# VIDEO SUMMARISATION USING KEYFRAME EXTRACTION

## INTRODUCTION

As part of this project, we use uniform sampling, image histograms, SIFT, and a convolutional neural network trained on ImageNet to extract low-level features for static keyframe extraction. In order to make the summary fluid and comprehensible for humans, we skim the video around the selected keyframes based on VSUMM.

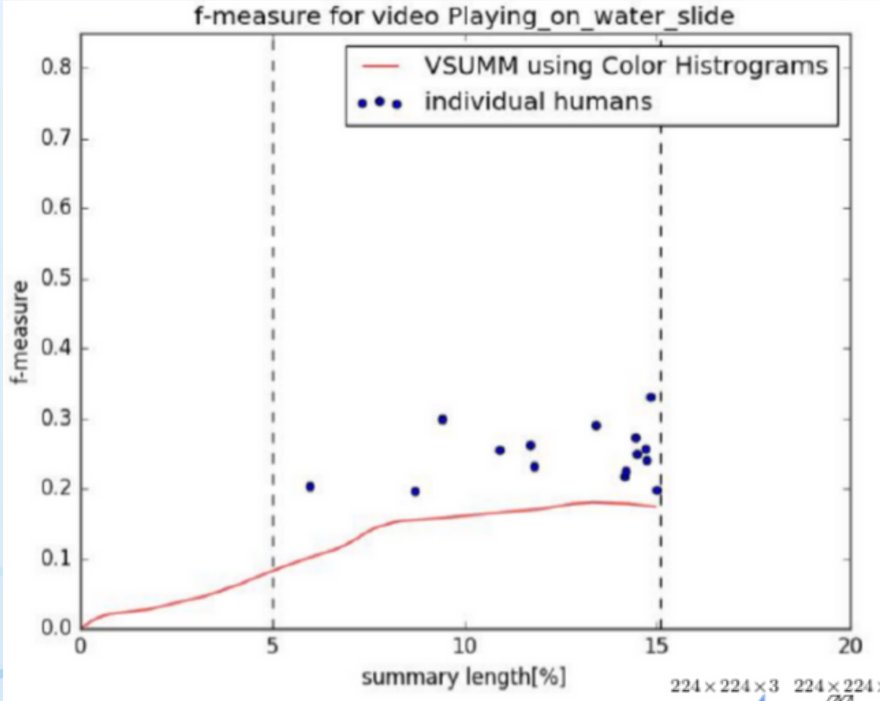
## METHODS

We use the below mentioned methods on our Dataset and comapeare the results to get the best working method.

- Uniform Sampling
- Image Histogram
- Scale Variant Feature Transform
- VSUMM
- ResNet16 on ImageNet

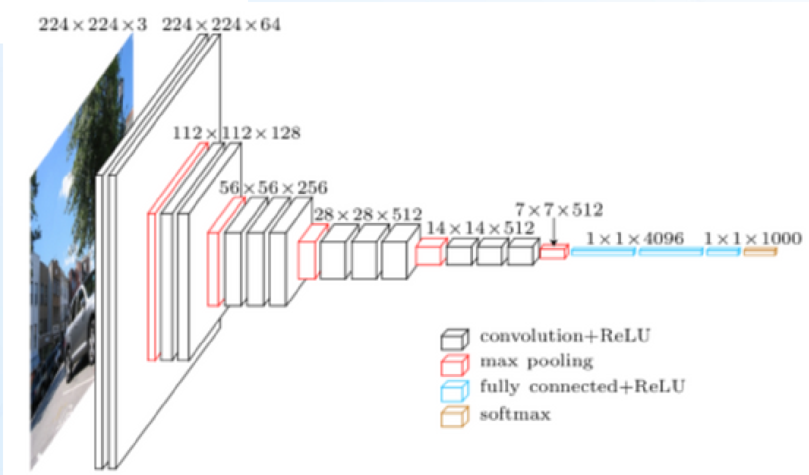

- Aritra Ghosh, Section A, Roll 74, 12021002002137
- Subhojit Ghosh, Section B, Roll 97, 12021002002160

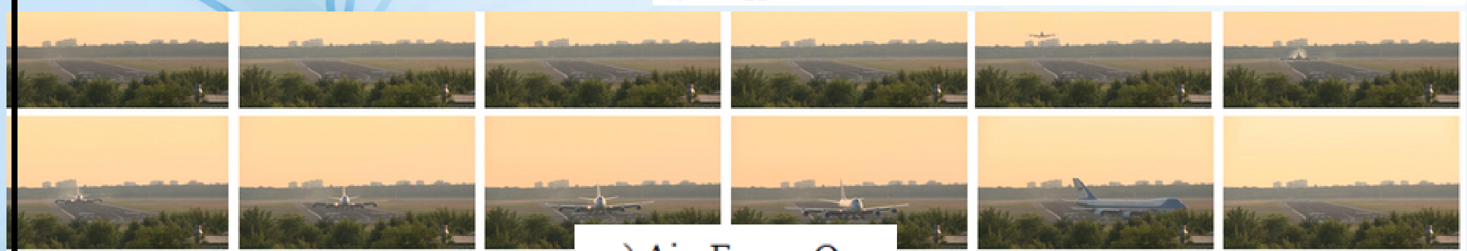

## RESULTS




F-Measure vs VSUMM

Sample CNN



a) Air_Force_One

b) Play Ball

## CONCLUSION

In conclusion, VSUMM effectively condenses lengthy videos into concise summaries but suffers from a drawback - heavy reliance on predefined features and heuristics, potentially missing nuanced content. Integrating natural language processing models can enhance summary quality through coherent and contextually relevant text descriptions.

## ACKNOWLEDGEMENT

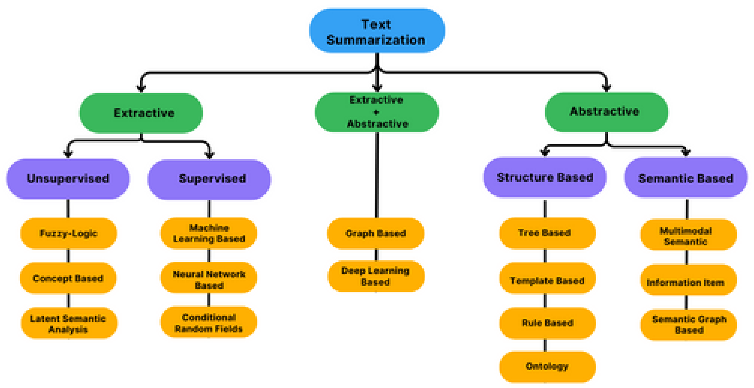

We would like to extend our heartfelt gratitude to our mentor, Dr. Anupam Mondal, for his unwavering guidance and support throughout this journey. We want to express our sincere thanks to our esteemed college, the Institute of Engineering and Management, Kolkata, for providing us with this incredible opportunity.

# An Overview of Summarization

## INTRODUCTION

Text summarisation is the method of reducing the source text into a compact variant, preserving its knowledge and the actual meaning. Here we present a comparative study of various methods of text summarization against our custom task-specific dataset and figure out their advantages and disadvantages. Our goal is to develop a new algorithm that can counter all existing drawbacks.

## METHODS




## RESULTS

| Task | Method | Dataset | Feature Extract | Accuracy | Limitation & Future Work |
|---|---|---|---|---|---|
| Text summarization by sentence selection [212] | Fuzzy logic | DUC 2002 | Tif, SL, TW, SP, TW, etc. | F-measure: 0.498, recall: 0.457 and precision: 0.471 respectively | - Fuzzy logic did not produce promising results<br>- Authors would like to use other learning with fuzzy logic in thei future work |
| Sentence similarity measure [101] | TF-IDF, Word2vec, Sectence2vec | SKE for English and Arabic | AE, VAE, ELM-AE | ROUGE-1: 0.6043 and ROUGE-2: 0.5771 | - Arabic abstractive summarization was not possible<br>- Authors would like to use auto encoders or attention encoders in their future work |
| Sequence labeling [235] | Conditional Random Fields (CRF) | DUC 2001 | SP, SL, Log-likelihood, TW, Indicator words etc. | ROUGE-2 :0.245, F1 Measure: 0.202 of | - Linguistic features were not used<br>- Authors would like to use rhetorical structures for robustness in their future work |
| Short representation from a long document [316] | Latent Semantic Analysis (LSA) | Corpus of Contemporary Arabic (CCA) | Euclidean Distance, Cosine Similarity, Jaccard Coefficient etc. | Entropy 0.1275(approx.), Purity 0.385(approx.) and | - NA |
| Query-based single-document summarization [99] | Restricted Boltzmann Machine (RBM) | SKE, BC3, TAC | Key phrase oriented, Subject oriented summarization | Precision: 0.1694, Recall: 0.2115 and F-score of AE: 0.1816 | - Did not cover multi-documents summarization into this system<br>- Authors would like to extend generic summarization by clustering |

| Task | Method | Dataset | Feature Extract | Accuracy | Limitation & Future Work |
|---|---|---|---|---|---|
| Two-stage extractive and abstractive framework [37] | Highlighting the relevant information | WikiSum (self-generated) | TF-IDF, TextRank, SumBasic, T-D, T-DMACA etc. | ROUGE-L: 38.8 | - Required performance measurement improvement<br>- Authors would like to conduct further research to improve performance |
| Decrease summarization redundancy with DISCOBERT [322] | BanditSum, NeuSum, JECS, BertSum, PNBert, HiBert | New York Times (NYT), CNN, and Daily mail (CNNDM) | EDU | ROUGE-1: 50.00, ROUGE-2: 30.38 and ROUGE-L: 42.70 | - Did not apply different task on long document encoding<br>- Authors would like to explore better encoding methods in their further research |
| To find Solution to information overload [323] | Sentence clustering | DUC 2001, 2002 | Stop words, Stemming | ROUGE-1: 0.47856, ROUGE-2: 0.18528 and F1 measure: 0.48324 | - Author tested one method only<br>- Authors would like to explore more datasets to evaluate the model |
| Enhancing criterion and objective functions [324] | Discrete differential evolution algorithm | DUC 2001, 2002 | Sentence similarity, Sentence clustering | ROUGE-1: 0.45412, ROUGE-2: 0.18982 | - NA |
| Quantitative and qualitative assessment of 15 algorithms [325] | Word scoring, Sentence scoring and graph scoring | CNN Dataset, Blog summarization dataset, SUMMAC dataset | Word Frequency, TF/IDF, Lexical Similarity, SL, SP, Text rank etc. | Average Recall: 0.47, Precision: 0.19, and F-measure is: 0.26 | - Authors would like to compare with other algorithms in their future work |
| [Do not touch I am working from 6.30pm again] Word-sentence co-ranking model [326] | CoRank model | DUC 2002 | Redundancy removal strategies | F-measure: 0.59, ROUGE-1: 0.69 and ROUGE-2: 0.60 | - Did not conduct query-oriented and event-driven summarization tasks<br>- Authors would like to extend corank model multi-document summarization |
| Joint Extractive and Compressive Summarizer (JECS) [327] | Neural Network model | CNN/Daily Mail, New York Times | Sentence and Document Encoder, Compression Encoder, Compression | ROUGE-1: 45.5, ROUGE-2: 25.3 | - NA |

## CONCLUSION

Existing methods face issues such as anaphora, cataphora, interpretability, and readability of lengthy texts. Proposed solutions include improving dataset quality by replacing outliers with human-generated inputs. As this field evolves, we anticipate further advancements in accurately and succinctly summarizing lengthy text documents, significantly impacting information retrieval and dissemination in the digital age.

## ACKNOWLEDGEMENT

We would like to extend our heartfelt gratitude to our mentor, Dr. Anupam Mondal, for his unwavering guidance and support throughout this journey. We want to express our sincere thanks to our esteemed college, the Institute of Engineering and Management, Kolkata, for providing us with this incredible opportunity.