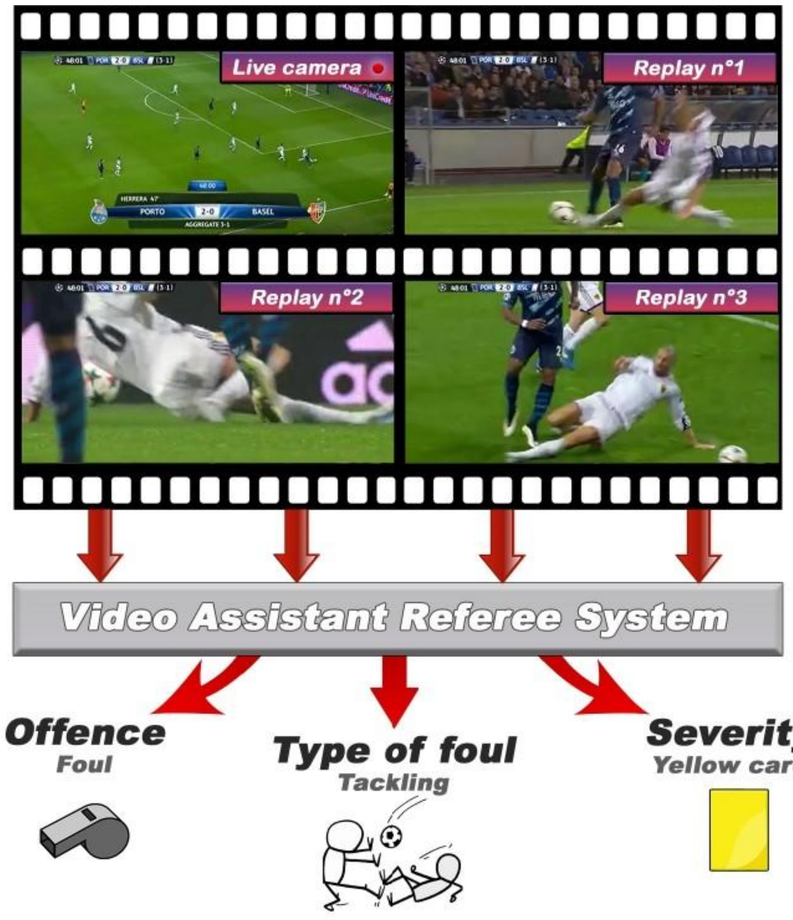
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

VARS: Video Assistant Referee System for Automated Soccer Decision Making from Multiple Views

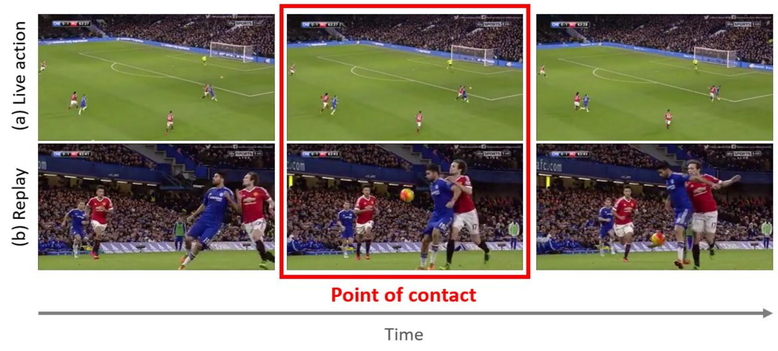
Jan Held*, Anthony Cioppa*, Silvio Giancola*,
Abdullah Hamdi, Bernard Ghanem and Marc Van Droogenbroeck

Code
www.github.com/SoccerNet/VARS

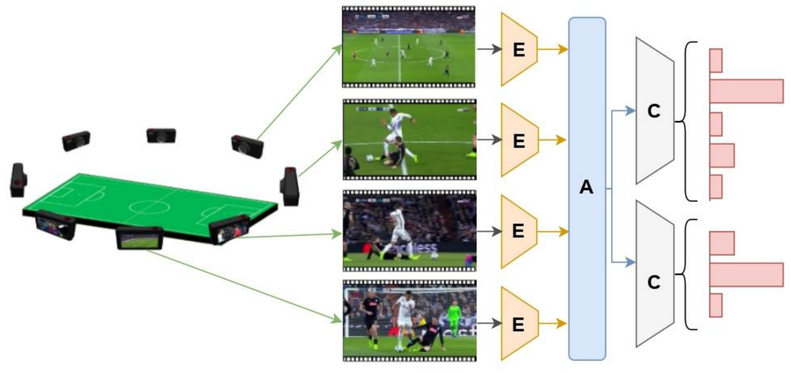


SoccerNet-MVFouls Dataset

- **3901 actions**, composed of **at least two video clips**.
- **10 annotated properties** describing the characteristics of the foul.
- Annotated by a **professional soccer referee**.



VARS architecture

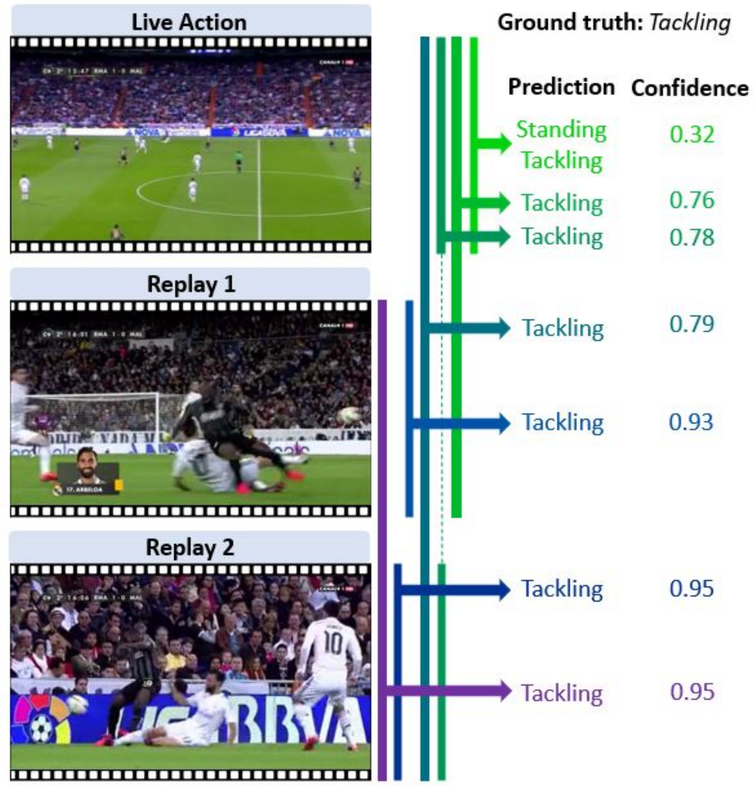


- **Feature encoder (E)**: extracts spatial + temporal features for each view using typical video encoders.
- **Aggregation (A)**: max or mean pooling of the view features to a single multi-view representation.
- **Classification heads (C)**: multi-task classification.

Single VS Multi-View Analysis

Performance	Viewing Setup				
	L	R1	L+R1	R1+R2	L+R1+R2
Acc_{T1}	0.31	0.47	0.50	0.56	**0.57**
Acc_{T1}@2	0.54	0.68	0.70	0.69	**0.72**
BA_{T1}	0.29	0.38	0.36	**0.44**	0.39
Acc_{T2}	0.38	0.39	**0.43**	0.39	0.40
Acc_{T2}@2	0.67	0.70	0.72	0.73	**0.75**
BA_{T2}	0.38	0.27	0.34	0.27	**0.39**

T1: Fine-grained foul classification, T2: Foul severity classification, L: Live view, R: Replay view, Acc: Accuracy, BA: Balanced Accuracy



Contributions

- We release **SoccerNet-MVFouls**, a new **multi-view video dataset** for soccer video understanding.
- We propose **two novel tasks** on this dataset:
  - Fine-grained foul classification.
  - Foul severity classification.
- We build a **V**ideo **A**ssistant **R**eferee **S**ystem (**VARS**), a new **multi-view video recognition model** for classifying fouls and their severity.
- We study the **effect of the type** and **number of views** on the performance of our VARS.

Conclusion

- VARS can **accurately** recognize foul properties.
- VARS offers an **unbiased** and **reliable** decision-making process for soccer matches.