part of the shot (of *max_segment_duration* seconds) for which its frames exhibit the maximum sum of importance scores. If this duration is lower than *min_segment_duration* then the shot is discarded. The procedure is repeated until the video summary has a duration that is very close to the target duration, and results in an array containing the start and end time of each selected segment. The array is sorted based on the time of appearance of each segment in the original video and is subsequently fed to a separate module that is responsible for decoding the original video, finding the respective selected segments, transforming them to fit the target aspect ratio and rendering them to a final summary video file.

The above-described processing pipeline is deployed as a REST service that: a) retrieves a video file, b) analyzes the video using the method of [1] to estimate frame-level importance scores, c) performs temporal segmentation of the video to shots, d) ranks the shots and selects a part from each of the top-ranked shots until the summary's specified time-budget is filled based on the determined thresholds and the selection process described above, e) transforms the video frames to the target aspect ratio, and f) renders the video summary. The REST service works through a 3-step process. The



On-line video summarization service

This service lets you submit videos (up to 10 minutes long) in various formats and generate summaries for use in various social media channels.

https://www.youtube.com/watch?v=8SJq1kRJ3BA

...OR upload local video (mp4, webm, avi, mov, wmv, ogv, mpg, flv, mkv)

OPTIONAL: Enter your email to get a 48-hours-active link to the summarization

I want to generate a summary for posting it on:

○ Twitter ○ Facebook (feed) ● Facebook (stories) ○ Instagram (feed) ○ Instagram (stories) ○ YouTube ○ TikTok

Summary characteristics: Summary video length up to 20 sec, output ratio 9:16.

A 3-minute tutorial of how to use this Web service

Example pre-processed videos: Click a thumbnail below to view the summarization results for the corresponding video

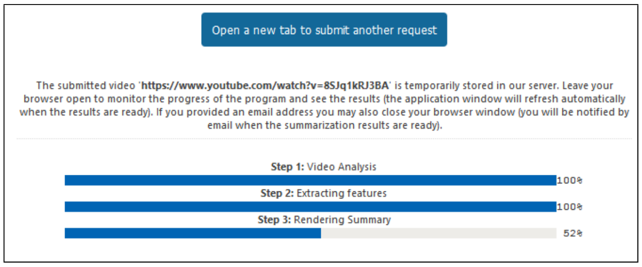
© CERTH-ITI, MKLab This demo is supported by the H2020 project ReTV (grant agreement No 780656).

Figure 1: The landing page of the frontend.

first step relates to an HTTP POST call that enables the submission of a video for analysis and the initiation of a relevant session in the REST service. The second step is associated to an HTTP GET call that queries the status of the initialized session and the progress of the analysis. Finally, the third step is performed by another HTTP GET call that enables the retrieval of the results of a successfully completed session.

3.2 Description of user interface

We have designed an interactive user interface (see Fig. 1) which allows the user to submit videos and generate summaries that are tailored for a user-specified social media channel. Video submission is performed on a one-by-one basis (i.e., no video collection analysis is supported) and, for demonstration purposes, the submitted videos are allowed to be up to 10 minutes long and 200MB in file size. In particular, to submit a video for summarization the user can either provide its URL or upload a local copy of it from his/her device. The supported on-line video sources include YouTube, Facebook, Twitter, Instagram, Vimeo, DailyMotion, LiveLeak and Dropbox. The service can handle videos in mp4, webm, avi, mov, wmv, ogv, mpg, flv, and mkv formats. After submitting a video, the user can monitor the progress of the summarization, and is also able to submit additional requests while the previous ones are being analyzed as shown in Fig. 2. The submitted video and the summarization results are cached in a server for 24 hours, and after this time period, the local copy or the video URL, the summarization results and the user's e-mail address (if provided) are automatically deleted from the server. When the analysis is completed, and after an automatic refresh of the user interface, the generated summary is presented to the user through the user interface presented in Fig. 3. Optionally, if the user provided an e-mail address she/he may close the Web browser and be notified by e-mail when the summary result is ready. The results page is an interactive webpage with two video players (for viewing the original and summarized videos) implemented using the HTML5 video tag. The players support all standard functionalities such as play/pause the video and toggle the video in full screen mode. Furthermore, the user is able to download the produced summary. The developed Web service for video summarization is fully compatible with Mozilla Firefox (>41.0), Chrome (>45.0), Opera (>32.0), Microsoft Edge (>77.0.200.1), and IE (>11.0).



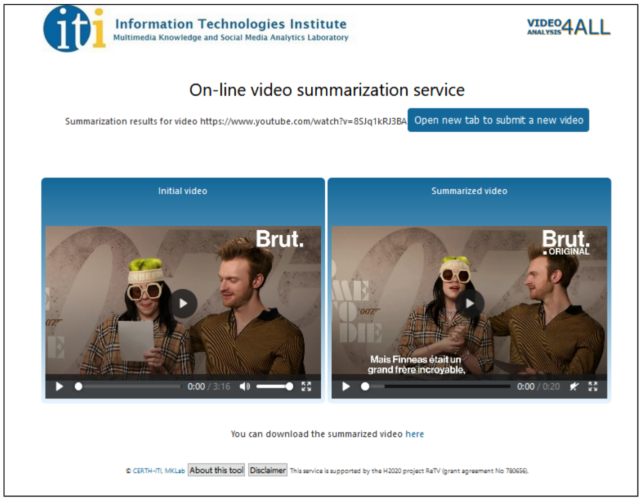
The submitted video '**https://www.youtube.com/watch?v=8SJq1kRJ3BA**' is temporarily stored in our server. Leave your browser open to monitor the progress of the program and see the results (the application window will refresh automatically when the results are ready). If you provided an email address you may also close your browser window (you will be notified by email when the summarization results are ready).

Step 1: Video Analysis 100%

Step 2: Extracting features 100%

Step 3: Rendering Summary 52%

Figure 2: The summarization progress bars of the frontend.


Information Technologies Institute
Multimedia Knowledge and Social Media Analytics Laboratory
VIDEO ANALYSIS 4ALL
Open new tab to submit a new video
Initial video
Summarized video
Brut.
Brut. ORIGINAL
Mais Finneas était un grand frère incroyable.
0:00 / 3:16
0:00 / 0:20
About this tool
Disclaimer

On-line video summarization service

Summarization results for video https://www.youtube.com/watch?v=8SJq1kRJ3BA

You can download the summarized video here

© CERTH-ITI, MKLab This service is supported by the H2020 project ReTV (grant agreement No 780656).

Figure 3: The results page of the frontend.

Aiming to help the user and relieve him/her from deciding the appropriate target duration and aspect ratio of the produced video summary, we have created a list of configurations for the most common social media and video sharing platforms. The user can select one using the respective option buttons found in the service's landing page (see Fig. 1). This action configures the service to produce a video summary that fully meets the prerequisites of the target platform, based on information gathered from [6, 10, 18, 19, 22, 34–36]. The specifications for five widely used social media and video sharing platforms are shown in Table 1, while an example of applying two different presets on the same input video is shown in Fig. 4. These parameters aim to maximize user engagement and

[1] This depends on the user account type; and additional limitations based on file size apply.

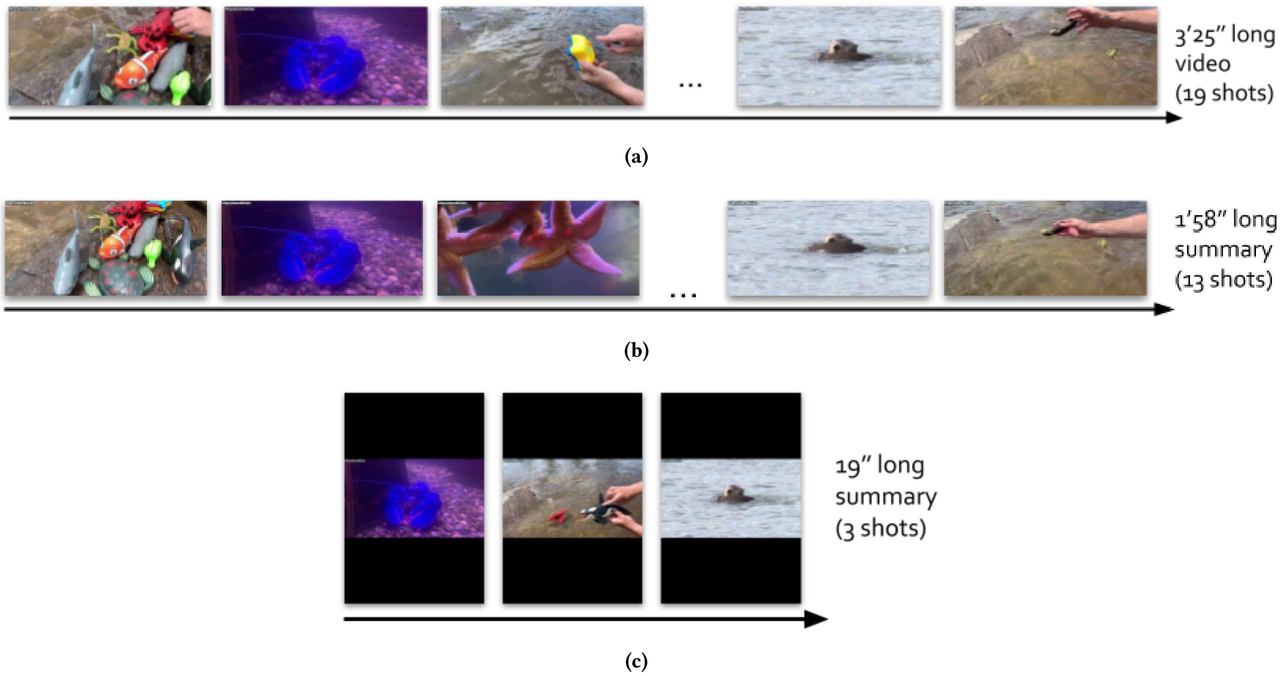

3'25" long video (19 shots)
(a)
1'58" long summary (13 shots)
(b)
19" long summary (3 shots)
(c)

Figure 4: Key-frames of shots for the: (a) original video with 19 shots and 205 seconds total length, (b) video summary for Facebook feed (16:9 aspect ratio) with 13 shots and 118 seconds total length, (c) video summary for Facebook story (9:16 aspect ratio) with 3 shots and 19 seconds total length.

Video sharing platform	Optimal summary length (seconds)	Video length hard limit (seconds)	Aspect Ratio
Twitter	30	140	2:1
Facebook (feed)	120	120	16:9
Facebook (stories)	20	20	9:16
Instagram (feed)	30	30	4:5
Instagram (stories)	15	15	9:16
YouTube	120	900 / unlimited[1]	original
TikTok	15	60	9:16

Table 1: Configuration settings for each social media platform. The developed service generates summaries that conform with the optimal summary length and aspect ratio listed above.

experience. Nevertheless, it is important to stress that these are best practices and they are subject to change based on various trends.

3.3 Description of frontend-backend communication

The frontend and backend components of our Web service are deployed as independent modules. Once a video is submitted to the frontend, a call to the backend service is initiated, which includes the specifications of the video summary to be generated. In turn, the backend instantiates a processing session, returns a unique session ID for reference to the frontend and inserts the session in the processing queue. The frontend periodically queries about the status of the backend session process. The response of the backend contains information about the position of the queried session in the queue or the progress, so that the user interface can provide the corresponding visual feedback to the user. Once the backend session is completed its status is set accordingly. The frontend can then retrieve the video summary as well as the original full length using another call, display the two videos side-by-side in the results interface and notify the user via email that the processing has been completed (if the user had opted to use this feature by submitting his/her email along with the original video).

4 DEPLOYMENT AND EXPERIMENTS

The original method of [1] that our work is based on, was evaluated on the SumMe [17] and TVSum [31] video summarization standard benchmark datasets. The findings show that the utilized method performs consistently well in both datasets, and is the most competitive one among the literature approaches. Details about that evaluation and an extensive analysis of the results can be found in [1].

The backend service of the demo is deployed on a PC with an Intel i7-4770K at 3.50 GHz, 32GB of RAM and a NVIDIA GeForce GTX 1660 graphics card. By exploiting the multi-thread and multi-core processing capabilities of the available CPU and GPU, the analysis is faster than real-time video processing, taking approximately 45% of the original video's playback time; though of course delays may be noticed if multiple analysis requests are submitted to the service, since these requests are processed on a one-by-one basis. Our Web service for video summarization can be accessed and tested at multimedia2.iti.gr/videosummarization/service/start.html.

As part of this work-in-progress, we ran experiments on publicly-available videos from social media platforms (e.g., to summarize a YouTube video for use in a Facebook story), beyond the experiments on benchmark datasets reported in [1]. These experiments verified that the Web service runs smoothly for videos of various formats, and visual inspection of the produced summaries showed that they are consistent with the summaries that a human would manually generate. Two such example summaries are shown in Fig. 4.

5 CONCLUSIONS AND NEXT STEPS

This paper presented the developed Web service for automatic generation of video summaries. Details about the use and functionalities of the service were given with the help of indicative snapshots of the implemented user interfaces. The integrated method for video summarization and the employed technologies for building the service were presented, and information about the performance of the developed service was given.

As next steps, the developed Web service will be tested with real users (primarily broadcasters and content archives staff as well as the general public) in the context of the ReTV project[2] (an EU Horizon 2020 research and innovation action), and based on a detailed analysis of these testing results will be further refined, extended and improved. In terms of development, future steps include giving the option to the user to employ editor-based rules (such as those briefly discussed in Section 3.1) and the integration of additional functionalities, such as generating multiple summaries for an input video and automatically posting the selected ones on the target platforms.

ACKNOWLEDGMENTS

This work was supported by the EU's Horizon 2020 research and innovation programme under grant agreement H2020-780656 ReTV. We would like to thank Rasa Bočytė of the Netherlands Institute for Sound and Vision for providing valuable information on the optimal summary length for the various video sharing platforms that are supported by our Web service.

REFERENCES

[1] Evlampios Apostolidis, Eleni Adamantidou, Alexandros I Metsai, Vasileios Mezaris, and Ioannis Patras. 2020. Unsupervised Video Summarization via Attention-Driven Adversarial Learning. In *International Conference on Multimedia Modeling*. Springer, 492–504.

[2] Evlampios Apostolidis, Alexandros I Metsai, Eleni Adamantidou, Vasileios Mezaris, and Ioannis Patras. 2019. A stepwise, label-based approach for improving the adversarial training in unsupervised video summarization. In *Proceedings of the 1st International Workshop on AI for Smart TV Content Production, Access and Delivery*. 17–25.

[3] Evlampios Apostolidis and Vasileios Mezaris. 2014. Fast shot segmentation combining global and local visual descriptors. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 6583–6587.

[4] Lorenzo Baraldi, Costantino Grana, and Rita Cucchiara. 2015. A deep siamese network for scene detection in broadcast videos. In *Proceedings of the 23rd ACM international conference on Multimedia*. 1199–1202.

[5] Wen-Sheng Chu, Yale Song, and Alejandro Jaimes. 2015. Video co-summarization: Video summarization by visual co-occurrence. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3584–3592.

[6] Giorgos Dimopoulos, Pere Barlet-Ros, and Josep Sanjuas-Cuxart. 2013. Analysis of YouTube user experience from passive measurements. In *Proceedings of the 9th International Conference on Network and Service Management (CNSM 2013)*. IEEE, 260–267.

[7] Naveed Ejaz, Irfan Mehmood, and Sung Wook Baik. 2014. Feature aggregation based visual attention model for video summarization. *Computers & Electrical Engineering* 40, 3 (2014), 993 – 1005. https://doi.org/10.1016/j.compeleceng.2013.10.005 Special Issue on Image and Video Processing.

[8] Naveed Ejaz, Tayyab Bin Tariq, and Sung Wook Baik. 2012. Adaptive key frame extraction for video summarization using an aggregation mechanism. *Journal of Visual Communication and Image Representation* 23, 7 (2012), 1031 – 1040.

[9] Mohamed Elfeki and Ali Borji. 2019. Video summarization via actionness ranking. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 754–763.

[10] Facebook video requirements (accessed: 2020-03-20). https://www.facebook.com/business/m/one-sheeters/video-requirements

[11] Jiri Fajtl, Hajar Sadeghi Sokeh, Vasileios Argyriou, Dorothy Monekosso, and Paolo Remagnino. 2018. Summarizing videos with attention. In *Asian Conference on Computer Vision*. Springer, 39–54.

[12] Jiri Fajtl, Hajar Sadeghi Sokeh, Vasileios Argyriou, Dorothy Monekosso, and Paolo Remagnino. 2019. Summarizing Videos with Attention. In *Computer Vision – ACCV 2018 Workshops*, Gustavo Carneiro and Shaodi You (Eds.). Springer International Publishing, Cham, 39–54.

[13] Tsu-Jui Fu, Shao-Heng Tai, and Hwann-Tzong Chen. 2019. Attentive and Adversarial Learning for Video Summarization. In *IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, HI, USA, January 7-11, 2019*. 1579–1587. https://doi.org/10.1109/WACV.2019.00173

[14] Marco Furini, Filippo Geraci, Manuela Montangero, and Marco Pellegrini. 2010. STIMO: STIll and MOving Video Storyboard for the Web Scenario. *Multimedia Tools Appl.* 46, 1 (Jan. 2010), 47–69.

[15] Boqing Gong, Wei-Lun Chao, Kristen Grauman, and Fei Sha. 2014. Diverse Sequential Subset Selection for Supervised Video Summarization. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2* (Montreal, Canada) *(NIPS'14)*. MIT Press, Cambridge, MA, USA, 2069–2077. http://dl.acm.org/citation.cfm?id=2969033.2969058

[16] Michael Gygli. 2017. Ridiculously fast shot boundary detection with fully convolutional neural networks. *arXiv preprint arXiv:1705.08214* (2017).

[17] Michael Gygli, Helmut Grabner, Hayko Riemenschneider, and Luc Van Gool. 2014. Creating summaries from user videos. In *European conference on computer vision*. Springer, 505–520.

[18] How Long Should Your Videos Be? Ideal Lengths for Facebook, Instagram, Twitter, and YouTube (accessed: 2020-03-20). https://blog.hubspot.com/marketing/how-long-should-videos-be-on-instagram-twitter-facebook-youtube

[19] How many seconds of video can I record on Instagram? (accessed: 2020-03-20). https://www.facebook.com/help/instagram/270963803047681

[20] Zhong Ji, Kailin Xiong, Yanwei Pang, and Xuelong Li. 2019. Video summarization with attention-based encoder-decoder networks. *IEEE Transactions on Circuits and Systems for Video Technology* (2019).

[21] Jie-Ling Lai and Yang Yi. 2012. Key frame extraction based on visual attention model. *Journal of Visual Communication and Image Representation* 23, 1 (2012), 114 – 125. https://doi.org/10.1016/j.jvcir.2011.08.005

[22] Christian Moldovan, Florian Wamser, and Tobias Hoßfeld. 2019. User Behavior and Engagement of a Mobile Video Streaming User from Crowdsourced Measurements. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 1–3.

[23] Online summarize tool (free summarizing) (accessed: 2020-03-20). https://www.tools4noobs.com/summarize/

[24] Online Text Summary Generator (accessed: 2020-03-20). http://autosummarizer.com/

[25] Mayu Otani, Yuta Nakashima, Esa Rahtu, and Janne Heikkila. 2019. Rethinking the evaluation of video summaries. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7596–7604.

[26] Rameswar Panda, Abir Das, Ziyan Wu, Jan Ernst, and Amit K Roy-Chowdhury. 2017. Weakly supervised summarization of web videos. In *Proceedings of the IEEE International Conference on Computer Vision*. 3657–3666.

[27] Danila Potapov, Matthijs Douze, Zaid Harchaoui, and Cordelia Schmid. 2014. Category-Specific Video Summarization. In *Computer Vision – ECCV 2014*, David

[2] https://retv-project.eu/

Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars (Eds.). Springer International Publishing, Cham, 540–555.

[28] Resoomer | Summarizer to make an automatic text summary online (accessed: 2020-03-20). https://resoomer.com/en/

[29] Sartaj Sahni. 1975. Approximate algorithms for the 0/1 knapsack problem. *Journal of the ACM (JACM)* 22, 1 (1975), 115–124.

[30] Yair Shemer, Daniel Rotman, and Nahum Shimkin. 2019. ILS-SUMM: Iterated Local Search for Unsupervised Video Summarization. *arXiv preprint arXiv:1912.03650* (2019).

[31] Yale Song, Jordi Vallmitjana, Amanda Stent, and Alejandro Jaimes. 2015. Tvsum: Summarizing web videos using titles. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 5179–5187.

[32] Text Compactor: Free Online Automatic Text Summarization Tool (accessed: 2020-03-20). https://www.textcompactor.com/

[33] Text Summarizer - Text Summarization (accessed: 2020-03-20). http://textsummarization.net/text-summarizer

[34] The Ideal Length For Every Online Content (accessed: 2020-03-20). https://seopressor.com/blog/the-ideal-length-for-every-online-content/

[35] The Ultimate Guide to TikTok Videos (accessed: 2020-03-20). https://clipchamp.com/en/blog/2019/ultimate-guide-to-tiktok/

[36] Twitter media upload best practices (accessed: 2020-03-20). https://developer.twitter.com/en/docs/media/upload-media/uploading-media/media-best-practices

[37] Video Length: 4 Tips That Will Help You Boost Engagement (accessed: 2020-03-20). https://meetmaestro.com/insights/how

[38] Li Yuan, Francis Eng Hock Tay, Ping Li, Li Zhou, and Jiashi Feng. 2019. Cycle-SUM: Cycle-Consistent Adversarial LSTM Networks for Unsupervised Video Summarization. In *2019 AAAI Conference on Artificial Intelligence (AAAI)*.

[39] HongJiang Zhang, Jianhua Wu, Di Zhong, and Stephen W. Smoliar. 1997. An integrated system for content-based video retrieval and browsing. *Pattern Recognition* 30 (1997), 643–658.

[40] Kaiyang Zhou and Yu Qiao. 2018. Deep Reinforcement Learning for Unsupervised Video Summarization with Diversity-Representativeness Reward. In *2018 AAAI Conference on Artificial Intelligence (AAAI)*.

[41] Kaiyang Zhou, Yu Qiao, and Tao Xiang. 2018. Deep reinforcement learning for unsupervised video summarization with diversity-representativeness reward. In *Thirty-Second AAAI Conference on Artificial Intelligence*.