

VARS: VIDEO ASSISTANT REFEREE SYSTEM FOR AUTOMATED SOCCER DECISION MAKING FROM MULTIPLE VIEWS

Ngô Tất Tố - 250201032

Tóm tắt

- Lớp: CS2205.SEP2025
- Link Github của nhóm: <https://github.com/tatto2k3/CS2205.SEP2025>
- Link YouTube video: <https://youtu.be/bIGfUikCiMs>



Giới thiệu

Bối cảnh

- **Công nghệ VAR** đã cải thiện đáng kể độ chính xác và tính công bằng trong bóng đá hiện đại.
- Tuy nhiên, **VAR** đòi hỏi chi phí cao và đội ngũ trọng tài chuyên nghiệp, nên chỉ được triển khai ở các giải đấu lớn.
- Nhiều giải bán chuyên và nghiệp dư chưa có **công nghệ VAR**, dẫn đến tranh cãi và sai sót trong các tình huống phạm lỗi.



Giới thiệu

Các phương pháp phân tích bóng đá hiện nay chủ yếu:

- chỉ sử dụng video từ một góc nhìn, hoặc
- chỉ nhận diện sự kiện mà chưa hỗ trợ ra quyết định theo luật thi đấu.

Trong khi đó, trọng tài VAR cần:

- xem lại cùng một tình huống từ nhiều góc camera,
- đồng thời xác định loại phạm lỗi và mức độ xử phạt (không thẻ, thẻ vàng, thẻ đỏ).

⇒ **Bài toán đặt ra:**

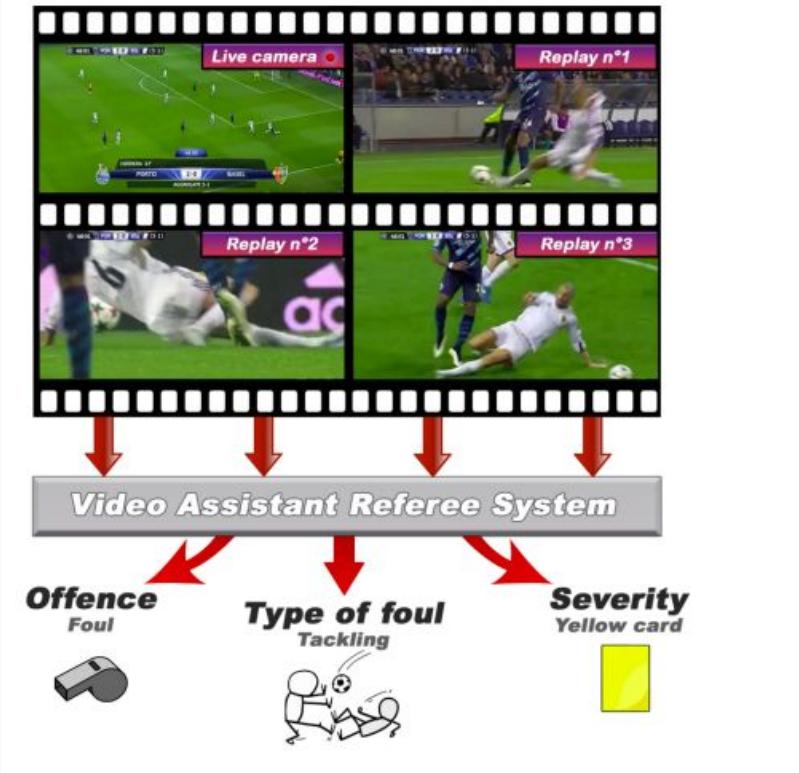
Làm thế nào để xây dựng một hệ thống có thể tự động hỗ trợ trọng tài đưa ra quyết định chính xác từ video đa góc nhìn, với kết quả đầu ra rõ ràng và phù hợp với luật bóng đá?

Giới thiệu

Giải pháp

Đề tài đề xuất VARS – Video Assistant Referee System, một hệ thống hỗ trợ trọng tài dựa trên:

- phân tích video đa góc nhìn (multi-view video analysis),
- học đa nhiệm (multi-task learning) để dự đoán đồng thời loại phạm lỗi và mức độ xử phạt.



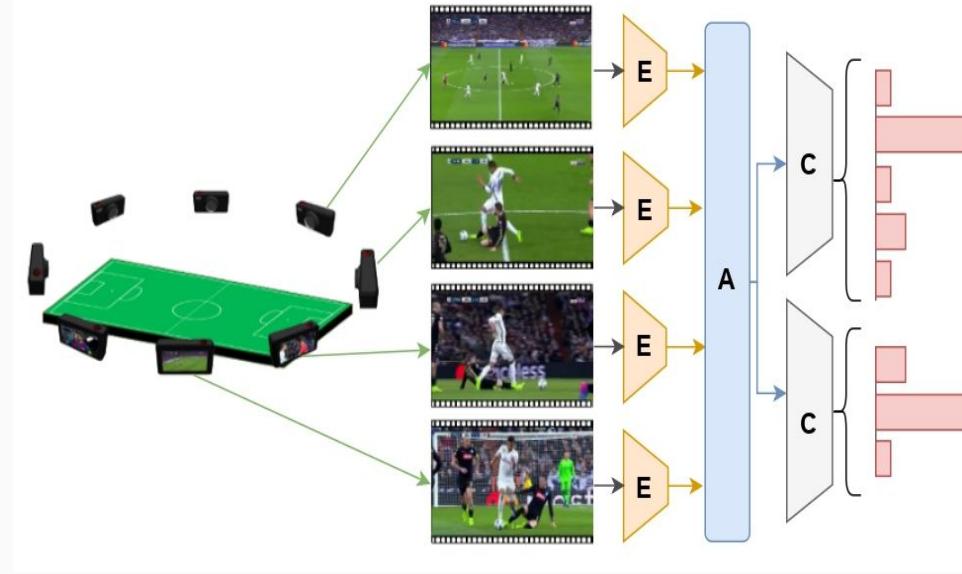
Mục tiêu

- **Xây dựng hệ thống** có khả năng **phân tích video đa góc nhìn** để hỗ trợ trọng tài trong các tình huống phạm lỗi.
- **Tự động dự đoán các quyết định theo luật bóng đá**, bao gồm:
 - Xác định có phạm lỗi hay không
 - Phân loại loại phạm lỗi
 - Đánh giá mức độ thẻ phạt
- **Đánh giá hiệu quả qua việc sử dụng nhiều góc camera** so với một góc nhìn, nhằm chứng minh tính cần thiết và lợi ích của việc phân tích video đa góc trong hỗ trợ ra quyết định trọng tài.

Nội dung và Phương pháp

Kiến trúc tổng quan

1. **Input:** một tình huống phạm lỗi được quay từ n góc nhìn khác nhau.
2. **E (Video Encoder):** mỗi góc nhìn đi qua encoder để tạo vector đặc trưng không gian–thời gian.
3. **A (Multi-view Aggregation):** gộp các vector đặc trưng thành một biểu diễn chung bằng *mean pooling* hoặc *max pooling*.
4. **C (Classifier Head):** đưa biểu diễn chung vào head phân loại để ra dự đoán cuối cùng.



Kết quả dự kiến

- Multi-view video analysis cho hiệu quả cao hơn single-view trong nhận diện phạm lỗi.
- Học đa nhiệm giúp cải thiện dự đoán mức độ xử phạt.
- Hệ thống đạt hiệu năng định lượng được và có khả năng ứng dụng thực tế trong hỗ trợ trọng tài.

Tài liệu tham khảo

- [1] Adrià Arbués Sangüesa, Adrián Martín, Javier Fernández, Coloma Ballester, Gloria Haro: Using Player's Body-Orientation to Model Pass Feasibility in Soccer. CVPR Workshops 2020: 3875-3884
- [2] Alejandro Cartas, Coloma Ballester, Gloria Haro: A Graph-Based Method for Soccer Action Spotting Using Unsupervised Player Classification. MMSports@MM 2022: 93-102
- [3] Anthony Cioppa, Silvio Giancola, Adrien Deliège, Le Kang, Xin Zhou, Zhiyu Cheng, Bernard Ghanem, Marc Van Droogenbroeck: SoccerNet-Tracking: Multiple Object Tracking Dataset and Benchmark in Soccer Videos. CVPR Workshops 2022: 3490-3501