



Báo cáo Đồ Án Cuối Kỳ Phân tích bóng đá

Môn: Nhập môn học máy

Lớp: 21KHDL1

Giảng viên: Bùi Tiến Lên

Bùi Duy Đăng

Trần Trung Kiên

Thông tin nhóm

Nhóm: 7

Thành viên:

- ❖ Trần Nguyên Huân – 21127050
- ❖ Doãn Anh Khoa – 21127076
- ❖ Nguyễn Minh Quân – 21127143
- ❖ Nguyễn Phát Đạt - 21127240



Nội dung



01

Giới thiệu dự án

02

**Mô hình
YOLOv8**



03

Flow Chart

04

**Huấn luyện
mô hình và
Thách thức**

05

**Hướng khắc
phục**



01

Giới thiệu dự án



Tổng quan

- Web phát triển để phân tích trận đấu bóng đá.
- Sử dụng các kỹ thuật Học sâu và Thị giác máy tính để nhận diện cầu thủ mỗi đội, trọng tài và quả bóng.
- Đánh giá hiệu suất của họ và tạo ra bản đồ chiến thuật.



Mục tiêu



- Nhận diện và phân loại cầu thủ: Nhận diện cầu thủ và quả bóng từ video trận đấu, sau đó phân loại cầu thủ vào các đội tương ứng.
- Tạo bản đồ chiến thuật: Ánh xạ vị trí của các cầu thủ và quả bóng trên video vào bản đồ chiến thuật trực quan.
- Theo dõi quả bóng và chiến thuật: theo dõi quả bóng qua các khung hình để phân tích các tình huống trận đấu.



02

Mô hình YOLOv8

Mô hình YOLOv8

.....

1

**Giới thiệu
YOLOv8**

2

**Kiến trúc và
thuật toán được
sử dụng**

3

**Ưu điểm và
hạn chế**

Giới thiệu YOLOv8

Là một mô hình nhận dạng đối tượng dựa trên mạng CNN được phát triển bởi Joseph Redmon và nhóm nghiên cứu của ông tại Đại học Washington.

Sử dụng kỹ thuật Attention để cải thiện khả năng nhận dạng đối tượng của mô hình

Tích hợp cơ chế tự động điều chỉnh tỷ lệ tăng kích thước của hình ảnh đầu vào (AutoScale).

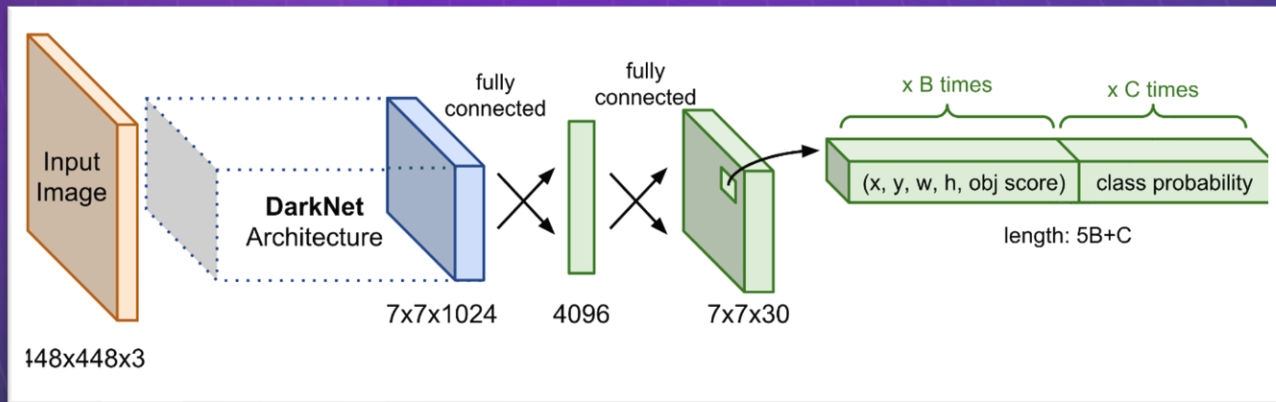
Giới thiệu YOLOv8

Hỗ trợ giám sát bằng video (Video Supervision)

Tích hợp công nghệ Ensemble.

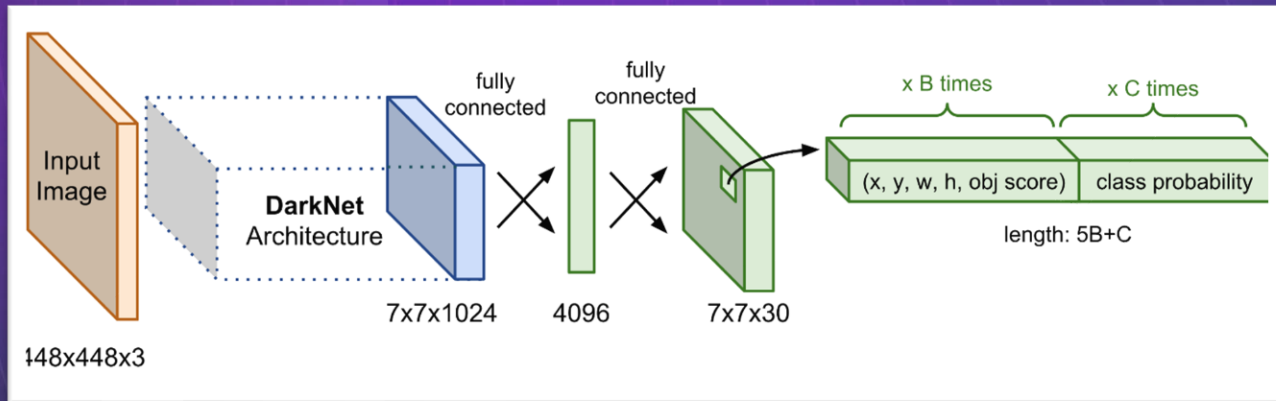
Tính năng điều chỉnh tỷ lệ tự động (AutoAnchor)

Kiến trúc mô hình YOLOv8



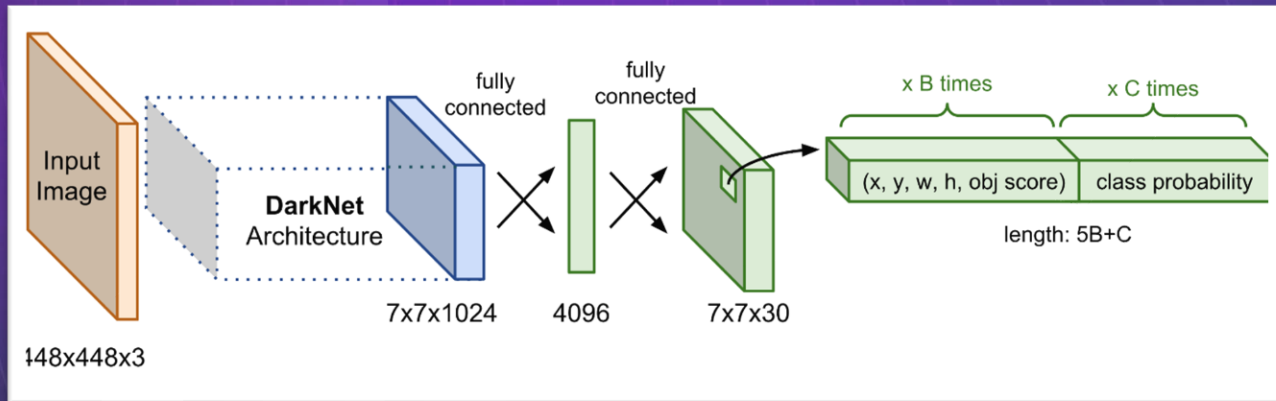
- Lớp đầu vào: nhận hình ảnh đầu vào.
- CNN layers: sử dụng mạng nơ-ron tích chập để rút trích đặc trưng của hình ảnh. Trong đó Darknet là cấu trúc mặc định cho YOLOv8.

Kiến trúc mô hình YOLOv8 (tt)



- Fully connected layers: tạo ra các dự đoán về đối tượng có mặt trong hình ảnh.
- Detection layer: dự đoán về vị trí và class các đối tượng được tạo ra.

Kiến trúc mô hình YOLOv8 (tt)



- Lớp đầu ra: trả về một tensor chứa các dự đoán về vị trí và class của các đối tượng.

Thuật toán trong YOLOv8

Hàm lỗi trong mô hình YOLO được tính trên việc dự đoán và nhãn mô hình, bao gồm 3 thành phần:

- Classification loss: Độ lỗi của việc dự đoán loại nhãn của đối tượng.

$$L_{classification} = \sum_{i=0}^{S^2} \mathbb{I}_i^{obj} \sum_{c \in Class} (p_i(c) - \hat{p}_i(c))^2$$

Trong đó:

- \mathbb{I}_i^{obj} : bằng 1 nếu ô vuông đang xét có đối tượng ngược lại bằng 0.
- $\hat{p}_i(c)$: xác suất có điều kiện của lớp c tại ô vuông tương ứng mà mô hình dự đoán.

Thuật toán trong YOLOv8

Hàm lỗi trong mô hình YOLO được tính trên việc dự đoán và nhãn mô hình, bao gồm 3 thành phần:

- Classification loss: Độ lỗi của việc dự đoán loại nhãn của đối tượng.
- Localization loss: Độ lỗi của dự đoán tọa độ tâm, chiều dài, rộng của boundary box (x, y, w, h).

$$L_{localization} = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2]$$

Thuật toán trong YOLOv8

Hàm lỗi trong mô hình YOLO được tính trên việc dự đoán và nhãn mô hình, bao gồm 3 thành phần:

- Classification loss: Độ lỗi của việc dự đoán loại nhãn của đối tượng.
- Localization loss: Độ lỗi của dự đoán tọa độ tâm, chiều dài, rộng của boundary box (x, y, w, h).
- Confidence loss: Độ lỗi của việc dự đoán bounding box đó chứa object so với nhãn thực tế tại ô vuông đó.

$$L_{confidence} = \sum_{i=0}^{S^2} \sum_{j=0}^B (\mathbb{I}_{ij}^{obj} + \lambda_{noobj} (1 - \mathbb{I}_{ij}^{obj})) (C_{ij} - \hat{C}_{ij})^2$$

- C_{ij} : điểm tin cậy của ô thứ i = P(contain object) * IoU (predict bbox, ground truth bbox)
- \hat{C}_{ij} : điểm tự tin dự đoán

Thuật toán trong YOLOv8

Hàm lỗi trong mô hình YOLO được tính trên việc dự đoán và nhãn mô hình, bao gồm 3 thành phần:

- Classification loss: Độ lỗi của việc dự đoán loại nhãn của đối tượng.
- Localization loss: Độ lỗi của dự đoán tọa độ tâm, chiều dài, rộng của boundary box (x, y, w, h).
- Confidence loss: Độ lỗi của việc dự đoán bounding box đó chứa object so với nhãn thực tế tại ô vuông đó.

$$L_{total} = L_{classification} + L_{localization} + L_{confidence}$$

Thuật toán trong YOLOv8 (tt)

Các thuật toán tối ưu trong mô hình bao gồm: SGD, Adam, AdamW, NAdam, RAdam, RMSProp, ...

Bên cạnh đó, YOLO sử dụng SGD với hệ số momentum và Adam với hệ số beta1.

YOLO có sử dụng weight decay (L2 regularization) để tránh overfitting.

Ngoài ra, YOLO cũng có thể sử dụng các kỹ thuật để tránh overfitting: data augmentation và dropout.

Ưu điểm của mô hình YOLOv8

Tốc độ

được đánh giá là ***nhANH chóng*** và ***thời gian phản hồi thấp***, giúp xử lý các tác vụ nhận diện đối tượng và phân-segment ảnh trong thời gian thực.

Độ chính xác

được xây dựng trên các tiến bộ về học sâu và thị giác máy tính, ***đảm bảo độ chính xác cao*** trong việc nhận diện đối tượng.

Sự linh hoạt

hỗ trợ việc nhận diện đối tượng và phân-segment trên cả ***GPU*** và ***CPU***, tận dụng các công nghệ như TensorRT của Nvidia và OpenVino của Intel.

Hạn chế của mô hình YOLOv8

Dữ liệu

Cần phải được huấn luyện trên một tập dữ liệu **đủ lớn** và **đa dạng** để đạt được hiệu quả cao nhất.

Tính toán

Yêu cầu các **tài nguyên tính toán cao** để đạt được tốc độ xử lý nhanh và chính xác.



03

Flow Chart

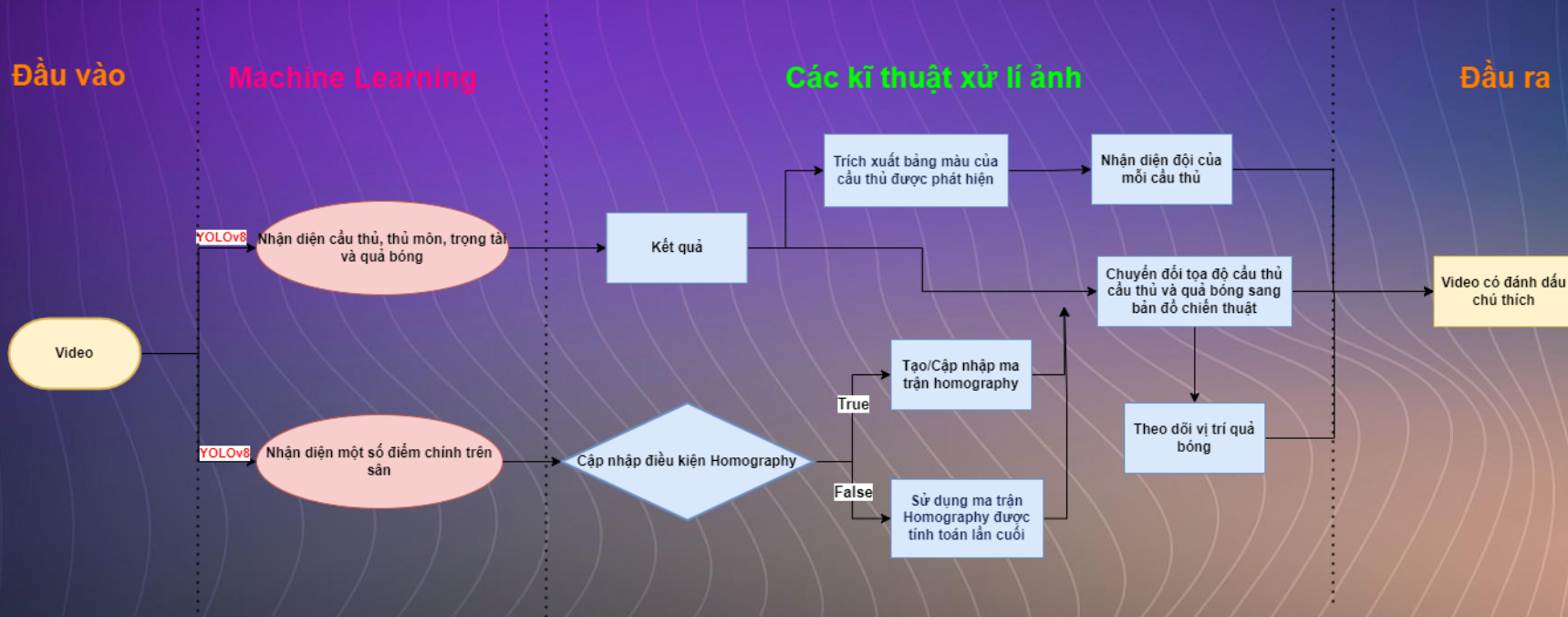
Flow Chart

Đầu vào

Machine Learning

Các kĩ thuật xử lí ảnh

Đầu ra



Mô tả các bước trong FlowChart

1

Video đầu vào

Video đầu vào được chuyển đến hai mô hình nhận diện đối tượng

2

Nhận diện đối tượng

1 mô hình sẽ nhận diện cầu thủ, trọng tài và quả bóng
1 mô hình sẽ nhận diện các điểm chính trên sân (keypoints)

3

Nhận diện đội

Kết quả mô hình thứ nhất được sử dụng để trích xuất màu sau đó tính toán khoảng cách
→ đưa ra dự đoán

Mô tả các bước trong FlowChart (tt)

4

Tạo bản đồ chiến thuật

Kết quả từ mô hình phát hiện điểm trên sân được sử dụng để ánh xạ vị trí vào bản đồ chiến thuật

5

Theo dõi quả bóng

Khi có tọa độ của cầu thủ và quả bóng trên bản đồ chiến thuật thì có thể dễ dàng theo dõi chuyển động của quả bóng

6

Video đầu ra

Hiển thị ra video với từng khung hình đã được chú thích và thông tin tương ứng trên bản đồ chiến thuật

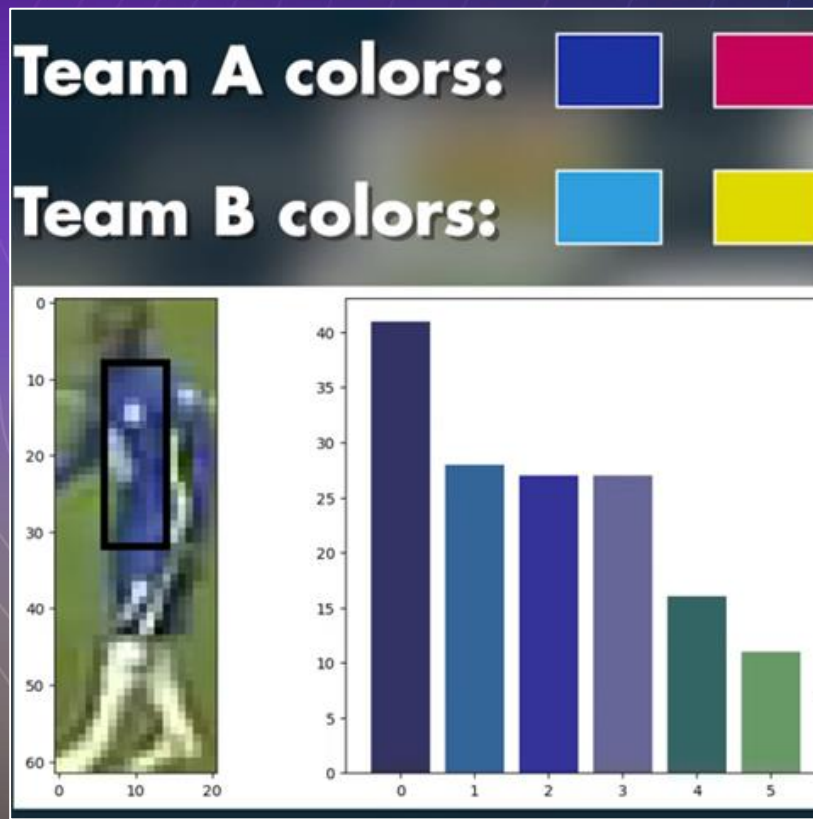
Mô hình nhận diện cầu thủ



Nhận diện đội của mỗi cầu thủ

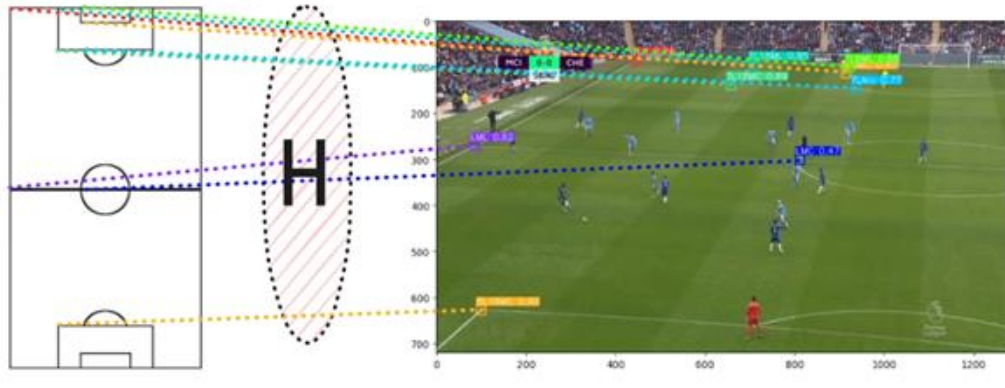
Bằng cách xác định màu sắc ban đầu của 2 đội và bảng màu trích xuất từ mỗi cầu thủ:

- Sau đó tính khoảng cách giữa các màu cầu thủ với màu đã được xác định của từng đội.
- Như vậy có thể gán mỗi cầu thủ vào đội một cách chính xác



Nhận diện các keypoints và trực quan trên bản đồ chiến thuật

$$\begin{bmatrix} wx' \\ wy' \\ w \\ \mathbf{p}' \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ \mathbf{H} \end{bmatrix} \begin{bmatrix} x \\ y \\ l \\ \mathbf{p} \end{bmatrix}$$

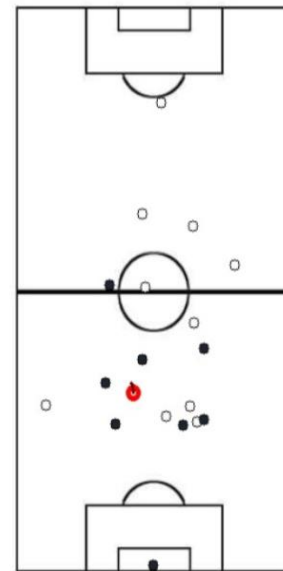


- Kết quả từ mô hình nhận diện các keypoints trên sân được sử dụng để tính toán một ma trận homography.
- Homography - ma trận ánh xạ vị trí các điểm chính trên bản đồ chiến thuật với vị trí các điểm được phát hiện trên khung hình.

Video xuất ra



Tactical Map



03

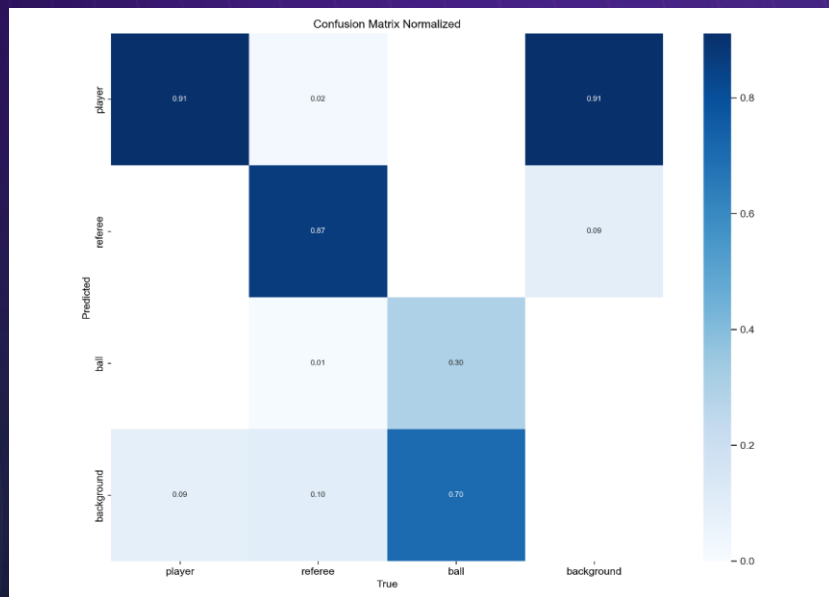
Huấn luyện mô hình và thách thức

Sơ lược về dữ liệu và mô hình sử dụng

- **Mô hình**: Pre-trained YOLOv8, đã được tinh chỉnh các siêu tham số.
 - Mô hình 1: Nhận diện cầu thủ, trọng tài và quả bóng
 - Mô hình 2: Nhận diện một số điểm chính trên sân (keypoints)
- **Bộ dữ liệu**: bao gồm 500 ảnh được trích từ video góc quay chiến thuật trong bóng đá và được gán nhãn thủ công trên Roboflow ([link video được trích xuất](#)).
- **Các lớp đối tượng**: Cầu thủ, trọng tài, quả bóng, và các keypoints trên sân

Khám phá Confusion Matrix của mô hình 1

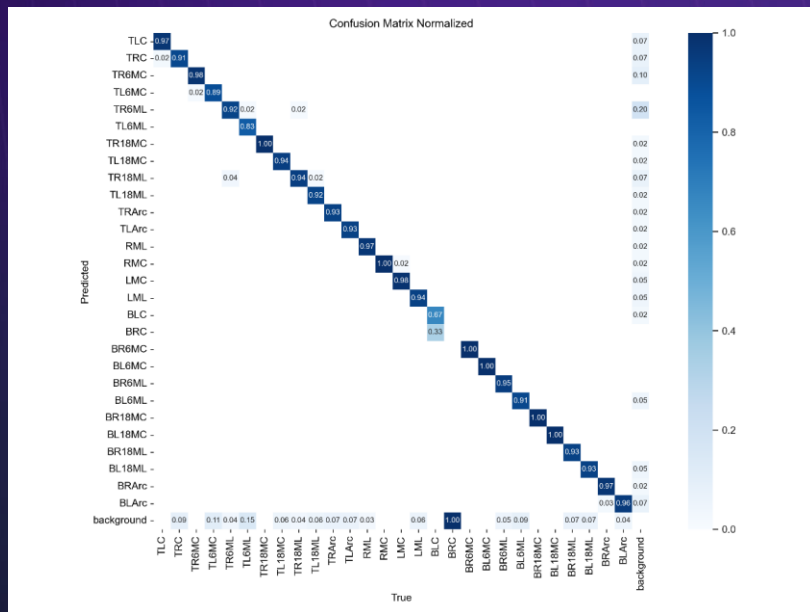
Players detection model



- Có thể thấy thông qua confusion matrix của mô hình 1 kết quả tương đối tốt cho tất cả các lớp khác nhau.
- Ngoại trừ lớp bóng nơi chỉ có 25% của số lượng bóng được dự đoán có mặt trong tập validation → Chỉ số Recall cho lớp bóng thấp
- Điều này cũng khá dễ hiểu vì dữ liệu về lớp bóng bị thiếu khá nhiều. → Đây là vấn đề của mất cân bằng dữ liệu

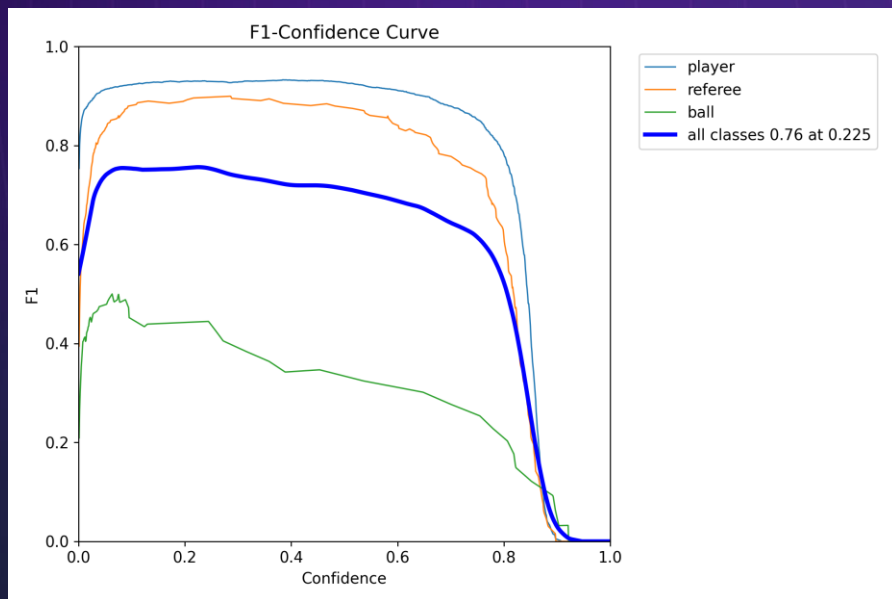
Khám phá Confusion Matrix của mô hình 2

Keypoints detection model



- Đối với confusion matrix của mô hình 2 ta có thể thấy kết quả tương đối tốt trên hầu hết các điểm.
- Chỉ riêng 2 lớp điểm BLC và BRC bị thiếu khá nhiều nên khả năng dự đoán cho 2 điểm này khá thấp
- Tuy vậy điều này không ảnh hưởng quá nhiều đến kết quả vì có tổng cộng 29 điểm trên sân đôi lúc sai từ 1 đến 2 điểm vẫn có thể trực quan thành bản đồ chiến thuật tốt.

Phân tích F1-Score của mô hình

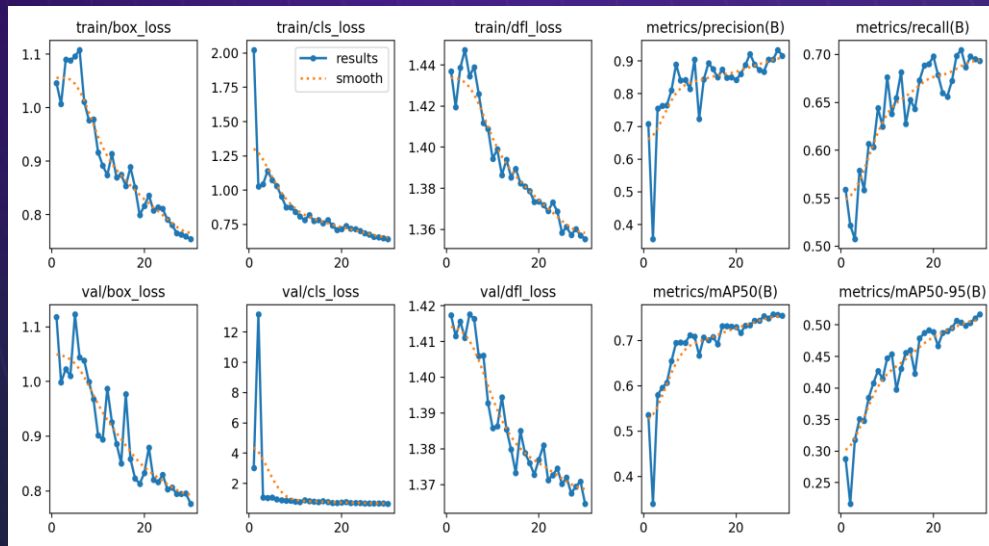


- Qua đường cong F1-Confidence biểu diễn điểm F1 so với mức độ tự tin của các lớp khác nhau.
- Ta có thể thấy lớp “Player” và lớp “Referee” có điểm F1 cao và ổn định qua các mức độ tự tin.
- Lớp “Ball” có điểm F1 thấp và biến đổi đáng kể qua các mức độ tự tin khác nhau.

Vấn đề gặp phải trong mô hình

- Dữ liệu lớp bóng bị thiếu khá nhiều trong tập dữ liệu dẫn đến hiệu suất mô hình cho lớp bóng thấp hơn đáng kể so với các lớp còn lại.
- Qua đó có thể thấy rằng sự cân bằng của lớp ảnh hưởng đặc biệt đến độ Recall cho lớp đó.

Đánh giá chung về kết quả của mô hình



Qua biểu đồ huấn luyện của mô hình có thể thấy mô hình có vẻ đã cải thiện hiệu suất theo thời gian:

- Box_loss và idf_loss giảm theo thời gian trong cả hai giai đoạn, cho thấy mô hình đang học và cải thiện theo dữ liệu khá tốt.
- Độ chính xác (precision), độ nhớ (recall) và mAP50 tăng lên, cho thấy mô hình đang cải thiện khả năng dự đoán của nó



04 Hướng khắc phục



Một số cách để khắc phục



Bộ dữ liệu lớn hơn

Huấn luyện ảnh mỗi lớp nên lớn hơn hoặc bằng 1500 ảnh

Đa dạng ảnh

Ảnh nên từ nhiều thời điểm, ánh sáng, góc nhìn,... khác nhau

Nhãn nhất quán

Tất cả các thể hiện của tất cả các lớp trong tất cả các ảnh đều phải được gán nhãn



Một số cách để khắc phục (tt)

Độ chính xác nhãn

Nhãn phải được gắn sát với mỗi đối tượng. Không nên có khoảng trống nào nên tồn tại giữa một đối tượng và khung giới hạn của nó (bounding box)

Xác minh nhãn

Xem thêm file train_batch*.jpg khi bắt đầu huấn luyện để xác minh nhãn của bạn xuất hiện chính xác

Ảnh nền

Nên huấn luyện thêm dữ liệu có hình ảnh nền – các hình ảnh không có đối tượng được thêm vào để huấn luyện để giảm False Positive (FP)



**Cảm ơn vì các
giảng viên đã
xem**

