

バスケットボールファウルデータセットとモデル認識 — (Basketball Foul Detection) —

劉靖[†] 北濱幹士[‡] 山田光穂[†] 星野祐子[†]

[†] [‡] 東海大学大学院情報通信学研究科 〒108-8619 東京都港区高輪 2-3-23

E-mail: [†] {0cjinm025, myamata, hoshino}@tokai.ac.jp, [‡] kitahama@tokai.ac.jp

あらまし About sports Video judgment is often used in ball games and athletics for arbitration. However, bad decisions are a frequent problem in sports, and the arbitration results tend to differ based on the subjectivity of the human referees; hence, a system that ensures objective judgment is required. In this paper, we propose a solution to judge fouls in basketball automatically.

キーワード Motion recognition, Deep learning, Basketball analysis, Sports science

1. はじめに

近年、コンピュータ人工知能技術の急速な発展に伴い、画像認識の能力と精度が大幅に向上した。人工知能もスポーツイベントに使われ始め、好成績を収めている。スポーツ分野では、動作分析によるフォームの修正、スキルアップの支援、対戦相手の戦術分析、選手のコンディション管理など広く応用されている。たとえば、球技、陸上、競泳など、スポーツ競技の各種目でも AI が浸透しつつある。今や球技の世界（野球、サッカー、バスケットボールなど）では戦術の高度化により、競技映像の分析サービスが一大ビジネスとなっており、記録を競う陸上や競泳などの世界では、競技者の体の動きを細かく分析し、適切にコーチや選手にその情報をフィードバックするシステムが活用されつつある。他に放送やストリーミングメディアの強化でも活用されはじめている。たとえば、2018 年のワールドカップは良い例である。世界中の何十億ものファンが画面の前に集まり、各試合を観戦した。FOX Sports は、ゲームをリアルタイムでブロードキャストするだけでなく、FIFA ワールドカップハイライトマシンプロジェクトを時間内に開始した。これは、ファンが数秒以内に独自のカスタマイズされたハイライトを作成して共有できるようにする AI テクノロジーに基づくプラットフォームである。

競技の場において、しばしば誤審が問題となっている。スポーツ競技への関心の高まりにより、誤審や判定の公平性をめぐって大きな議論が巻き起こることも珍しくない。正確かつ公平な判定を支援するために AI 活用、あるいは判定自体の自動化も議論されている。陸上競技を始め、球技でもビデオ判定が利用されている。テクノロジーの発展により、テニスなどの球技ではビデオ判定のみならず、ホークアイシステムが使われている[1]。競技技術の高度化に伴い、審判員に要求される技能レベルも上がっている。また、審判の主観によって判定結果に差が出やすく、より客観的な判断

が必要となる。スポーツの審判では、体操競技で AI による採点システムの開発が始まっており、最近、サッカーの試合でもビデオ判定やホークアイシステムが導入された。しかし、試合が時々中断されて、観戦感覚が悪くなってしまう場合がある。試合の中断が多くなると特にストリーミングメディア等で観戦しているファンにとってライブ感が失われ、観戦感覚が悪くなることが考えられる。

近年、日本ではバスケットボールのプロ化が実現し、人気のスポーツとなってきているが、審判の判断を巡って争うこともある。また、日常スポーツや趣味としてバスケットボールをする人も多いが、そのようなときには審判がいない場合もある。そこで、本研究では公正なゲームプレイを確保し、審判による誤判断を減らし、ライブ感を減退させないことを目的として、バスケットボールのファウル判断モデルと、AI と画像認識を使用した対応するシステムの提案と試作を行う。提案された手法は、正確な判断を迅速に実行することにより、レフリーを支援することが期待されている。

2. 関連研究

2.1 動作検出と人物動作解析

ボール追跡手法の 1 つに安田らによるカメラの 3 次元位置と姿勢を用いたボール認識の手法がある。この手法は、オクルージョンに頑強なボール認識である。この手法ではカメラの位置と姿勢推定を行った後、ガウスモデルを用いた前景抽出、色空間での検出を用いたボール候補の検出、エピソード拘束を用いた多視点映像の統合の順で処理を行う。この処理により検出されたボール位置を映像の前後のフレームを用いて補完し、3 次元位置を推定し、出力する[2]。この手法によって、ボールが隠れてしまうオクルージョンに対して、複数視点映像を用いることで頑強なボールの検出方法について有効な手段あることを示し、正しく検出できることを確認した。

他のボール追跡手法として、高橋らによる機械学習

を利用した複数視点映像からのボール追跡がある[3]. この手法では機械学習を用いた、複数の映像から追跡したボール位置を実空間上に投影し、各カメラ映像での処理結果を統合することで、高い精度のボール位置推定を実現した.

2.2 動作検出と人物動作解析

動作検出手法の1つに高橋らによる手の動きに基づいた不審動作検出がある[4]. この手法は、スーパーマーケット内でカゴをもって買い物をしている際に、人物が商品を保持したあとの手の動きの特徴量を用いた不審動作検出手法である. CHLAC 特徴量を算出し、主成分分析によって通常動作部分空間を作成することで異常度を数値化し不審動作の判定基準とした検証実験である. これは、通常動作を商品の追跡および人物の検出と追跡を行い、動作特徴量算出を行う. この動作特徴量を用いて判定基準作成を行っている.

他に田摩らによる身体的特徴変化を統計的に解析したモデルを用いた人物追跡技術の手法がある[5]. この手法には統計的モデリングを用いた人物追跡と人物姿勢推定が含まれている. 人物追跡手法では, Active Ω Model を用いて, 頭部外輪郭形状の推定を行う. 人物追跡では, 部分的な遮蔽や姿勢変動に対応するために, 対象を 5 つの基幹部位領域(頭部, 胴体部, 腰部, 上脚部, 下脚部)に分割して追跡する. 姿勢推定において基幹リンクモデルの作成に用いた姿勢データには, 主に, 「しゃがむ」「前傾姿勢」の姿勢変動が含まれている. 3次元姿勢推定では, まず腕エリア特徴量抽出を行い, 次にモデル照合を行う. その後動作解析を行う. 人物動作解析において, 高速化や使用メモリ削減に有効な人物検出, 人物の大まかな姿勢や肩位置を抽出しながら姿勢変動や遮蔽に頑健な人物追跡, 行動理解に重要な腕部の姿勢推定など方法を提案している.

2.3 スポーツモーション認識

スポーツビデオ内で発生するイベントを認識するための手法は主に視聴覚コンテンツ分析に基づいて行われている. 視聴覚コンテンツ分析を使用した手法の多くはブロードキャスト映像で 사용되는ことが多く, すべてのイベントは手作業で抽出される.

先行研究としては, Alexandre Alahi らによる固定カメラのビデオ入力からオブジェクト(プレイヤー, 審判)を追跡する手法があるがイベントの認識がうまく行っていない[6]. 他に Jungsoo Lee らによる, バスケットボールの試合でプレイヤーまたはジャッジの動作に焦点を当て, イベントを自動的に抽出する手法がある. これらのイベントは, プレイヤーとレフリーの事前定義された動きを追跡および分析することによって抽出される[7].

2.4 スポーツデータセット

スポーツデータセットとして, 野球のキャッチやスロー, サッカーのジャグリングなどのスポーツ関連アクションが含まれた UCF101[8], ACTIVITYNET [9]などの多くのビデオベースのベンチマーク人間活動認識データセットが公開されている. これらのデータセットには, ビデオと, ビデオ内の被験者によって実行されたさまざまなアクティビティの画像が含まれており, ビデオの分類に大きく貢献している. UCF Sports [10], BASKETBALL-51 [11]のデータセットには, 野球, バスケットボール, テニスなどのさまざまなスポーツの画像とビデオが含まれており, バドミントン, サッカー, 乗馬, ランニングなど, さまざまなスポーツを分類するために使用されている.

しかし, 特定のスポーツに関連するすべてのアクションが含まれていないため, 実際のアプリケーションに適用して, 単一のスポーツで実行されるさまざまなアクションを分類することは困難である. 例として, データセット UCF101 と BASKETBALL-51 には, バスケットボールダンクやバスケットボール射撃などのダンクアクションに関連するアクティビティのみがある.

3. 研究方法

3.1 ファイル

バスケットボールにおける違反には, 大きく分けて「バイオレーション」と「ファウル」の2種類がある. バイオレーションにはドリブル, トラベリング, 3秒ルール, 8秒ルール, 24秒ルールなどが含まれる. ファウルには, 大きく2種類あり, ひとつは, プレイヤー同士の体の接触によって起こる「パーソナルファウル」, もうひとつは, スポーツマンらしくない言動や行為に対して与えられる「テクニカルファウル」である.

本研究のシステムは「パーソナルファウル」のうち以下の2種類のファウルを対象とする.

「ブロック」: ディフェンス側のプレイヤーによるファウルである. 相手プレイヤーがボールをコントロールしている・いないにかかわらず, ディフェンス側のプレイヤーが手や足を広げたり, 体ごとぶつかったりして相手の進行を妨げるとブロックと判断される.

「イリーガルユースオブハンズ」: 「規則に反している」という意味を持つ. 「イリーガルユースオブハンズ」とは不当な手の扱いを反則とするもので, 相手プレイヤーの動きを妨げることを目的に, 手を使って相手のプレイヤーをつかんだり叩いたりする動きを指す.

3.2 データ収集と学習データの作成

モデル作成用データの作成手順は以下の通り.

Step1: 試合データの収集

YouTube などから NBA バスケットボールの試合のライブビデオを収集する。プレーヤーのアクションに加えて、ビデオにはスコア、リプレイ、ハイライト、インタビューに関する情報も含まれている。ビデオ品質は主に 2kHD である。

Step2: ラベルの保存

ゲームの記録から、ビデオ内のファウルアクションのタイムスタンプを手動で検出し、タイムスタンプのリストを作成する。このリストには、ビデオの時間 (HH)、分 (MM)、および秒 (SS) で構成され、対応するファウルタイプがマークされている。

Step3: クリップの生成

得られたタイムスタンプにより、編集ソフトウェアをしようして約 5 秒のビデオクリップを生成する。このビデオクリップには、ファウルプレーヤーとファウルアクションが含まれている。20 本のバスケットボールゲームビデオからクリップを生成する。合計 554 個のクリップを生成した。収集されたデータセットのサンプルを図 1 に示す。



図 1 ルール違反のデータ

データセットには、上記 554 パターンのファウルビデオクリップが含まれている。これらのビデオは、スポーツ放送メディアで使用されるカメラによって撮影された完全に第三者の視点である。データは個人ファウルとショットファウルの 2 種ラベルに分類される。個人ファウルとショットファウルに分類された動画クリップの数量を図 2 に示す。

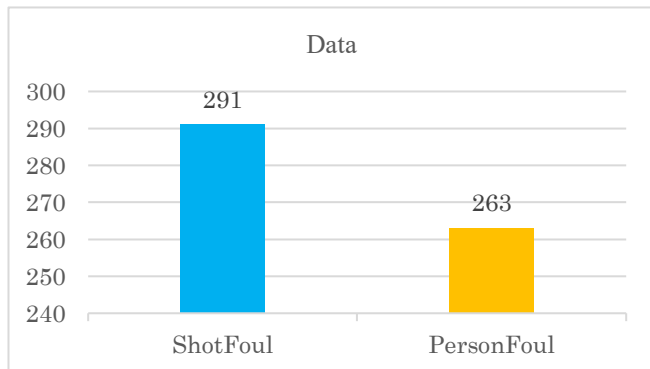


図 2 データセット分布

3.3 ファウルモデルの構築

TensorFlow とは、Google が開発したソフトウェアライブラリであり、機械学習アプリケーション用のオープンソースツールとしてリリースされている [12]。このライブラリを使用して、ニューラルネットワークモデルを構築する。

上記で定義した 2 つの判定対象のモデルを TensorFlow で 3DCNN と LSTM 用いて構築する。TensorFlow は深層学習に基づいてファウルモーションの分類器をトレーニングするために使用する。モデルのプロセスを図 3 と図 4 に示す。

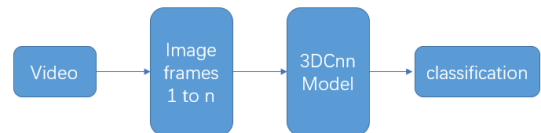


図 3 3DCNN モデル流れ

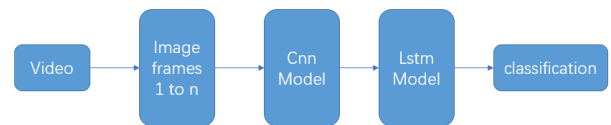


図 4 LSTM モデル流れ

4. 実験

この節では、今回作成したバスケットボールのデータセットを使用して評価実験を実施し、ファウルアクション認識の特性と課題を説明する。ビデオ分類には最も一般的に使用されているディープラーニングアーキテクチャの 1 つである、LSTM を使用した。このタスクの目的は、ファウルタグを含むアクションビデオクリップを正しく分類することである。ここでは、NBA ビデオから抽出されたクリップを使用して分類器をトレーニングし、分類器に基づいてアルゴリズムのパフォーマンスを評価する。

実験では、データセットはランダムに分割され、80% をトレーニングに使用し、20% をテストに使用した。トレーニングデータセットとテストデータセットの両方に、すべてのラベルが含まれている。

データセットの 2 種類のトピック関連分類の混同行列と分類レポートを図 5 と表 1 に示す。分類レポートには、個人ファウルとショットファウルが含まれている。ここでは、個人ファウルの再現率の値が最も低く、約 36% であり、ショットのファウルのカテゴリの再現率の値が最も高く、約 56% となっている。LSTM モデル全体の正解率 48% となっている。3DCNN モデル全体の正解率 55% となっている。分類精度が低いのは、データが不足している可能性がある。

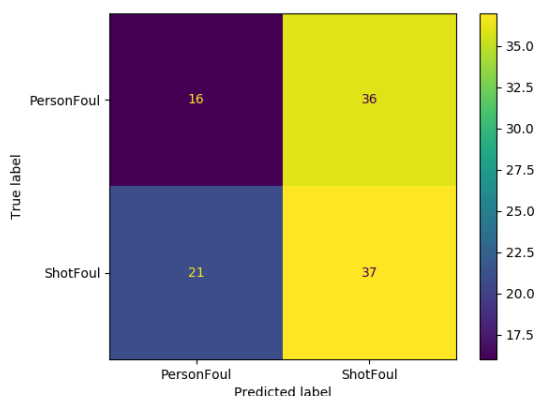


図 5 Confusion Matrix

LSTM	Precision	Recall	F1-Score	Support
PersonFoul	0.43	0.31	0.36	52
ShotFoul	0.51	0.64	0.56	58
Accuracy				
LSTM	0.48			
3DCNN	0.55			

Table 1. The classification report

5. おわりに

本研究では、TensorFlow を用いたファウルモデルを構築し、それに基づく判断モデルを検証した。しかし、モデルをトレーニングするためのデータ数が多くないため、モデル認識の精度に問題がある。この問題を解決するため、より多くのファウルデータを収集し、モデルを改善する必要があると考える。

本研究のシステムを審判の補助として使用することで、素早く正しい判定を伝えることができ、審判の負担を減少し、他の判定に集中することができると考える。また、全てのファウルやルール違反の適応に発展することで、人の審判がいなくても、AI テクノロジーによる公平公正且つ、クレームによる中断の少ない試合が視聴者に見せられると考えられる。

参考文献

- [1] 柏原全孝."判定者について:審判と判定テクノロジーをめぐる社会学的考察."追手門学院大学社会学部紀要 9(2015):1-15.
- [2] 安田航也,井高 悠斗,何宜欣,田川 憲男,大久保寛.『バスケットボール解析における多視点映像を用いたボール追跡手法』映像情報メディア学会技術報告 2016.9
- [3] 高橋正樹,中村 俊之,三科 智之.『機械学習を利用した複数視点映像からのサッカーボール追跡』映像情報メディア学会技術報告 2014.12
- [4] 高橋裕樹,滑川 崇.『手の動き特徴に基づいた不審動作検出』映像情報メディア学会技術報告 2015.3
- [5] 田摩雅基,青木義満.『映像監視ソリューションの実現にむけた人物動作解析技術*』精密工学会誌 Vol.79,No.11,2013
- [6] Alexandre Alahi, Yannick Boursier, Laurent Jacques, and Pierre Vanderghenst, Sport players detection and tracking with a mixed network of planar and omnidirectional cameras, in Distributed Smart Cameras, 2009. ICDSC 2009. Third ACM/IEEE International Conference on Distributed Smart Cameras. IEEE, 2009, pp. 1–8.
- [7] Jungsoo Lee, Jiwon Lee, Sungwon Moon, Dowon Nam, Wonyoung Y, Basketball event recognition technique using Deterministic Finite Automata (DFA), in 2018 20th International Conference on Advanced Communication Technology (ICACT), 2018.
- [8] K. Soomro, A. R. Zamir, and M. Shah, "{UCF101:} {A} Dataset of 101 Human Actions Classes From Videos in The Wild," CoRR, vol. abs/1212.0, 2012.
- [9] B. G. Fabian Caba Heilbron Victor Escorcia and J. C. Niebles, "ActivityNet: A Large-Scale Video Benchmark for Human Activity Understanding," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 961–970.
- [10] M. D. Rodriguez, J. Ahmed, and M. Shah, "Action MACH a Spatio-temporal Maximum Average Correlation height filter for action recognition," 2008.
- [11] Sarbagya Ratna ShakyaChaoyang, ZhangChaoyang, ZhangZhaoxian , Basketball-51: A Video Dataset for Activity Recognition in the Basketball Game,11th International Conference on Computer Science and Information Technology (CCSIT 2021)
- [12] MartínAbadi,PaulBarham,JianminChen,ZhifengChen, AndyDavis,JeffreyDean,MatthieuDevin,SanjayGhemawat,GeoffreyIrving,MichaelIsard,ManjunathKudlur, JoshLevenberg,RajatMonga,SherryMoore,DerekG.Murray,BenoitSteiner,PaulTucker,VijayVasudevan,Pete Warden,M System for Large Scale Machine Learning.Proceeding softhel2th USENIX Symposiumon Operating System sDesignand Implementation(OSDI'16).November2–4,2016• Savannah,GA,USA.
- [13] BasketBall Data,<https://www.sportingnews.com/>.
- [14] JBAOfficialBasketballRules2019,<http://www.japanbasketball.org/>.