COSE474-2024F: Final Project Proposal
"Video Summarization "

Muhammad Azfar

1. Introduction

Video content is growing rapidly across platforms like YouTube and Netflix making it hard for users to watch long videos in their entirety without losing and missing out important part of the content. Summarizing videos helps users to quickly grasp the main content of the video.

2. Problem definition & challenges

The task of video summarization involves automatically generating an accurate and informative summary of a video.

Some of the challenges with these task is how diverse a video's content is, based on different type of video such as action, dialogues, documentaries video. Each of these video types has different characteristics, and summarizing them requires different techniques

3. Related Works

Pre-trained models has significantly impacted video summarization. Models like **CLIP (Radford et al., 2021)**, which learns from both images and textual descriptions, provide a powerful multimodal framework that can be adapted for video summarization. By using CLIP, it becomes possible to generate both visual and textual summaries based on a video's content. .

4. Datasets

TVSum is a popular dataset for video summarization that contains 50 YouTube videos across 10 categories (e.g., news, sports, documentaries). The dataset is annotated with importance scores provided by human annotators, making it ideal for keyframe-based and segment-based summarization tasks.

5. State-of-the-art methods and baselines

CLIP (Radford et al., 2021) for Multimodal Summarization. CLIP's ability to align images with text descriptions makes it an excellent baseline for video summarization tasks, particularly when the goal is to generate multimodal summaries.

6. Schedule

Research begin and dataset is collected, set up environment. Next preprocess the dataset. Then, Implement video summarization using CLIP and fine-tune Llama for textual summaries. Evaluate results and compare with state-of-the-art methods. A week before submition, refine the model, perform final evaluations, and prepare the report.