

Kaggleのアメフトコンペで負けた話

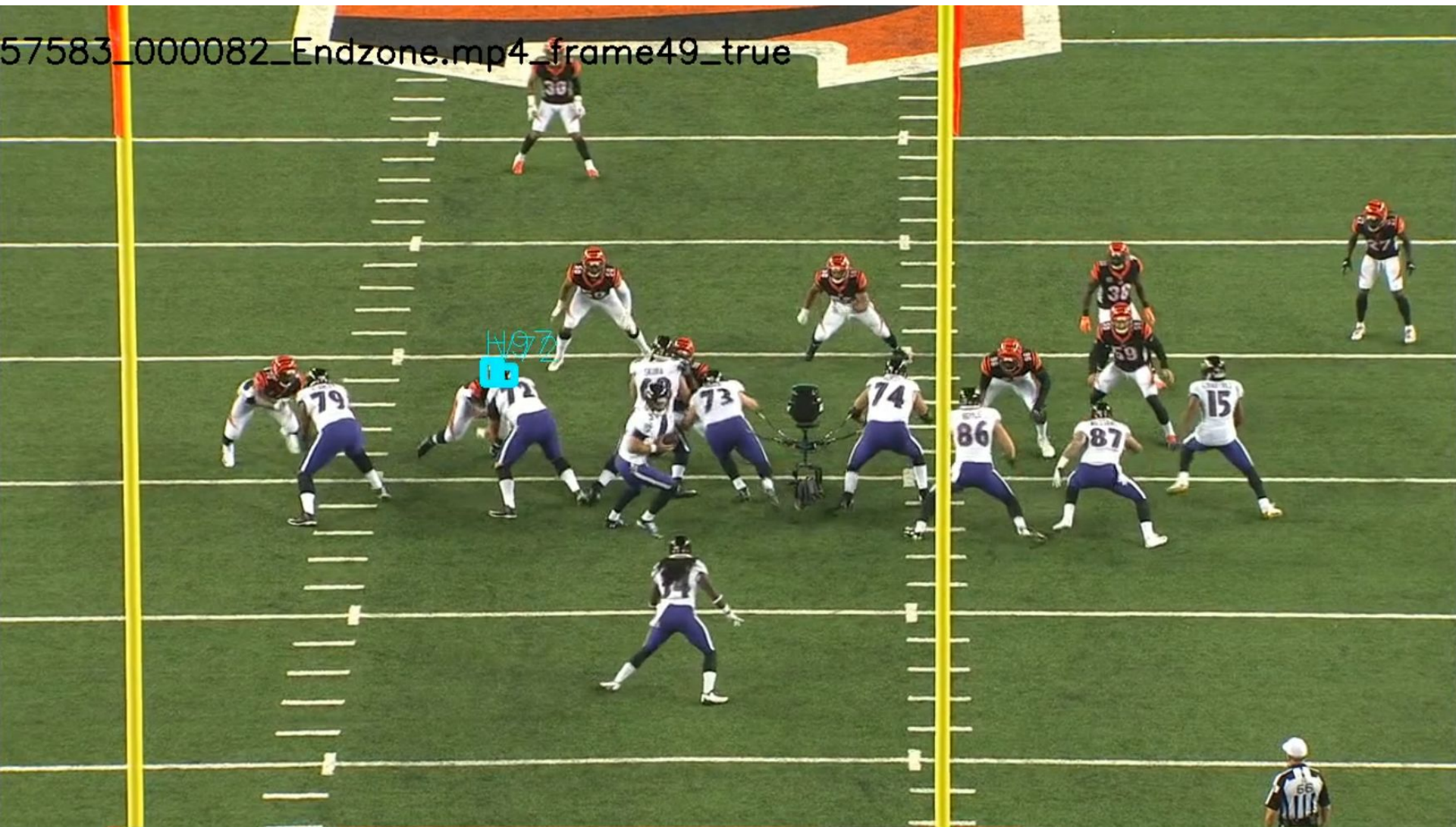
もくじ

1. コンペ概要
2. 結果
3. 私の解法
4. なんで負けたか
5. 上位の解法
6. おすすめコンペ紹介

1.コンペ概要

アメフトのプレイ動画中でヘルメットの衝突した瞬間をBBOX付きで予測

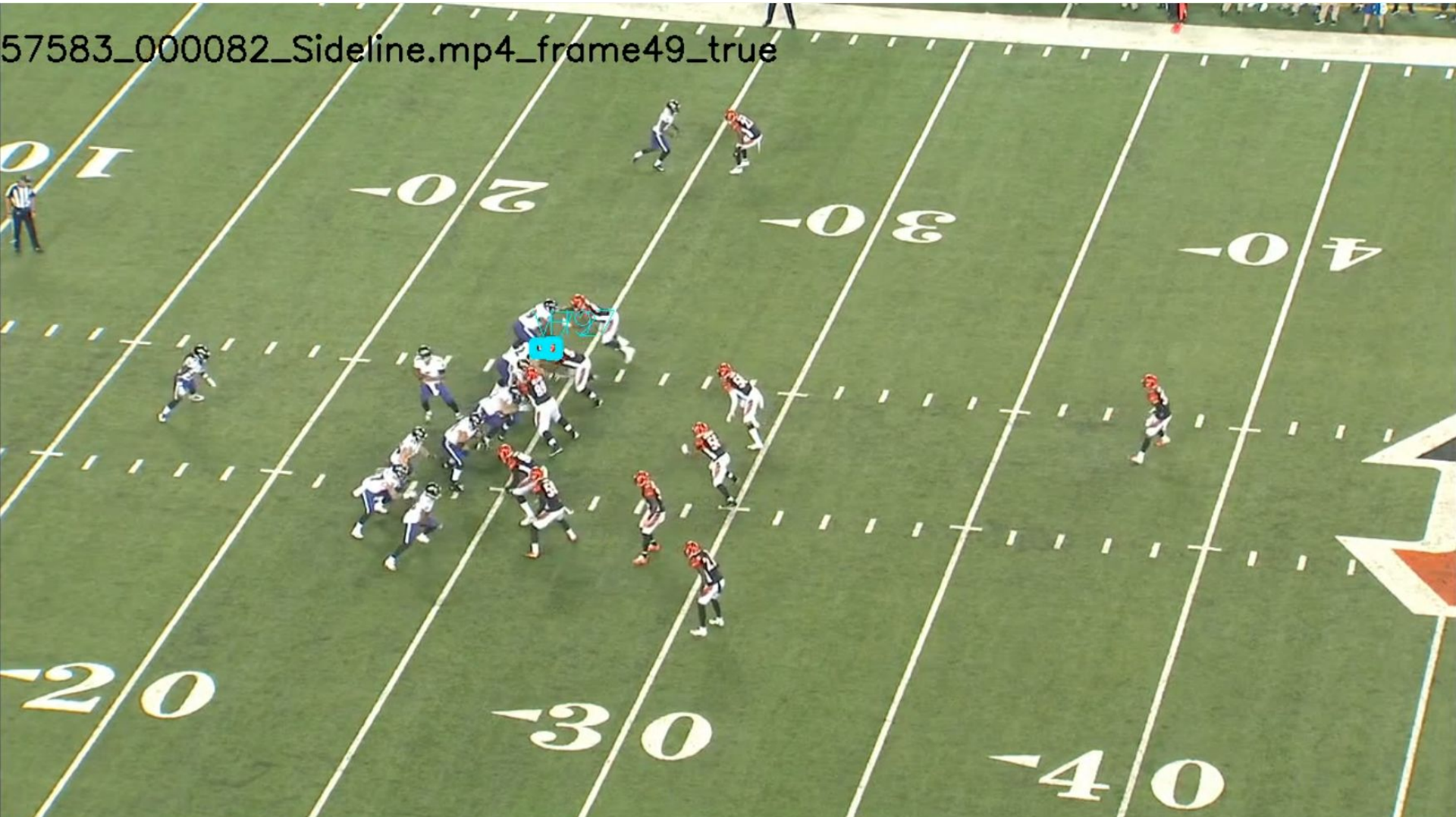
<https://www.kaggle.com/samhuddleston/nfl-1st-and-future-getting-started>



1.コンペ概要

カメラが二つありEndzoneとSidelineから同時に撮影

予測はそれぞれの画に対して行う

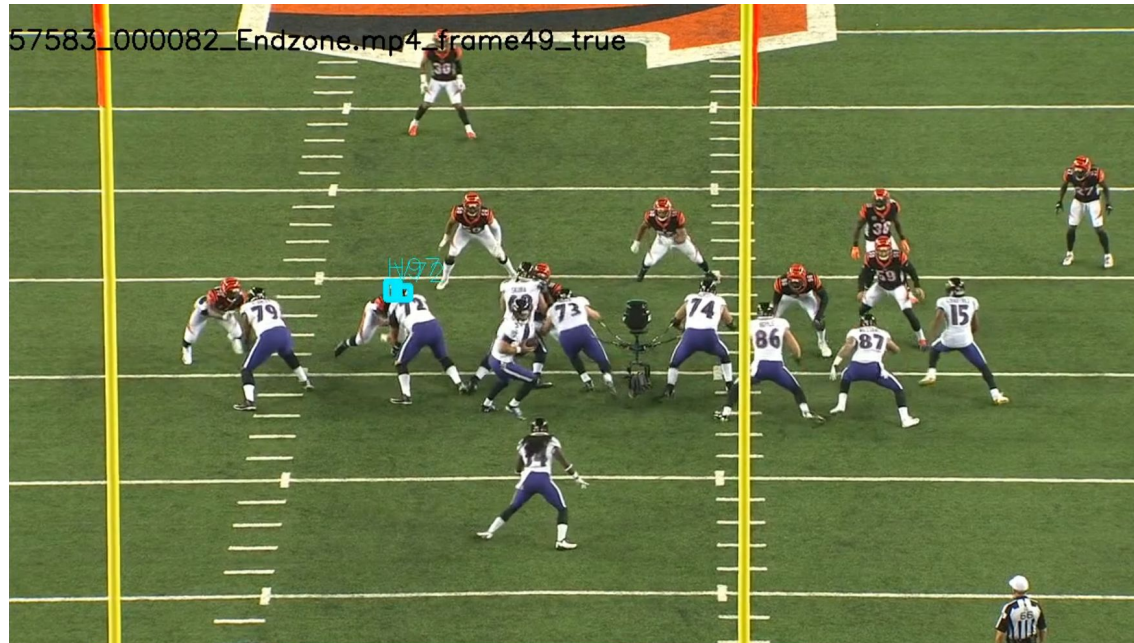


1.コンペ概要

その他入力データ

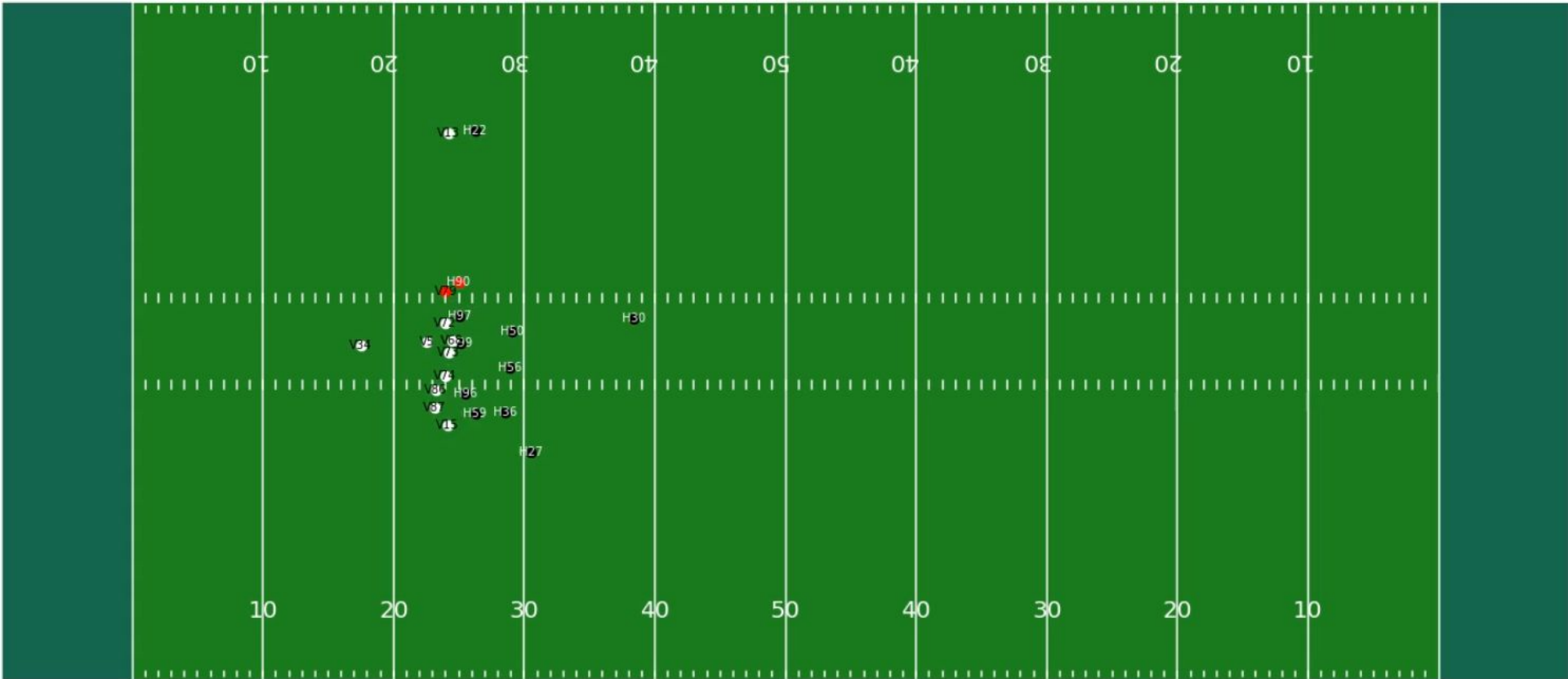
動画の各フレームに対して

- 全プレイヤーのヘルメットのBBOX
- 全プレイヤーのID
- チーム(ホームかアウェイか)
- ヘルメットが衝突しているかどうか



1.コンペ概要

そのほかGPSセンサーのトラッキング情報も(下図はそのプロット)



* カメラとの位置関係の結びつけが難しく、上手く使えていたチームは結果的になかった

1.コンペ概要

評価指標

F1 score@IOU>=0.35

ただしframeが+-4までずれていてもOK

時系列要素が絡むので結構特殊

コードが公開されておらず評価指標の謎を解くところから始まる

The IoU of a proposed bounding box and a ground truth bounding box is calculated as:

$$IoU(A, B) = \frac{A \cap B}{A \cup B}.$$

The metric treats any IoU of at least 0.35 as a true positive.

F1 is calculated as follows:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

where:

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

2.結果

15/459位

<div><div></div>In the money<div></div>Gold<div></div>Silver<div></div>Bronze</div>							
#	△pub	Team Name	Notebook	Team Members	Score	Entries	Last
1	—	Dmytro Poplavskiy			0.7527	6	3mo
2	▲1	Happy New Impact			0.7079	158	3mo
3	▼1	Nvidia Football League			0.6981	86	3mo
4	—	Azat Davletshin			0.6355	7	3mo
5	▲3	Ruslan Grimov			0.6095	64	3mo
6	—	fantastic_hirarin			0.5910	59	3mo
7	▼2	AlphaGoal			0.5753	73	3mo
8	▲5	Guanshuo Xu			0.5217	5	3mo
9	▲1	Deep Impact	<> RCNN Full Train + ...		0.5153	176	3mo
10	▲1	nvnn			0.5137	70	3mo
11	▼2	bestfitting			0.5126	46	3mo
12	▲5	-			0.5038	38	3mo
13	▼6	GMになるのだ🐶			0.5017	186	3mo
14	▼2	kzykmyzw			0.4766	12	3mo
15	—	camaro			0.4556	98	3mo
16	▲3	yu4u			0.4541	14	3mo
17	▼3	The A-Team			0.4359	131	3mo
18	▼2	French Flair			0.4244	169	3mo
19	▲7	PaulG			0.4017	14	3mo
20	▲2	No Free Lunch			0.3693	39	3mo

2.結果

LBの推移が手に入れば・・・

3.私の解法

Baseline: Multiple Object Tracking + 2 additional branch

- Helmet impact classification
- Team feature

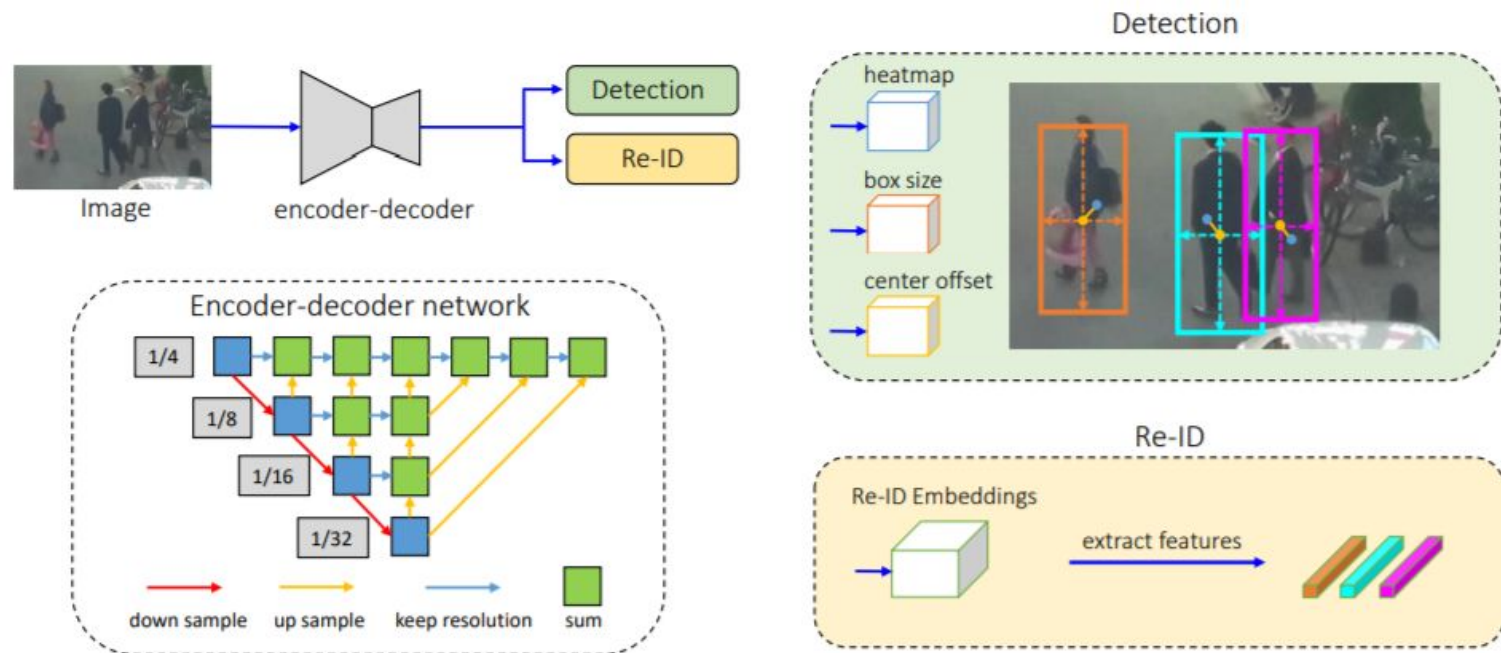


Fig. 2. Overview of our one-shot tracker *FairMOT*. The input image is first fed to an encoder-decoder network to extract high resolution feature maps (stride=4). Then we add two homogeneous branches for detecting objects and extracting re-ID features, respectively. The features at the predicted object centers are used for tracking.

3.私の解法

選手ごとの予測

1. Helmetを全てtrackingする
2. 選手ごとにヘルメットが衝突したかどうかの確率を時系列で出す
3. 前後nフレームの極大値だけを採用、さらに閾値で切る

後処理(とにかくFalse positiveを消すのが重要)

1. 敵味方のBBOXがあるIOU以上なら、高いほうの確率を採用する
2. 最も近い敵のプレイヤーがある閾値以上遠ければ不採用
3. 前後mフレーム含めて一つしかない予測は消す

閾値だらけ！

→[optuna](#)でブラックボックス最適化

4.敗因

- 1.左右のviewのアソシエーションにハマる
- 2.アンサンブルできず
- 3.時系列要素を組み込めなかった

4.敗因

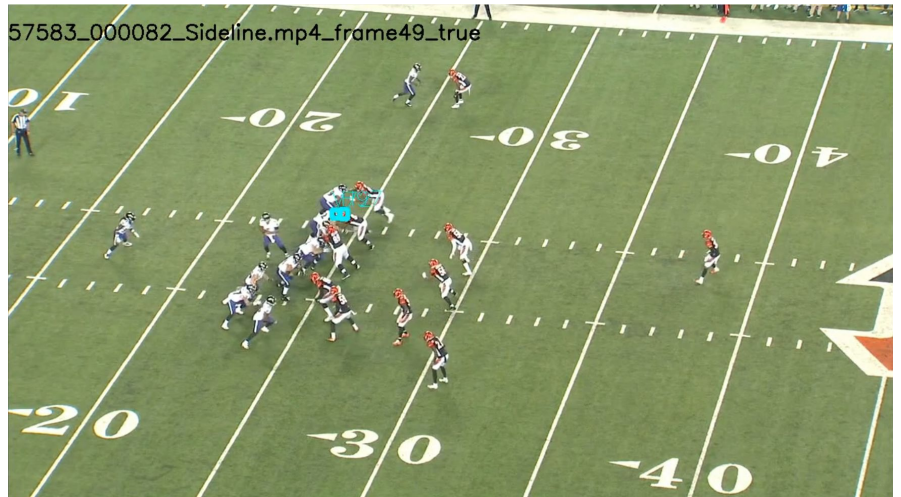
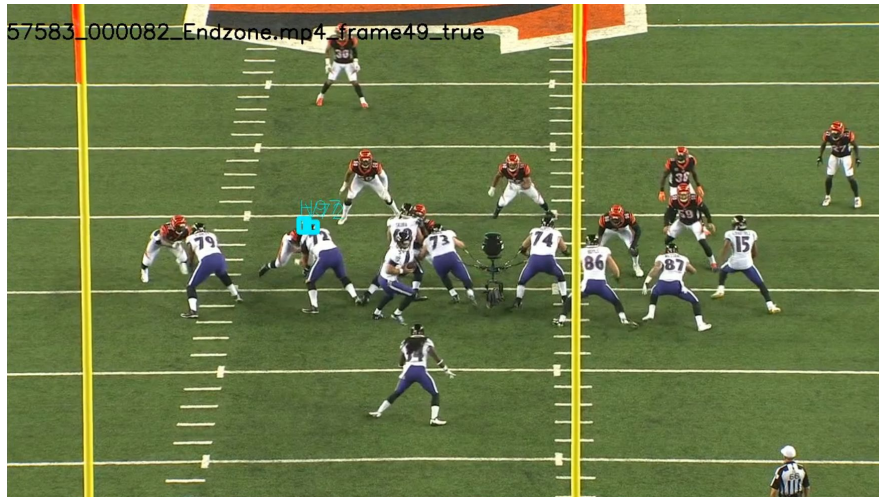
1.左右のviewのアソシエーションにハマる

ある視点の情報はもう一方の視点での予測に有用なはず

(例えば左図では重なっていても右図で見れば実際は空間が開いている、など)

幾何的な制約を考慮すれば1対1でプレイヤーを対応させられるのでは？

→結局スコアの向上につながるほどまともにならなかった。



4.敗因

1.左右のviewのアソシエーションにハマる

2.アンサンブルできず

→パイプラインが複雑で後処理依存が強く、アンサンブルできるところがなかった

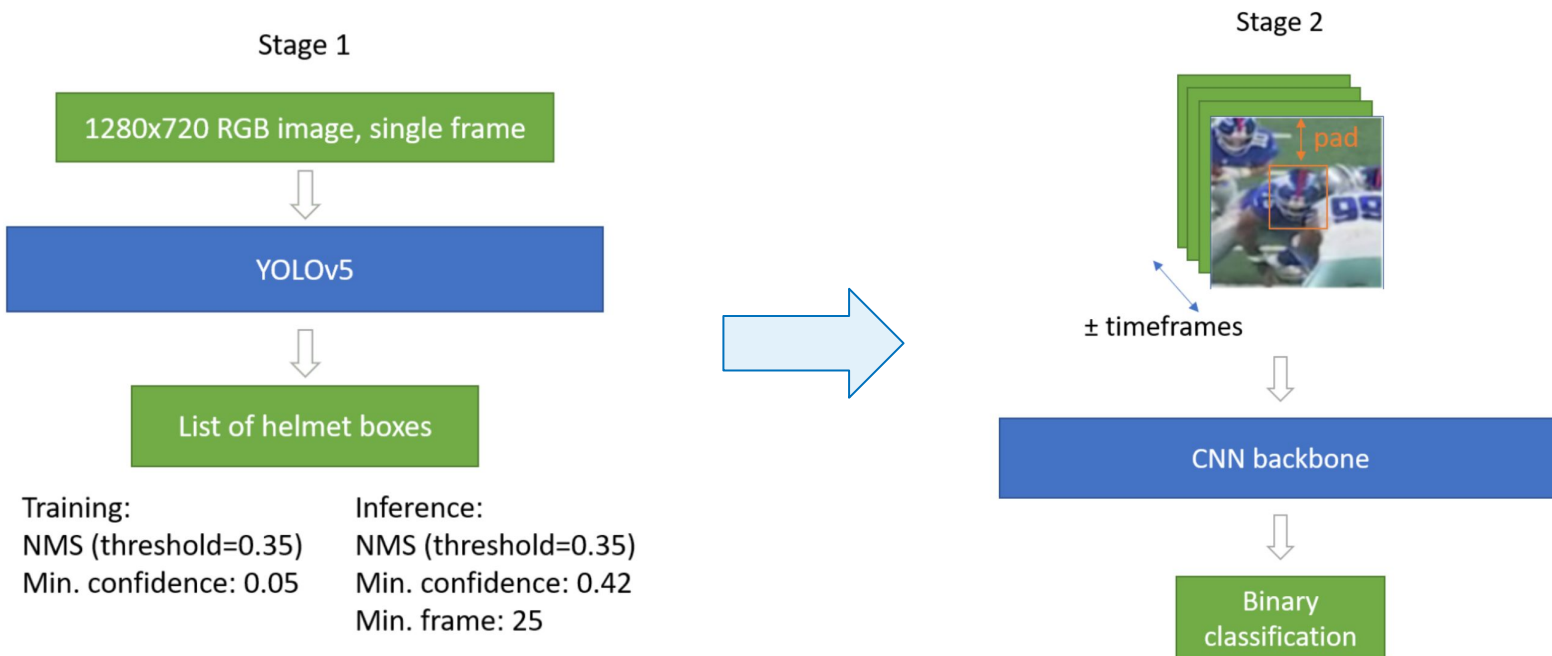
3.時系列要素を組み込めなかった

→他チームの解法で

5.上位の解法

大体みんな2stageでの予測

1. HelmetのObject Detection
2. 切り出した前後数フレーム合わせた3D image classification



2位の解法より

<https://www.kaggle.com/c/nfl-impact-detection/discussion/208979>

6.おまけ

おすすめ現行コンペ

1.Bristol-Myers Squibb – Molecular Translation

OCRコンペ！

化学物質の構造式をInChI(International Chemical Identifier)に変換

<https://www.kaggle.com/c/bms-molecular-translation/overview>

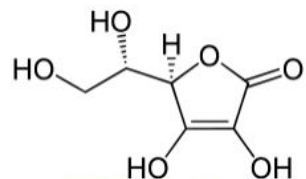
例：

CH₃CH₂OH

エタノール

InChI=1/C2H6O/c1-2-3/h3H,2H2,1H3

InChI=1S/C2H6O/c1-2-3/h3H,2H2,1H3 (standard InChI)



L-アスコルビン酸

InChI=1/C6H8O6/c7-1-2(8)5-3(9)4(10)6(11)12-5/h2,5,7-10H,1H2/t2-,5+/m0/s1

InChI=1S/C6H8O6/c7-1-2(8)5-3(9)4(10)6(11)12-5/h2,5,7-8,10-11H,1H2/t2-,5+/m0/s1 (standard InChI)

<https://ja.wikipedia.org/wiki/InChI>

6.おまけ

おすすめ現行コンペ

2.Shopee - Price Match Guarantee

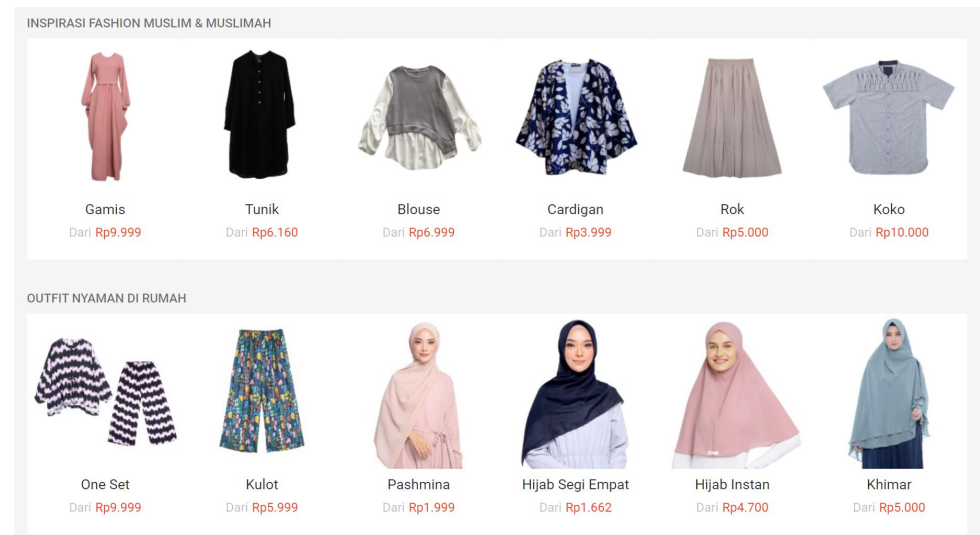
商品の画像とタイトルのテキストから同じ商品をグループ分け
数万枚の画像があって同じグループは最大でも50個しかない

<https://www.kaggle.com/c/shopee-product-matching>

例:

インドネシアのEコマース

なのか、こんな感じのが多い



<https://shopee.co.id/Fashion-Muslim-cat.43>