

フィールドスポーツにおける 選手と審判以外の背景除去に関する研究

姜文淵† 山本雄平‡ 田中ちひろ† 坂本一磨†‡ 中村健二‡‡

田中成典‡‡‡ 鳴尾丈司† 肖智葳‡‡‡ 岡寄雄也‡‡‡ 福井真子‡‡‡

関西大学先端科学技術推進機構† 関西大学環境都市工学部‡ 公立小松大学生産システム科学部†‡

大阪経済大学情報社会学部‡‡ 関西大学総合情報学部‡‡‡ 関西大学大学院総合情報学研究科‡‡‡

1. はじめに

我が国では、総務省のスポーツ×ICT ワーキンググループ[1]の設置や内閣府の未来投資戦略2018[2]が発表されるなど、スポーツ選手の競技力向上に向けた ICT の効果的な利活用が推進されている。このような背景の下、フィールドスポーツにおいて、画像処理による選手領域の検出及び選手追跡に関する研究[3]が行われている。選手領域の検出には、一般的に背景差分法や深層学習による選手検出手法が用いられる。しかし、背景差分法では、日照環境の変動や風の影響による背景の変化から正確な選手領域の検出ができない。また、深層学習による選手検出手法では、背景に含まれる人以外の物体を選手として誤認識する課題がある。そこで、本研究では、GAN (Generative Adversarial Network) を利用した画像生成モデルの一種である pix2pix[4]を用いて、選手と審判以外の背景を除去するシステムを開発する。これを実現することにより、日照環境やコート外の影響を受けずに選手領域の検出が可能になり、高精度な選手追跡が期待できる。

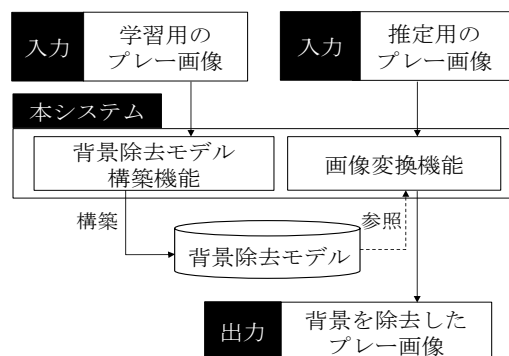


図1 本システムの概要



図2 学習データのペア画像

2. 研究の概要

本システムの概要を図1に示す。本システムは、背景除去モデル構築機能と画像変換機能により構成される。入力データは、学習用のプレー画像と推定用のプレー画像とする。それぞれのプレー画像は同一試合から抽出し、学習用に前半部分、推定用に後半部分を使用する。出力データは、背景を除去したプレー画像とする。

2.1 背景除去モデル構築機能

本機能では、敵対するペアの画像から画像間の特徴関係を学習することで、対となる画像を生成する pix2pix を用いて背景を除去するモデルを構築する。モデルを構築するための学習データ (図2) は、背景を含む画像と、手動で選手と審判以外の背景を除去した画像の敵対するペアの画像を使用する。

Research for Removing Background Excluding Players and Referees in Field Sports

† Wen Yuan Jiang, Chihiro Tanaka and Naruo Takeshi
Organization for Research and Development of Innovative Science and Technology, Kansai University

‡ Yuhei Yamamoto
Faculty of Environmental and Urban Engineering,
Kansai University

†‡ Kazuma Sakamoto
Faculty of Production Systems Engineering and Sciences,
Komatsu University

‡‡ Kenji Nakamura
Faculty of Information Technology and Social Sciences,
Osaka University of Economics

‡‡‡ Shigenori Tanaka and Mako Fukui
Faculty of Informatics, Kansai University

‡‡‡‡ Zhiwei Xiao and Yuya Okazaki
Graduate School of Informatics, Kansai University

2. 2 画像変換機能

本機能では、推定用のプレー画像を背景除去モデルにより画像変換することで、背景を除去したプレー画像を出力する。まず、入力データからフィールド内の人物のみを抽出するために、フィールド上の4つのコーナーを基準点とした処理領域を切り出す。次に、抽出した領域内を `pix2pix` の学習画像サイズに分割し、背景除去モデルを用いて背景を除去した画像に変換する。そして、変換した画像をそれぞれ分割前の座標位置に戻すことで、選手と審判以外の背景を除去したフィールド全体のプレー画像を出力する。

3. 実証実験

3. 1 実験内容

本実験では、サッカーの試合映像を対象にし、選手検出精度を比較することで、提案手法の有用性を検証する。本実験では、YOLOv4[5]を用いて、背景除去前後の画像のそれぞれで選手検出モデルを構築する。また、実験に使用する検証データは、フレームレート 30fps で撮影したプレー映像から 300 フレーム間隔に抽出した 100 フレームを用いる。

3. 2 結果と考察

実験結果を表 1 に示す。背景除去後において適合率と再現率と、F 値の全てにおいて僅かな精度向上が見られた。提案手法により精度が向上した成功例 (図 3) と精度が低下または不変であった失敗例 (図 4) を示す。図 3 では、ゴールポストやフィールド上のライン、コート外の物体から影響を受けた誤検出が改善していることがわかる。このことから、提案手法は、背景に含まれる物体を選手として誤検出する課題に対し、有用であることが確認できた。一方で、図 4 では、2人の選手間を誤検出する課題に対し、改善は見られなかった。これは、YOLOv4 が近接する複数の物体を 1 つの物体として検出する傾向があることが原因と考えられる。また、背景除去時に選手が消えることにより精度が低下した。これは、カメラから離れた位置の選手が小さく、変換時に背景と区別ができないことが原因である。そのため、学習データにカメラから離れた位置のプレー画像を追加することで、提案手法の背景除去の精度が向上するかを確認する必要がある。

4. おわりに

本研究では、選手と審判以外の背景を除去する手法を開発した。そして、実証実験により、背景を除去することで誤検出を改善できることがわかった。今後は、本研究で明らかになった課題を解決するとともに、多様なプレー画像において適用可能かを検証する。

表 1 実験結果

背景	適合率	再現率	F 値
除去前	0.86	0.70	0.77
除去後	0.88	0.72	0.79

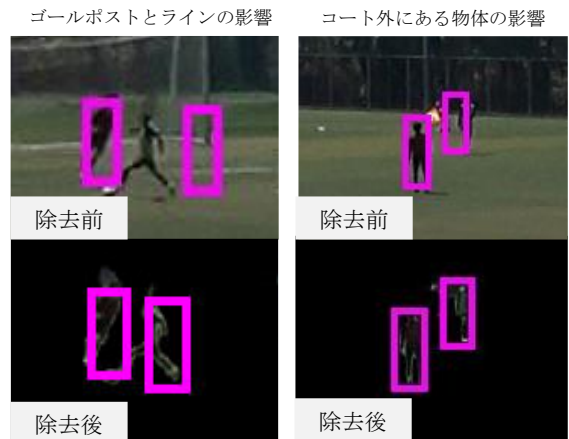


図 3 成功例

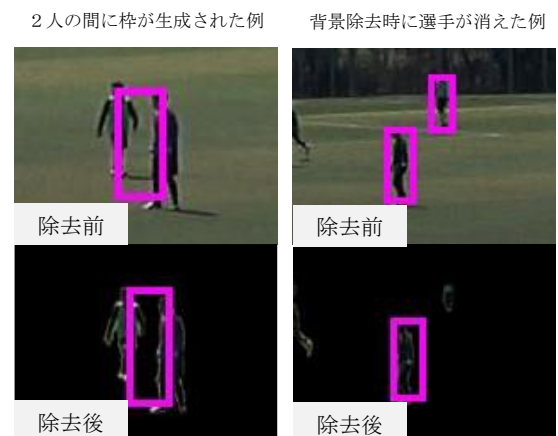


図 4 失敗例

参考文献

- [1] 総務省：スポーツ×ICT WG の立上げについて，総務省（オンライン），入手先〈https://www.soumu.go.jp/main_content/000369372.pdf〉（参照 2022-1-7）。
- [2] 内閣府：未来投資戦略 2018，内閣府（オンライン），入手先〈https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/miraitousi2018_zentai.pdf〉（参照 2022-1-7）。
- [3] 田中成典，山本雄平，姜文淵，中村健二，清尾直輝，田中ちひろ：複数視点からの映像を用いたスポーツ選手のトラッキングに関する研究，知能と情報，日本知能情報ファジィ学会，Vol.32，No.4，pp.821-830，（2020）。
- [4] Isora, P., Zhu, J., Zhou, T. and Efors, A.: Image-to-Image Translation with Conditional Adversarial Networks, *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.5967-5976 (2017).
- [5] Bochkovskiy, A., Wang, Y. and Liao, M.: YOLOv4: Optimal Speed and Accuracy of Object Detection, arXiv (online), available from 〈<https://arxiv.org/pdf/2004.10934.pdf>〉 (accessed 2022-1-7).