

サッカーにおける AI 技術を統合したファウルの自動判定

2020SE016 岩佐拓斗

指導教員：野呂昌満

1 はじめに

近年のサッカーの判定においては、VAR[1] の導入により試合中の誤審は格段に減少したが、未だ試合結果を左右する誤審は多数存在する。VAR は、試合を長時間中断することによる様々な悪影響も指摘されており、試合への介入は最低限であることが望まれる。主審は、これらを考慮しながら、VAR 介入の是非を 1 人で判断しなければならない。

本研究の目的は、ファウルの高精度な自動検出により、ファウル判定の正確性に資する要因を考察することである。第 1 段階として、ファウル行為の最も基礎的な例の 1 つであるハンドリングの自動検出を試みる。

実験科学の方法論に則り問題解決を図るという観点から、以下の通り研究課題を定義する。

- **RQ1.** 画像から取得する座標精度の向上
- **RQ2.** ハンドリング検出モデルの定義
- **RQ3.** ハンドリング検出モデルの実用化に関する考察

2 先行研究

Jingchao ら [2] は、VAR により試合が長時間中断する仕組みが、選手の熱意と審判の権威に大きな影響を与えることを指摘し、ハンドリングのリアルタイム検出の実現を目的に研究を行った。

Jingchao ら [2] は、物体検出ライブラリ YOLOv5 と姿勢推定ライブラリ AlphaPose を用いたハンドリング検出アルゴリズムを提案した。ファウル判定に必要なキー情報

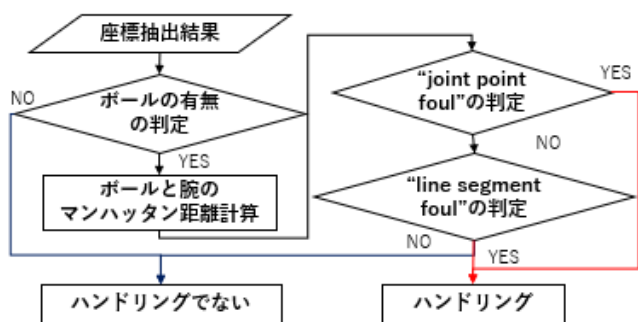


図 1 ハンドリング判定アルゴリズム ([2] より参照)

として、ボールの中心と半径、選手の両腕の関節点 (肩, 肘, 手首) を定義し、その座標を抽出する。ボールのキー情報は、YOLOv5 のボール検出ボックスのパラメータ (中心 x 座標, 中心 y 座標, 幅, 高さ, 確信度) を用いて求める。この際、ボール検出ボックスの精度を高めるために、YOLOv5 のモデルの転移学習を事前に行う。選手のキー情報は、AlphaPose の骨格抽出により座標を取得する。判定モデル

は、ボールの有無の判定、ボールと距離が近い最大 6 選手の特選、ハンドリング判定の 3 つに分けられる (図 1)。ハンドリング判定は、2 次元空間上で点、線、円が交差するかどうかの問題に変換する。"joint point foul" の判定 (図 2) では、各関節点 (m,n) とボールの中心 (x,y) との長さ、ボールの半径 r_0 を比較し、"line segment foul" の判定 (図 3) では、各関節点 (m,n) を結ぶ線分に対してボールの中心 (x,y) から引かれる垂線 d の長さ、ボールの半径 r_0 を比較し、2 段階で判定を行う。実験結果では、ハンドリング検

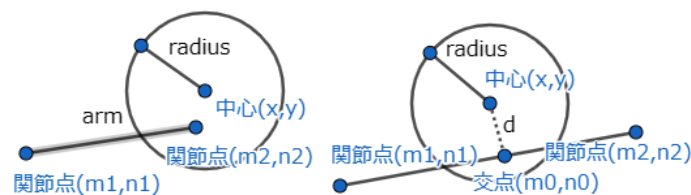


図 2 "joint point foul" の判定 ([2] より参照)

出の再現率は 0.840、適合率は 0.913、正解率は 0.996 であり、ビデオの判定速度は 8.2frame/s である。

3 課題解決のアプローチ

本研究では、既存研究の検出モデルを基に、ボール座標抽出精度の向上、ボール検出ボックスの修正、手首より先のハンドリング判定の 3 つを行い、より実用的なハンドリング検出モデルの提案を行う。座標精度の向上には、物体検出の分野で一定の検出精度を示し、最も速度に優れた性能を示す YOLO の学習済みモデルの転移学習を行う。検出ボックスの修正では、ボールと選手の重なりにより誤ったボール検出ボックスが生成されることを想定し、検出ボックス座標を修正する。ハンドリング判定では、既存研究における両腕の 6 つの関節点に加え、手首より先の関節点に対する判定を行う (図 4)。判定箇所は両腕と両手の 4 つに場合分けし、ボールと最も距離が近い部分を、最もハンドリングの可能性が高い部分とみなし、その部分のみ判定を行う。腕の判定においては、既存研究と同様に "joint point foul", "line segment foul" の 2 段階で行い、手の判定では、"joint point foul" のみを行う。この理由は、手の関節同士の間隔は腕に比べて非常に短く、"line segment foul" の有無は、検出精度に大きく影響を与えないと考えられるからだ。

4 実装

転移学習で用いるデータセットは、Kaggle 上で公開されている "Football Vs Rugby Image Classification" にお

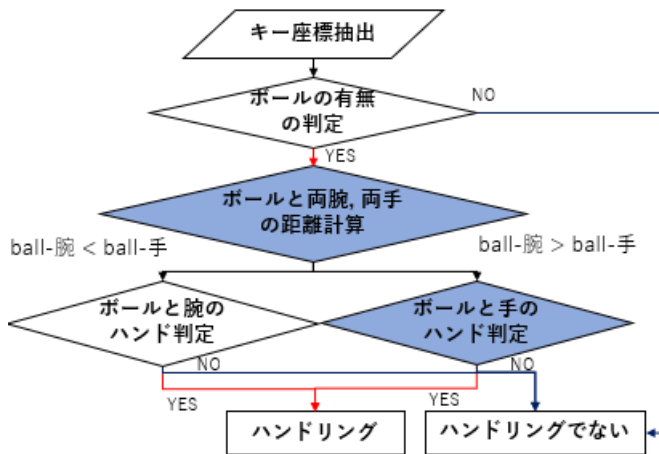


図4 ハンドリング判定の概要図

ける 1224 枚の football 画像を用いる。ボールに対する検出ボックスをアノテーションとして付与し、訓練用 70 %、評価用 20 %、テスト用 10 % にデータセットを分割する。YOLOv5 の学習済みモデルは、最も高速な YOLOv5s を採用する。

検出ボックスの修正では、ボールと選手の検出ボックス座標から、ボールの隠れた部分を特定し修正する。ハンドリング検出では、テスト用データセットとして iPhone12 mini で撮影した 3198 枚のサッカー画像を用意する。ボールのキー座標抽出には、転移学習済みの YOLOv5s モデルによるボール検出ボックス座標を用い、選手のキー座標抽出は、AlphaPose[3] で紹介された、COCO Wholebody dataset で学習されたモデルにより、腕と手を含む全身 133 個の関節点座標を用いる。比較対象である既存研究の手法の検証には、MSCOCO dataset で学習されたモデルを用い、手を含まない全身 17 個の関節点座標を用いる。本研究では、ハンドリング判定の数理モデルの精度を正確に評価するために、ボールを含む画像のうち YOLOv5 でボール検出に失敗した 92 枚は、データセットから除く。

5 考察

5.1 実験結果の考察と今後の課題

転移学習では、MAP0.5 において学習前では 0.563、学習後では 0.617 とボール検出精度が向上することが確認された。ハンドリング検出実験では、既存研究の手法と比べて、検出精度が大きく向上した (表 1,2,3)。実験結果からは、

表 1 腕のみの検出結果

正解ラベル	検出結果	
	foul	no-foul
foul	TP758	FN571
no-foul	FP83	TN1694

表 2 腕と手の検出結果

正解ラベル	検出結果	
	foul	no-foul
foul	TP1169	FN160
no-foul	FP107	TN1670

適合率の値がどちらも約 9 割あることから、ハンドリングと予測した時の検出精度は非常に高いことがわかり、再現率の値に顕著な差があることから、手首より先のハンドリ

表 3 検出精度の比較

	腕のみ	腕と手の両方
再現率	0.570	0.880
適合率	0.901	0.916
正解率	0.789	0.914

ング検出が行えるか否かが、ハンドリング判定において重要な要素であることが確認された。

今後の課題は、3次元空間を考慮したファウル検出と検出速度の向上の 2 つを挙げる。前者の理由は、誤検出画像の傾向から、選手とボールや選手同士の重なりにより隠れた部分が存在する画像に対して誤検出が多いからである。

5.2 他のファウル検出への応用可能性についての考察

サッカーにおけるファウル行為は、主に他者やボールとの接触により引き起こされる。ハンドリングファウルはボールとの接触により発生する、ファウルの最も基礎的な例である。他者との接触により発生するファウルの中には、ボールとの接触の有無が重要な要素として考慮されるものもあり、ボールとの接触検知は、他のファウル検出にも直結する要素である。本研究では、ファウル判定における基本技術である接触の検知において、ボールと選手との接触を高精度で検出することに成功した。これらのことから本稿の提案手法は、他者との接触により発生する他のファウル判定にも、応用可能であると考えられる。

6 おわりに

本稿は、サッカーにおけるファウルの高精度な自動検出により、判定の正確性に資する要因を考察することを目的とし、YOLOv5 と AlphaPose を用いたハンドリング自動検出手法を提案した。実験では、再現率 0.880、適合率 0.916、正解率 0.914 の高い精度が示された。ファウル判定の高精度な自動検出の実現は、VAR のデメリットを最小限に抑えながら、判定の正確性を向上させることに繋がる。

参考文献

- [1] JFA—公益財団法人日本サッカー協会. “ルールを知ろう!” JFA.jp, <https://www.jfa.jp/rule/var.html>, 2020. (Accessed 2023.1.9)
- [2] Jingchao Xu, et al. “Real-time detection of game handball foul based on target detection and skeleton extraction,” *2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, 2021, pp.41-46.
- [3] Hao-Shu Fang, et al. “AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, pp.7157-7173.