**CHAPTER 1**

**INTRODUCTION**

Sports Video summarization is the condensing of lengthy sports videos into shorter summaries that capture the important events of a game. In an age where people do not have the time to sit for hours together to watch their favorite game, this can help viewers efficiently enjoy the entire game in very little time.

Traditional approaches to summarizing sports videos rely heavily on manual efforts by video editors who sift hundreds of hours of video footage from multiple camera angles and positions in order to identify the significant events and then manually edit this into a highlights video. This process is not only expensive but also time consuming. It also lacks an element of audience engagement. However, this approach can be made better by leveraging the advancements in computer vision, deep learning and natural language processing. Furthermore, the audience reactions can be captured from various social media platforms to help create a better experience for the sports viewers.

The goal of this project is to create a system that can efficiently summarize sports videos utilizing the various advancements in computer vision, deep learning and natural language processing. Further, this project also aims to capture user sentiments by analyzing the tweets related to the game, thereby enabling us to effectively capture audience reactions and insights to create more engaging highlights for the viewer.

The core of this project involves a multi-modal approach that seamlessly integrates multiple modals of data - visual, audio and textual in addition to the twitter data of that game. This ensures that our system is accurate and effective in generating summaries of sports videos and also adds audience engagement and insights into our video summarization process.

Audio Analysis will be crucial for our system in order to help identify events based on valuable insights obtained from the audio track such as crowd intensity, commentator fervor, and player reactions, adding the summaries with a deeper layer of context and emotion.

Visual Analysis will play a key role in ensuring that no event goes unnoticed. The system will analyze visual elements such as scoreboards, player actions, and replays to identify pivotal moments and highlights within the videos.

Textual Analysis will be performed both on the commentary data of the game as well as on the twitter feed provided. This will ensure that there is a human element in summaries generated. The commentary will be analyzed using a large language model to classify the event as significant or not significant.

By incorporating various modalities into our project we aim to create a superior system than the traditional systems being used to generate highlights and additionally by adding audience reactions and insights from twitter we aim to generate a more enriching summary of the sports video.