

Real-Time Intelligent Sports Game Highlight Creation

Vishal Muthuraja, Dev Thakkar, Rohan Malavathu

vmuthura@purdue.edu, duthakka@purdue.edu, rmalavat@purdue.edu

GitHub: <https://github.com/Rmalavathu/Open-Source-Software-Senior-Design-Project>

ECE 49595O (Open Source Software Senior Design Proposal)

Introduction and Background

Problem Statement

Creating compelling highlight reels or summaries of sports plays involves time-consuming video editing and voiceover work, requiring manual selection of key moments, cutting the video, and adding narration. This process often demands extensive human effort, leading to inefficiencies, time constraints, and potential inconsistencies in the final output.

To streamline and expedite this process, there's a need for an AI-driven solution that automates the compilation of sports play highlights. The solution should incorporate advanced video analysis, automated video cutting, and intelligent voiceover capabilities to efficiently generate engaging highlight reels or summaries, reducing the manual effort and ensuring consistency and quality in the final output.

Solution Statement

Our solution is designed specifically to cater to the fast-paced demands of halftime sports analysis. By leveraging cutting-edge AI technologies, our platform enables instant creation of highlight reels during halftime breaks or live sports events. Through real-time video analysis algorithms, our system swiftly identifies and extracts key moments from the ongoing game or match. These pivotal highlights are automatically compiled into engaging and concise reels within moments, ensuring timely availability for halftime analysis.

Furthermore, our AI-driven voiceover functionality provides instant narration or commentary, adding depth and context to the highlighted plays. This feature facilitates comprehensive analysis and enhances viewer engagement during the brief halftime window. The user-friendly interface allows sports analysts or broadcasters to seamlessly access, customize, and present these highlight reels within minutes, enabling dynamic and informative halftime discussions. Our solution significantly reduces the time and effort traditionally required for video editing and voiceover work, ensuring that valuable halftime moments are maximized for insightful analysis and audience engagement.

Our systems would be trained using full-length sports games, producing comprehensive highlight reels. Feedback from pre-existing shortened games, such as those found on NBA or NFL YouTube channels, refines our model's grasp of impactful moments. This data, coupled with user interactions like crowd noise, commentator tone/volume, and manual labeling, enhances our machine learning algorithms. This iterative process guarantees our system consistently delivers high-quality sports play highlights.

Competitive Analysis

There is a definite interest in the field of AI-generated highlights, and the idea has proven to be meritorious. Tools such as WSC Sports, Magnifi, and Cognitive Mill are currently being utilized by sports leagues and organizations to generate highlights of the most significant moments of a game using AI. Additionally, there is a growing interest in generating highlight reels for various forms of content, with tools like ContentGroove, Valossa AI, and Pictory employing AI to generate highlights. An example of real-world usage includes FOX Sports' "Catch Up with Highlights" feature, which utilized AWS Rekognition to create automated highlights for the games played during the most recent soccer World Cup.

However, the existing tools in the market lack one crucial element – AI-generated commentary integrated with the highlights. This idea remains largely unexplored, and we believe there is significant potential to develop a tool that can provide a competitive edge over existing solutions. AI-generated highlights with AI-synthesized commentary can greatly enhance the user experience by adding a storytelling aspect and providing instantaneous analysis. This can offer users insights that are not apparent with the highlights generated by currently available tools.

Use Cases

- a. User gives a video and has no options. This will create a highlight reel with commentary of a quarter-time
- b. User gives a video with no commentary. This will create a highlight reel with no commentary and video quarter length time
- c. User gives a video with the option of top 5 clips. The program will create the top 5 clips and send them to the user.
- d. Users give a pre-cut video, and the program will create commentary and provide the video with commentary.

Requirements

- Users should be able to upload a video of sports matches via a web portal without feedback explaining the progress of the upload to the user.
- Users should have options to customize the end product, such as:
 - The length of the final highlights reel
 - Toggling the voiceover commentary on or off
 - Specifying a specific number of clips to include in the highlight reels, etc.
- The product should pick out key moments of the game and stitch together a highlights reel.
- The product should intelligently generate commentary to add to the final highlights reel.
- Users should be able to view the end product (highlights reel) upon completion of the processing in the web portal.

Software Architecture

Context Diagram

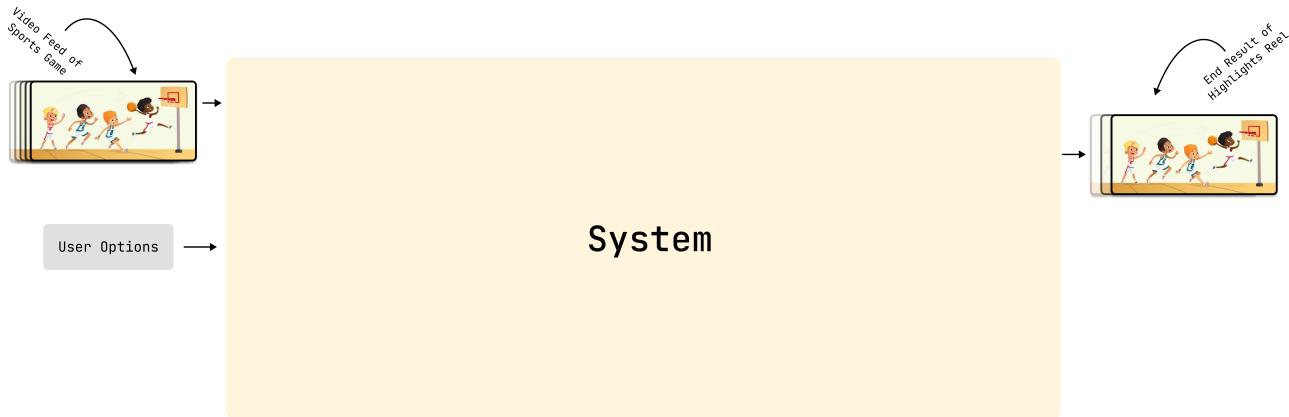


Figure 1: Context Diagram of our Solution

The context diagram describes the interaction between the user input and our system. The user inputs a video feed for a sports game and provides some user options, which can include the length of the highlights clip, voiceover toggle, etc. Our software processes these inputs and provides the user with a final highlight reel that can be downloaded/viewed.

Component Diagram

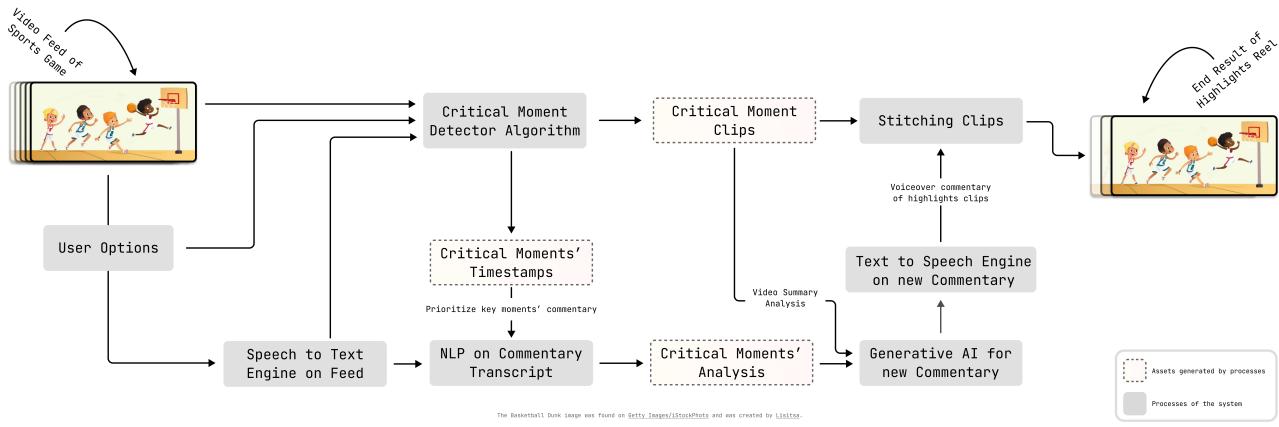


Figure 2: System Diagram of our Solution

The component diagram elaborates on the major parts of the application.

User Options: The User Options feature enables users to customize their output preferences. Users can choose from options such as including commentary, adding commentary only, or selecting only the top N highlights. The default when given no options is a highlight reel with commentary.

Speech to Text Engine of Feed: Transcripts the commentary, so that it can be utilized to analyze the game

NLP on Commentary Transcript: Processing the transcribed commentary to understand the core moments in the game. Given the timestamps for the critical moment detector, it will cut the transcript to just that critical moment, and provide analysis of the clip.

Critical Moment Detector Algorithm: Utilizes the processed transcript as well as analyzes characteristics of the feed to identify critical moments' timestamps

Stitching Clips: Processes the original video based on the critical moments' timestamps to generate cuts and stitches. Additionally, overlays the video with the AI-synthesized commentary once the critical moments are analyzed.

Generative AI for New Commentary: Using the critical moment analysis, utilizing a LLM, generate new commentary for the clips

Text to Speech Engine on new Commentary: Converts the text from the Generative AI, to speech to overlay over the critical moment.

Backend

Host the program on a web server, hosted on the cloud to give an interface for users to input their videos and select options, while also giving them an interface to get their generated video at the end.

Library Analysis

Library	Component	Description
Ffmpeg (OSS)	Stitching Clips	Utilized to cut and stitch clips, and overlay the AI synthesized commentary
Ffmpeg (OSS)	NLP on Commentary Transcript	Cuts the audio using timestamps for the NLP commentary
Whisper V3	Text to Speech Engine on New Commentary	Using the transcript, will create audio commentary
OpenAI	Generative AI for New Commentary	Utilized the OpenAI API and passed in the video clip and critical moment analysis to create commentary
Project Deep-Speech (OSS)	Speech-to-text engine on feed	Converts the commentary from the game to text

Table 1: Library Components and Descriptions