Soccer Highlight Generator

Anirudh Ramesh Narayanan {an9425@rit.edu}

Mentored by: **Leon Reznik Ph.D.** {lrvcs@rit.edu} and **Mr. Sergei Chuprov** {sc1723@rit.edu}

Introduction

Problem: Soccer while being the most popular sport in the globe, lags notably behind in embracing artificial intelligence for key functions like event detection and statistics tracking.

Motivation: Catering to soccer enthusiasts constrained by time or those who find watching full-length matches impractical, our project seeks to encapsulate the thrill of soccer into a condensed and accessible format. While existing literature predominantly concentrates on audio-visual cues within the game for event classification, our endeavor shifts the paradigm by harnessing broadcaster-augmented data, which provides a rich tapestry of information, to draw insights and conclusions. This innovative approach enhances the precision of event detection.

Solution: Use a video classifier and Yolov8 model to identify the highly exciting moments in the game, and combine them to create an engaging highlight reel.

Methodology

For the experiments in this project, we utilized the SoccerNetv2 dataset, which encompasses video footage from 550 matches, enriched with annotations for 17 distinct types of events.

VideoMAE: The project involved fine-tuning a video classification model, specifically tailored to identify three critical soccer events: corners, free-kicks, and penalties.

YOLOv8: A YOLOv8 object detection model was trained to recognize broadcaster-augmented information, enabling the detection of substitutions and bookings. The model was adept at identifying scoreboards to track goals during the game.

Architecture

The project's workflow is structured as follows:

1. Image Classifier:
   a. Divide input video into 3 parts using the image classifier.
   b. This includes: Overview, Close-up, replays.
   c. Add replays extracted to the highlight reel.
2. VideoMAE:
   a. Use the Overview frames extracted by the image classifier to perform video classification.
   b. If event identified, add clip to highlight reel.
3. Yolov8:
   a. Optimize event identification by processing every tenth frame.
   b. Send Scoreboard ROI to AWS Rekognition to extract text in image to track goals every 10 seconds (250 frames).
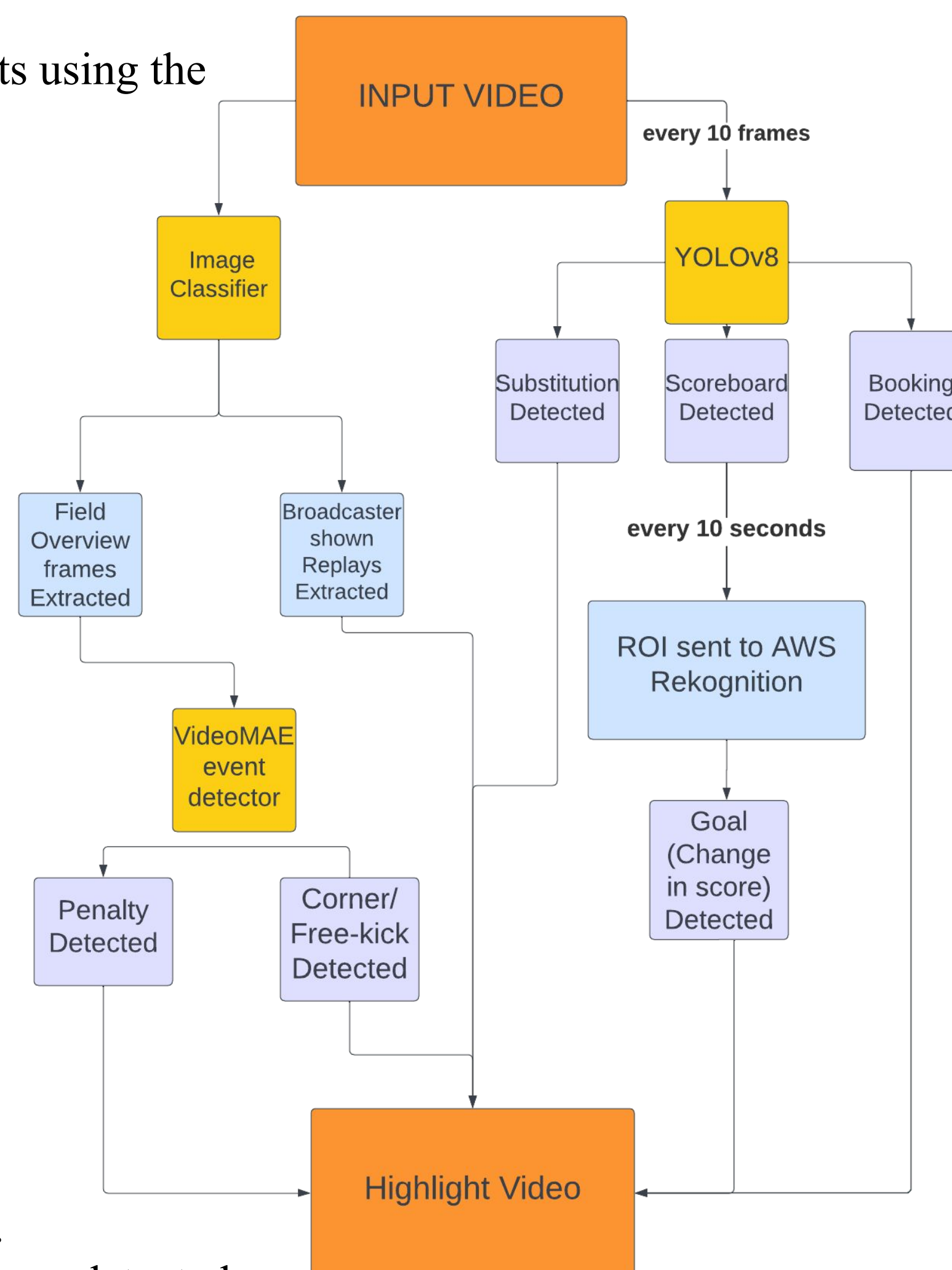   c. If event identified or score change detected, add clip to final highlight reel.

Figure : Project Pipeline

Results

- Image Classifier: The image classifier achieved an accuracy of 97% on the test set and when applied to 45 minutes test video, achieved precision of 97% in extracting overview frames, and 100% precision in extracting replays.
- VideoMAE: The Video Classifier achieved accuracy of 92% on the test set and when applied to a 45 minute test video, achieved a precision of 67% in identifying the events.
- YOLOv8: The Yolov8 model achieved an accuracy of 91.2% on the test set, and when applied to a 45 minutes test video, extracted events with a precision of 82%.

See the Implementation Details

Repository available at : https://github.com/Anirudhrn98/SoccerHighlightGenerator

Comparison

Video Information	Original Highlight	Methodology Results
Highlight Length	2:07	10:35
Goal	2	2
Foul	2	2
Free-Kick	0	3
Corner	2	5
Card	1	5
Substitution	1	4
False-Positives / Repeat	0	6
Attacks	0	18

Broadcaster-Augment Information

(a) Substitution

(h) Transition

(b) Booking

(e) Scoreboard

Importance of Data Quality

Class	144p Accuracy (%)	240p Accuracy (%)
Setpiece	75.70	92.86
Transition	96.09	100.00
Penalty	80.37	91.67
Openplay	71.21	98.70
Average	78.97	95.34

Conclusion

The combination of the YOLOv8 model with a finely tuned, advanced video classifier, pretrained on a comprehensive dataset, demonstrated effective results in generating soccer game highlights. Research indicates that employing higher quality video datasets could further enhance these outcomes.

References

[1] He, Kaiming, et al. "Masked autoencoders are scalable vision learners." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
[2] Cioppa, Anthony, et al. "Camera calibration and player localization in soccernet-v2 and investigation of their representations for action spotting." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.