2023 26th International Conference on Computer and Information Technology (ICCIT)
13-15 December, Cox's Bazar, Bangladesh

ROUGE Score Analysis and Performance Evaluation Between Google T5 and SpaCy for YouTube News Video Summarization

Nurshat Fateh Ali, Jamal Uddin Tanvin, MD Rifat Islam,
Jubair Ahmed, M. Akhtaruzzaman
Department of Computer Science and Engineering, Military Institute of Science and Technology,
Mirpur Cantonment, Dhaka-1216, Bangladesh
E-mail: nurshatfateh@gmail.com, jamaluddintanvin@outlook.com, rifat010bushra@gmail.com,
jubairahmed1678@gmail.com, akhter900@cse.mist.ac.bd

***Abstract*—The project aims to compare the performance between Google T5 and SpaCy for YouTube video textual summarization based on the ROUGE score. The project utilizes two distinct techniques: abstractive summarization using Google T5 and extractive summarization using SpaCy. The CNN/DailyMail dataset serves as the primary data source for training the model and the YouTube video description sections are used for evaluating the results. The project evaluates the effectiveness of each technique separately. The strengths and weaknesses of abstractive and extractive summarization approaches are discussed, providing valuable insights into their respective trade-offs and applicability for YouTube news video summarization. From the observations, it is found that Google T5 outperforms SpaCy every single time. The overall ROUGE-1 score for Google T5 is found to be 0.339 while the SpaCy scored 0.232. The findings demonstrate the potential of AI-based techniques for automating the summarization process for YouTube news videos.**

***Index Terms*—Google T5, SpaCy, ROUGE, abstractive summarization, extractive summarization, YouTube videos.**

I. Introduction

YouTube has emerged as a vast repository of news content [1], including videos and talk shows on a wide range of news categories in today's digital era. However, the sheer volume of information can be overwhelming for viewers seeking specific knowledge within a limited time frame. To address this challenge, video summarization methods are used to make information easily graspable.

The primary objective of this project is to evaluate the performance of two approaches of text summarization using ROUGE [2] scores. Two distinct techniques are employed: abstractive summarization using Google T5 and extractive summarization using SpaCy. Abstractive summarization involves generating new phrases and sentences that capture the essence of the original video, going beyond the mere extraction of sentences. This technique leverages the power of language models and contextual understanding to produce coherent and concise summaries. On the other hand, extractive summarization focuses on selecting and organizing important sentences from the video based on their relevance and importance.

To train the T5 model, the CNN/DailyMail dataset consisting of news articles and corresponding summaries is used. As a highly curated dataset, it provides a valuable resource for training the summarization model. Separate evaluations are conducted for both abstractive Google T5 summarization and extractive SpaCy summarization, analyzing the strengths and weaknesses of each technique, and providing valuable insights into their applicability and performance for YouTube news video summarization.

The outcomes of this project hold significant potential for enhancing the news viewers' experience on YouTube. By choosing the right approach for the summarization process, viewers can save time and effort to gain a comprehensive understanding of the video content. Additionally, this project contributes to the advancement of automatic summarization [3] in the news domain, providing insights and guidance for future developments in the field. In this article, a comprehensive analysis of the process including the methodologies, dataset, evaluation metric, and the result is conducted. A discussion on the implications of the findings and highlighting areas for future research and improvement is also presented. This work mainly aims to determine more efficient and accurate summarization approaches for news and journal resources, ultimately enhancing knowledge accessibility and learning outcomes in the digital age. The highlights of this article are:

- Google T5 summarization shows better unigram overlapping with the summaries generated by humans.
- Google T5 presents better results (10% higher) than SpaCy in terms of the ROUGE-1 score.
- The potential of Google T5 in summarizing YouTube news videos is very promising.

II. Literature Review

YouTube video summarization has gained prominence as a means to enhance the accessibility and utility of educational content on the platform. Bhagat *et al.* [4] introduced an approach with a user-friendly interface for video text summarization. They employed Flask as the backend framework to handle API calls and process user requests. The backend server retrieved video subtitles from YouTube, and Natural Language Processing (NLP) techniques were employed to process the received transcript. They implemented multiple summarization methods such as the Text Rank algorithm, Luhn algorithm, Latent Semantic analysis, and Frequency-based algorithm. This approach offered users the ability to download, translate, share, and listen to audio summaries. However, it heavily relied on YouTube videos having proper closed captions, and the translated text did not support all file formats. Additionally, the summarization process did not capture the essence of the video as desired by all users.

A modern approach to video summaries with emphasis on YouTube content was proposed by Bhandare *et al.* [5]. They used the Pytube library for retrieving videos from YouTube, while the LSA algorithm was implemented to extract the necessary features for the summarization. For clip concatenation, they used the MoviePy library. Their proposed method demonstrated the effectiveness of producing accurate video summaries, which was supported by metric analysis.

An automatic sharing of summarized transcripts through WhatsApp and email was introduced by Kumari *et al.* [6] which leverages Python APIs for text transcription and NLP techniques. This approach effectively summarized video content but relied on videos with properly prepared closed captions. It did not support translations in all file formats, and summarization effectiveness varied with video content.

Emad *et al.* [7] focused on generating informative video descriptions and timestamps. Combining video frames, audio transcriptions, and emotions enhanced summary accuracy and coherence. Challenges included emotion detection accuracy and performance variations across different video types.

An LSTM-based approach that combined RGB and motion history image (MHI) features for video description generation was proposed by Zhang *et al.* [8]. The method demonstrated promising results but relied on the quality and diversity of training data and tended to favor more superficial sentence structures.

The majority of the related works emphasized extractive summarization using frequency and ranking-based algorithms. Moreover, they used LSTM or CNN models for abstractive summarization. There were no specialized pre-trained deep learning models used in any of the works. Furthermore, none of the reviewed works were done in a specific sector, such as news or journal video summarization. This article presents the utilization of an advanced Google T5 model that has been expertly fine-tuned and trained on a carefully curated dataset in a specific sector, offering a powerful and highly effective solution. The article also shows the comparison between the Google T5 model and SpaCy.

III. Study Design

A. Data collection

For data collection, the CNN/DailyMail dataset [9], a widely popular dataset for text summarization from the Hugging face is used. The dataset contains over 300k unique news articles with corresponding summaries. Though the original dataset was created for machine reading, comprehension, and abstractive question answering, the current version supports both extractive and abstractive summarization. Using this dataset, a wide range of topics and writing styles are provided for the foundation of the model to grasp the essence of different subjects. Although the dataset is not specifically tailored to video content, it is a valuable resource for training and evaluating the summarization models.

B. Collecting YouTube Transcript

One of the foremost steps is to collect the YouTube link of the video that requires summarization. This link serves as the source for retrieving the video transcript. Next, the unique ID is extracted, which is essential for retrieving the transcript using YouTube Transcript Api. The API provides access to closed captioning or automatic speech recognition (ASR) generated transcripts. The retrieved transcript is then converted into coherent paragraphs for further analysis.

C. Abstractive Summarization

Abstractive summarization methods generate new text that is not placed in the original text [10]. Abstractive methods build an internal semantic representation of the original content (often called a language model) and then use this representation to create a summary that is closer to what a human might express. Abstraction may transform the extracted content by paraphrasing sections of the source document to condense a text more strongly than extraction.

The methodology for abstractive summarization in this project involves several key steps to develop an efficient and accurate summarization system for YouTube videos. The procedure is illustrated in Figure 1. The process begins with selecting a suitable pre-trained model, such as Google T5 [11]. Google T5 is an encoder-decoder model pre-trained on a multitask mixture of unsupervised and supervised tasks for which each task is converted into a text-to-text format. It is basically a text-to-text transfer transformer model for abstractive summarization. T5 is trained on the unlabeled large text corpus called C4 (Colossal Clean Crawled Corpus) using deep learning. C4 is the web extract text of 800Gb cleaned data. The cleaning process involves deduplication, discarding incomplete sentences, and removing offensive or noisy content.

After selecting the Google T5 model, the model is fine-tuned and trained using the collected dataset, ensuring that the model adapts to the specific requirements of abstractive summarization. The dataset consists of diverse news articles and summaries, providing a range of topics and writing styles to train the model effectively.

To train the model, the dataset is split into training and validation sets. The training set is used to optimize the model's parameters through iterative training iterations. During training, the model's performance is monitored using the validation set to prevent overfitting and select the best-performing model.

Once the model is trained, the summary generation stage is started. The video text is preprocessed to generate abstractive summaries, ensuring it is formatted appropriately for input to the model. The preprocessed text is then fed into the trained

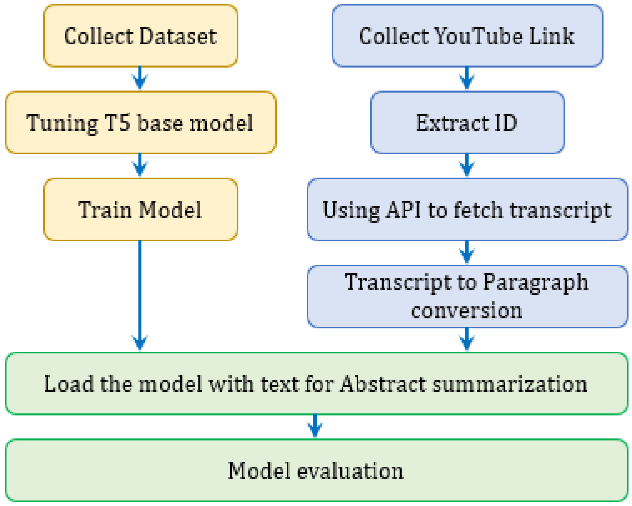

Figure 1: Abstractive summarization flow diagram.

model, which generates concise and coherent summaries that capture the essence of the original videos.

D. Extractive Summarization

Extractive summarization is an NLP technique that involves extracting sentences or phrases from a document to create a concise summary. SpaCy, an open-source library written in Python and Cython, is used for this task. It offers fast syntactic analysis, Named Entity Recognition, and access to word vectors. SpaCy is effective for large-scale information extraction tasks and offers tokenization, sentence boundary detection, Parts of Speech tagging, syntactic parsing, integrated word vectors, and alignment into the original string. SpaCy converts texts into tokens, calculates individual word frequencies, and ranks sentences based on importance. Top-ranked sentences are selected and merged to create a summarized textual paragraph. The procedure is illustrated in Figure 2.

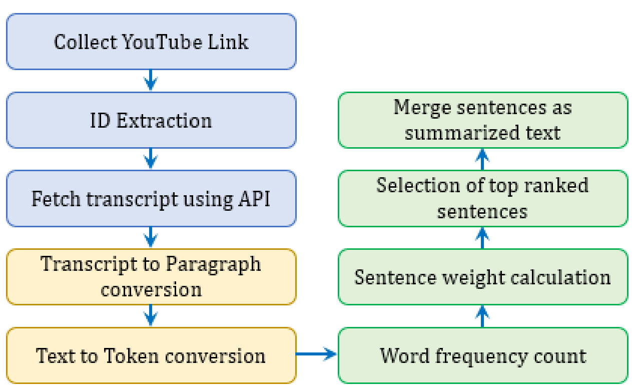

Figure 2: Extractive summarization flow diagram.

Tokenization is applied to the paragraphed transcript, breaking it down into individual words or tokens. This tokenization process enabled the analysis of word frequencies and weights in the transcript. By calculating the frequencies of each word, insights into the importance of specific words were gained within the video context.

The next step involved assigning weights to each sentence based on the cumulative frequencies of the words it contains. Sentences with higher weight are considered more important. These weights helped to determine the significance of each sentence in the transcript.

The top-ranked sentences are selected based on their calculated weights to create the extractive summary. Typically, a predetermined percentage of sentences, such as 10% of the total, is chosen. These selected sentences represent the most critical and relevant information extracted directly from the video transcript. Finally, the selected sentences are merged to form the final extractive summary.

IV. Test Case Analysis

To evaluate the result, YouTube video summarization from the YouTube video's description section is manually collected and then compared against both system-generated extractive and abstraction summarizations using the ROUGE score.

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) is a set of metrics used for evaluating the quality of machine-generated summaries by comparing them to reference (human-generated) summaries. ROUGE scores operate on the basis of comparing n-grams (contiguous sequences of n words) and other linguistic units between the generated summary and the reference summary. The most commonly used ROUGE metrics include ROUGE-N (precision, recall, and F1 score for n-gram overlaps), ROUGE-L (measuring longest common subsequence), and ROUGE-W (weighted F1 score for word overlaps).

For the evaluation of this project, the following variations of ROUGE scores are taken into consideration.

1) ROUGE-1 (ROUGE-N for n=1)
2) ROUGE-2 (ROUGE-N for n=2)
3) ROUGE-L (ROUGE-Longest)
4) ROUGE-Lsum (ROUGE-Longest for summary level evaluation)

Each of these ROUGE metrics focuses on different aspects of summarization quality, from matching individual words (ROUGE-1 and ROUGE-2) to assessing content overlap and coherence (ROUGE-L and ROUGE-Lsum).

ROUGE-N is an n-gram recall between a candidate summary and a set of reference summaries. ROUGE-N is computed as follows:

$$ROUGE-N = \frac{\sum\limits_{S \epsilon \{RefSum\}} \sum\limits_{g_n \epsilon S} C_m(g_n)}{\sum\limits_{S \epsilon \{RefSum\}} \sum\limits_{g_n \epsilon S} C(g_n)} \qquad (1)$$

Here, $RefSum$ indicates Reference Summaries, n is the length of the $n-gram$ (g_n), and $C_m(g_n)$ is the maximum number of $n-gram$ co-occurring in a candidate summary and a set of reference summaries. $C(g_n)$ is the number of n-grams in a sentence.

On the other hand, **ROUGE-L** is mainly the Longest Common Subsequence (LCS). A sequence $Z = [z_1, z_2, ..., z_n]$ is a subsequence of another sequence $X = [x_1, x_2, ..., x_m]$, if there exists a strict increasing sequence $[i_1, i_2, ..., i_k]$ of indices of X such that for all $j = 1, 2, ..., k$, we have $x_{ij} = z_j$ [12]. If two sequences X and Y are given, LCS of X and Y is a common subsequence with maximum length. To apply LCS in summarization evaluation, a summary sentence is viewed as a sequence of words. The general principle is that the longer the LCS of two summary sentences is, the more similar the two summaries are. ROUGE-L is simply using LCS-based F-measure to estimate the similarity between two summaries X of length m and Y of length n, assuming X is a reference summary sentence and Y is a candidate summary sentence, as follows:

$$R_{lcs} = \frac{LCS(X,Y)}{m} \tag{2}$$

$$P_{lcs} = \frac{LCS(X,Y)}{n} \tag{3}$$

$$F_{lcs} = \frac{(1+\beta^2)R_{lcs}P_{lcs}}{R_{lcs}+\beta^2 P_{lcs}} \tag{4}$$

Here $LCS(X,Y)$ is the length of the longest common subsequence of X and Y, and $\beta = P_{lcs}/R_{lcs}$. Equation 4 is considered as the ROUGH-L score.

ROUGE-Lsum is applied to summary-level. The union LCS matches between a reference summary sentence, r_i, and every candidate summary sentence, c_j. Given a reference summary of u sentences containing a total of m words and a candidate summary of v sentences containing a total of n words, the summary-level LCS-based F-measure can be computed as follows:

$$R_{lcs} = \frac{\sum_{i=1}^{u} LCS_{\cup}(r_i, C)}{m} \tag{5}$$

$$P_{lcs} = \frac{\sum_{i=1}^{u} LCS_{\cup}(r_i, C)}{n} \tag{6}$$

$$F_{lcs} = \frac{(1+\beta^2)R_{lcs}P_{lcs}}{R_{lcs}+\beta^2 P_{lcs}} \tag{7}$$

Here $LCS_{\cup}(r_i, C)$ is the LCS score of the union longest common subsequence between reference sentence r_i and candidate summary C and $\beta = P_{lcs}/R_{lcs}$. Equation 7 is considered as the ROUGH-Lsum score.

Three test cases are considered and corresponding Rouge metrics are presented in Tables I, II, and III.

A. Test Case-1

YouTube link:
https://www.youtube.com/watch?v=6VZfeexme4c
Video Title: North Korean leader Kim Jong Un enters Russia to visit President Putin - BBC News
Provided Human Summary: North Korea's leader Kim Jong Un has crossed the border into Russia for a meeting with President Vladimir Putin. They are likely to discuss an arms deal as Russia faces a Ukrainian counter offensive, a US official said. South Korea's defence ministry confirmed that Mr Kim's armoured train entered Russia early on Tuesday morning.
T5 Summary: Kim Jong-un's train entered Russia on his way to meet with President Putin. Mr Kim is believed to be meeting to finalize an arms deal with Russia. North Korea has become increasingly isolated, according to reports.
SpaCy Summary 10%: It's also thought that Mr Kim might ask in return for these weapons, might ask Mr Putin to hand over sensitive weapons technology that he could then use to make breakthroughs in his own nuclear weapons program, though this isn't something that officials here are especially concerned about. it's taking place in Vladivostok and, according to the Kremlin, at the meeting with Kim Jong-un, will be taking place at the conclusion of that event.

Table I: ROUGE scores for Test case-1

ROUGE Scores	Google T5	SpaCy
ROUGE-1	0.458	0.221
ROUGE-2	0.191	0.075
ROUGE-L	0.270	0.103
ROUGE-Lsum	0.271	0.103

B. Test Case-2

YouTube link:
https://www.youtube.com/watch?v=_pHglSTvgbI
Video Title: The Russians leaving their country for Finland - BBC News
Provided Human Summary: Some Russian people are anxious to get out of the country because there has been a persistent rumour that President Vladimir Putin's government might soon introduce martial law to deal with demonstrations against the invasion of Ukraine. With flights to Europe halted, the only way out of the country is by car - crossing this border - or by train.
T5 Summary: People in finland want to give up old neutrality and join the western alliance. there are rumors that president putin will soon introduce martial law. for russians coming to finland is an escape from the dangers of life there.
SpaCy Summary 10%: For russians coming to finland is an escape from the dangers of life there, but for people on this side of the border there's real fear that the tensions within russia could boil over and engulf finland itself.

Table II: ROUGE scores for Test case-2

ROUGE Scores	Google T5	SpaCy
ROUGE-1	0.327	0.265
ROUGE-2	0.083	0.021
ROUGE-L	0.286	0.122
ROUGE-Lsum	0.286	0.122

C. Test Case-3

YouTube link:
https://www.youtube.com/watch?v=TfKSOIhLpNc
Video Title: Soyuz rocket launches to take astronauts to International Space Station - BBC News
Provided Human Summary: A new crew is being taken to the International Space Station, on a joint Russian-US rocket which took off from Kazakhstan. The three members of Expedition 70 will relieve astronauts currently on the space station, and will remain there for six months to a year in the orbiting lab. The Soyuz MS-24 spacecraft launched from the Baikonur Cosmodrone in Kazakhstan.
T5 Summary: Roscoe Cosmos soyuz ms-24 rocket will launch three astronauts to the International Space Station. They will then meet the ISS, the International Space Station, and dock at Rendezvous. BBC News: Launch was successful around the world and across the UK and up in space.
SpaCy Summary 10%: Three astronauts and NASA astronaut and two Roscoe- Roscoe Cosmos astronauts, cosmonauts, to the International Space Station where they will be docking soon after taking off. then we will see the three astronauts entering the ISS joining the Space Station's Expedition 69 crew.

Table III: ROUGE scores for Test case-3

ROUGE Scores	Google T5	SpaCy
ROUGE-1	0.392	0.340
ROUGE-2	0.152	0.134
ROUGE-L	0.243	0.226
ROUGE-Lsum	0.243	0.226

D. Overall ROUGE Score

Individual ROUGE scores of the ten test cases are given in Table IV. By reviewing these ten human-provided summarizations and comparing them to the summary created by the project, the overall ROUGE score for both the Google T5 model and SpaCy is computed, as shown in Table V.

V. Result and Discussion

The study introduced two procedures for textual extractive and abstractive summarization of YouTube news videos. The study also illustrates the performance comparison between Google T5 and SpaCy summarization using ROUGE metrics. The procedure takes a YouTube link as input. Then it uses API to collect the transcript of the video.

To analyze the effectiveness and the comparison of performance, the system-produced summaries are compared with the

Table IV: Individual ROUGE score for Google T5 and SpaCy of ten samples

Youtube Link and Title	ROUGE Scores	T5	SpaCy
https://www.youtube.com/watch?v=6VZfeexme4c **Title:** North Korean leader Kim Jong Un enters Russia to visit President Putin - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.458 0.191 0.270 0.271	0.221 0.075 0.103 0.103
https://www.youtube.com/watch?v=bFZHLYLzJ-Q **Title:** US President Joe Biden's son Hunter indicted on federal gun charges - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.396 0.043 0.230 0.230	0.242 0.041 0.152 0.152
https://www.youtube.com/watch?v=3tKpbsHe1NQ **Title:** Russia pilot tried to shoot down RAF aircraft in 2022 - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.321 0.104 0.231 0.231	0.274 0.026 0.128 0.128
https://www.youtube.com/watch?v=XYDRuEQRTzE **Title:** Libya floods leave more than 5,000 people dead - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.268 0.063 0.144 0.144	0.150 0.027 0.102 0.102
https://www.youtube.com/watch?v=SRrgb8oxMxc **Title:** US and Vietnam elevate diplomatic ties as President Biden visits Hanoi	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.351 0.083 0.189 0.189	0.301 0.044 0.194 0.194
https://www.youtube.com/watch?v=Faj9EhkgBIY **Title:** South China Sea tensions: Confrontation between China and Philippines	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.368 0.141 0.299 0.299	0.220 0.075 0.165 0.165
https://www.youtube.com/watch?v=ULlGENlkMB8 **Title:** UK accuses China of interfering in its democracy - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.252 0.017 0.134 0.134	0.173 0.022 0.092 0.092
https://www.youtube.com/watch?v=_pHglSTvgbI **Title:** The Russians leaving their country for Finland - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.327 0.083 0.286 0.286	0.265 0.021 0.122 0.122
https://www.youtube.com/watch?v=TfKSOIhLpNc **Title:** Soyuz rocket launches to take astronauts to International Space Station - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.392 0.152 0.243 0.243	0.340 0.135 0.226 0.226
https://www.youtube.com/watch?v=2EimOYbvnnI **Title:** Lampedusa: 7,000 migrants arrive on Italian island in three days - BBC News	ROUGE-1 ROUGE-2 ROUGE-L ROUGE-Lsum	0.244 0.045 0.155 0.155	0.128 0.015 0.084 0.084

Table V: The overall ROUGE scores for Google T5 and SpaCy

ROUGE Scores	**Google T5**	**SpaCy**
ROUGE-1	0.339	0.232
ROUGE-2	0.092	0.047
ROUGE-L	0.219	0.137
ROUGE-Lsum	0.218	0.136

YouTube-provided summaries from the YouTube description section for both approaches. ROUGE-1, ROUGE-2 (ROUGE-N for n=2), ROUGE-L, and ROUGE-Lsum scores are calculated. ROUGE-1 refers to the overlap of unigrams between the system summary and reference summary. For both approaches the ROUGE-1 scores are above the acceptable mark. In every single case, it is seen that the T5 model produced results with higher ROUGE scores than the SpaCy method. The overall ROUGE-1 score for T5 is 0.339 while the score for SpaCy is 0.232. It shows that T5 model-generated summaries have better unigrams overlapping with human summaries. The same pattern followed for other ROUGE metrics too. T5 outperforms SpaCy in all tests. From the scores, it can be quickly concluded that the T5 model produced good summarizations as per standards but the SpaCy approach produced summaries that can be described as average at best.

VI. Conclusion and Future Scope

The project focused on implementing and comparing the performance of both abstractive and extractive summarization techniques for YouTube news video summarization. For the ROUGE score comparison of Google T5 and SpaCy, it is found that the abstraction summarization method using Google T5 (ROUGE-1: 0.339) outperforms SpaCy (ROUGE-1: 0.232) in every single metric. For all types of ROUGE evaluation, higher scores for Google T5 are observed certifying the superiority of the transformer-based models for summarization.

Future work of this project can be focussed on enhancing the effectiveness and applicability of the summarization system by developing a browser extension. It will allow users to generate on-demand summaries while browsing YouTube videos. This extension will provide users with quick access to summarized content, enhancing their browsing experience and facilitating efficient information consumption. Also, extending the system to support multiple languages will enhance its accessibility and usefulness globally.

Acknowledgment

The authors extend sincere thanks to the Ministry of Defence, the CSE department of the Military Institute of Science and Technology, particularly the Quantum Robotics and Automation Research Group (QRARG). Gratitude is also expressed to the R&D Section of DynREAM Robotics Ltd. for valuable insights in enhancing the article.

References

[1] J. Burgess and E. Hurcombe, “Digital Journalism as Symptom, Response, and Agent of Change in the Platformed Media Environment,” Digital Journalism, vol. 7, no. 3, pp. 359–367, Mar. 2019, doi: 10.1080/21670811.2018.1556313.

[2] C.-Y. Lin, “ROUGE: A Package for Automatic Evaluation of Summaries,” Meeting of the Association for Computational Linguistics, pp. 74–81, Jul. 2004.

[3] A. Nenkova, “Automatic Summarization,” Foundations and Trends® in Information Retrieval, vol. 5, no. 2, pp. 103–233, 2011, doi: 10.1561/1500000015.

[4] A. Bhagat, M. Singade, P. Anjankar, S. Surbhik, and N. Sambhe, “Text Summarization on YouTube Videos in Education Domain,” BioGecko, vol. 12, no. 03, pp. 859-866, 2023.

[5] Ms. K. M. Bhandare, A. A. Chigare, U. U. Patil, and S. B. Sangle, “YouTube Transcript Summarizer,” International Research Journal of Modernization in Engineering Technology and Science, vol. 04, no. 03, pp. 2495-2299, Mar. 2022, Available: www.irjmets.com.

[6] P. V. Kumari, M. C. Keshava, C. Narendra, P. Akanksha, and K. Sravani, “Youtube Transcript Summarizer Using Flask And Nlp,” Journal of Positive School Psychology, vol. 6, no. 7, pp. 1204–1209, 2022, Available: http://journalppw.com

[7] A. Emad, F. Bassel, M. Refaat, M. Abdelhamed, N. Shorim and A. AbdelRaouf, “Automatic Video summarization with Timestamps using natural language processing text fusion,” 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2021, pp. 0060-0066, doi: 10.1109/CCWC51732.2021.9376115.

[8] C. Zhang and Y. Tian, “Automatic video description generation via LSTM with joint two-stream encoding,” Dec. 2016, doi: 10.1109/icpr.2016.7900081.

[9] “cnn_dailymail · Datasets at Hugging Face,” Jan. 24, 2001. https://huggingface.co/datasets/cnn_dailymail

[10] C. Zhai and S. Massung, “Toward A Unified System for Text Management and Analysis,” Jun. 2016, doi: 10.1145/2915031.2915052.

[11] D. Garg, K. Wong, J. Sarangapani, and S. K. Gupta, Advanced Computing: 10th International Conference, IACC 2020, Panaji, Goa, India, December 5–6, 2020, Revised Selected Papers, Part I. Springer Nature, 2021, vol. 1367

[12] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, Introduction to Algorithms. Mit Press, 1990.