

ディープラーニングを用いたバレーボールにおける 各動作割合の分析

中川樹^{†1} 樋口好彦^{†1} 大野拓摩^{†1} 重永貴博^{†1} 中井一文^{†1} 江崎修央^{†1}

概要: 近年、スポーツにおける IT の活用が注目されており、実際に試合分析や指導に用いられている。バレーボール競技においては、狭いコート上で選手が位置を移動しながらボールを追い、各ポジションの役割をこなす。各選手の動作の質や速さ、移動量などで今後の試合展開に大きな影響を与える。本研究では、試合動画からトスやレシーブ、スパイクなど選手の動作を検出し 1 試合あたりの動作割合を分析する。バレーボール競技の試合動画を、選手の動作を学習させたモデルにかけ、選手の動作を検出する実験を行った。動画からの選手位置・動作の検出にはディープラーニングの物体検出アルゴリズムのひとつである YOLO[1]を用いた。結果として、選手の位置と動作を検出することができた。

キーワード: スポーツ, ディープラーニング, 物体検出アルゴリズム

Analysis of Operating Ratios in Volleyball using Deep Learning

ITSUKI NAKAGAWA^{†1} YOSHIHIKO HIGUCHI^{†1} TAKUMA ONO^{†1}
TAKAHIRO SHIGENAGA^{†1} KAZUFUMI NAKAI^{†1} NOBUO EZAKI^{†1}

Abstract: In recent years, utilization of information technology in sports has been drawing attention, and it is actually used for game analysis and guidance. In volleyball competition, the player follows the ball while moving the position on a narrow court, and performs the role of each position. The quality and speed of each player's movement, the amount of movement and so on will have a major impact on future game development. This research detected actions of players such as Toss, Receive, Spikes from game videos and analyze the operation ratio per game, and conducted experiments to detect the actions of players by using a model that learned the actions of players against the volleyball game match videos. YOLO, which is one of deep learning object detection algorithms, was used to detect player position, movement from motion pictures. As a result, this experiments were able to detect the position and movement of the player.

Keywords: Sports, Deep learning, Object detection algorithm

1. はじめに

近年、スポーツにおける IT の活用が注目されており、実際に試合分析や指導に用いられている。日本のバレーボール競技においてもタブレット端末とスコアによる分析ソフトを用いることで戦術を的確に切り変えるなど、指示を明確に伝えている。しかし、こういった分析には入力に時間や人手が掛かり、また動画と各選手のスコアによる分析を行うだけでは、得点率の高い選手や選手同士のコンビネーションの調子などの情報しか得られない。1 プレー中の選手の位置、次に何の動作の準備をしているかなどの具体的な情報はすぐに可視化できず、選手側と指導者側のイメージが上手く一致しない。選手の技術向上には、より客観的な分析を可能にしなければならないと考えられる。

バレーボール競技においては、狭いコート上で選手が位置を移動しながらボールを追い、各ポジションの役割をこなす。選手はスパイクやトス、レシーブ、ブロックなどのプレーや、ジャンプやダッシュ、飛び込みなど多種に渡る

動作を行う。動作の質や速さ、移動量などで今後の試合展開に大きな影響を与えるため、選手の各動作について分析を行う必要がある。本研究ではバレーボールの試合動画に対して AI による動作分析を行い、選手の動作データを数値や文字に起こすことで可視化し、選手側と指導者側のイメージ一致を容易にすることを目的とする。これにより、部活動などにおける技術指導の質が向上することを想定している。

本稿では、試合動画からトスやレシーブ、スパイクなど選手の動作を検出し 1 試合あたりの動作割合を分析することを目的としている。動画からの選手位置・動作の検出にはディープラーニングの物体検出アルゴリズムのひとつである YOLO を用いる。バレーボール競技の試合動画を、選手の動作を学習させたモデルにかけ、検出する実験を行う。実験結果から、動作検出の成功例・失敗例を考察し、1 試合あたりの動作の割合から、バレーボール競技における動作分析の適用可能性を示す。

^{†1} 鳥羽商船高等専門学校
National Institute of Technology, Toba College

2. 関連研究

2.1 非熟練者と熟練者の違い

バレーボール競技のレシーブにおいて、熟練のための要因はいくつか挙げられる。熟達するには、指導者が選手の各動作始動タイミングの指摘をすることや、レシーブ位置までのステップ動作にも着目して指導を行う必要性がある[2]。

ボールの動きや相手選手の動作に対して、選手が予備動作を行うことで次のプレーに繋がりがやすく、また相手のフェイントなど予期せぬプレーにも反応しやすくなる。例えば、相手選手がボールを返す動作が予測できたときに、あらかじめレシーブの構えができるよう姿勢を整えておくことが予備動作となる。レシーブに限らず、予備動作や構えを行うことで、次の攻撃や選手間のコンビネーションの上達にも繋がりが、チーム全体の熟達に貢献する。

本研究においても、動作始動のタイミングや選手の予備動作の有無、構えなどに注目し指導や試合分析を行う予定である。

2.2 技術指導のためのシステム

スポーツ選手の技術向上のために画像処理を用いたシステム[3]がある。バレーボール競技、卓球競技、スキー競技の3つに対応しており、それぞれ分析対象は異なる。バレーボール競技、卓球競技においてはボールの速度計測、回転数の計測を行い、指導に活かすというものである。スキー競技では、滑走中の選手をストロボ撮影した画像を印刷して指導を書き込む。

しかし、バレーボール競技において球速の分析が技術向上に直結するわけではない。また、静止画を用いての指導では動作の流れがわかりにくく指導者の時間を取ってしまう。

そこで本研究においては、動画に対しての処理を行うことで指導効率の向上を図る。動画での指導では、指導者が指摘したい部分だけでなく、その直前の動作や、その動作をした後の試合展開を選手が直感的に掴みやすくなる。結果として、指導者側の指摘箇所と選手側の改善点が一致しやすく、指導効率が向上する。また、ディープラーニングを用いて動作検出の自動化を行うことで指導者の負担を軽減することができる。

3. 学習モデルの構築

3.1 物体検出アルゴリズムの概要

今回、動作認識には、物体検出アルゴリズムである YOLO を用いた。図 1 は手法全体の流れを示している。画像を $S \times S$ 個のセルに分け、多数の情報を持った矩形領域を保持している。それぞれのセルはクラスごとにそのクラスであるという確率を推測し、最も信頼度の高い枠線を出力すること

で物体検出を行う。また、YOLO 以前のモデルでは物体領域候補を多く挙げ、それぞれに対してクラス分類を行っていたが、YOLO は物体領域候補を選び出さずに 1 つの CNN(Convolution Neural Network)で物体の検出からクラス分類まで予測するという特徴がある。CNN とは、AI が画像分析を行うための学習手法の 1 つである。20 クラスの物体をリアルタイム検出が可能な速度で認識することができる。図 2 に示すように、アーキテクチャは 24 層の CNN が使われており、過学習を防ぐために epoch 数に応じて学習率を変化させているのも特徴である。

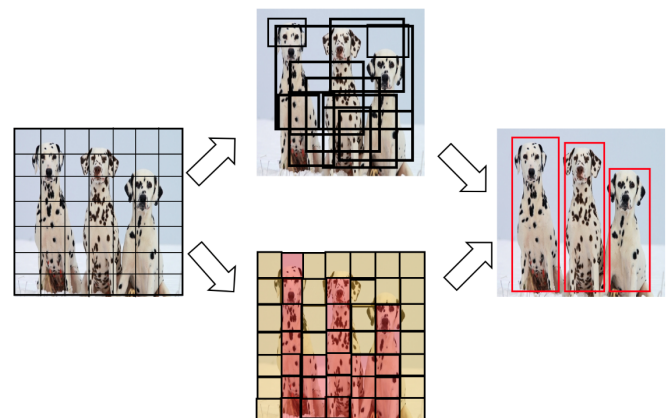


図 1 YOLO アルゴリズム全体の流れ
Figure 1 Flow of the YOLO algorithm

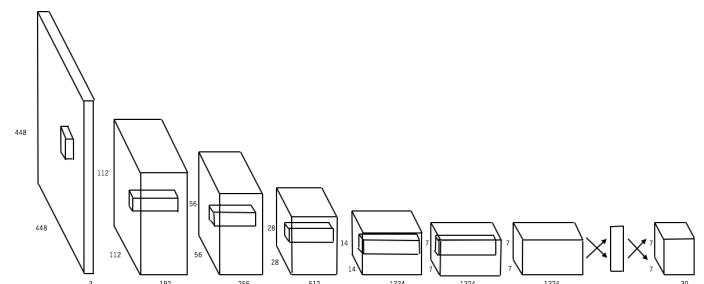


図 2 モデルのアーキテクチャ
Figure 2 Model architecture

3.2 バレーボール競技の試合動画の撮影

バレーボール競技の試合を撮影し、撮影動画から切り出した画像を用いてデータセットの作成を行う。作成したデータセットを用いて機械学習を行い、バレーボール競技における試合中の選手の動作検出を行う。

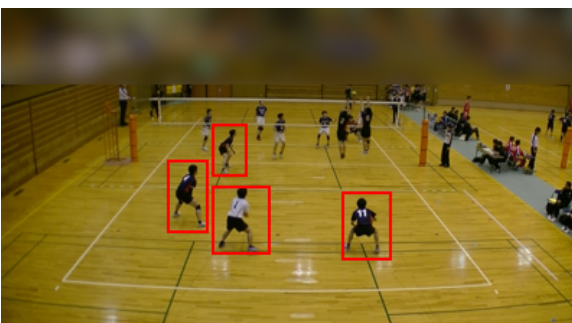
図 3 に示すような画角で撮影を行った。コート手前側の選手を対象者(16~19歳の男子6名)とする。撮影動画のピクセル数は1920×1080であり、フレームレートは30FPSとなっている。



図3 バレーボール競技の試合撮影
Figure 3 Movie of volleyball competition



(a) スパイク
(a) Spike



(b) アンダーハンド
(b) Underhand



(c) オーバーハンド
(c) Overhand

図4 検出対象とする動作
Figure 4 Movement to be detected

3.3 データセットの作成

YOLOの学習には検出したい物体が写っている画像データと、その物体のクラス名、学習させたい範囲の座標が入力されているテキストデータが必要である。撮影した動画から1フレームごとに静止画の切り出しを行い、切り出した静止画の中からスパイク、アンダーハンドでのプレー、オーバーハンドでのプレーといった3クラスの動作を行っている画像を無作為に抜き出した。スパイク、アンダーハンドでのプレー、オーバーハンドでのプレーはバレーボール競技において重要な動作の大部分を占め、得点率やラリーの質に大きく寄与する。

図4(a)~(c)のような画像を各クラス約20枚前後用意し合計64枚とした。次節で述べるが、学習データは学習の際に増幅できるので、1クラス20枚前後で問題はない。

図4(a)では画像左側の選手がスパイク、図4(b)では画像手前側の選手4人がアンダーハンドでのプレー、図4(c)では手前コート中央の選手がオーバーハンドでのプレーを行っている。

図4(a)~(c)の四角い枠で囲んだ部分の4つ角の座標を学習モデルに与える。クラス番号は学習させたい各クラスに対応した番号を割り振っていく。座標データは基の座標データを0~1の範囲内に正規化したものである。図4に示す各クラスの例で与えた座標データは表1のようになっている。

表1 動作種別と学習に用いる座標データの例

Table 1 Number of actions and example of points of detections by learning model

	クラス 番号	中心の X座標	中心の Y座標	中心座標 からの幅	中心座標 からの高さ
スパイク	0	0.58359	0.65555	0.06406	0.17222
アンダー ハンド	1	0.37578	0.44305	0.03593	0.16666
オーバー ハンド	2	0.53984	0.43472	0.02968	0.18055

3.4 動作の学習

3.3節で作成したデータセットをYOLOで学習させる。今回の学習では各画像を2枚に増幅しており、合計128枚のデータセットとなっている。教師データとテストデータの割合は8:2とした。図5は学習中の平均損失(current average loss)の変化を示している。縦軸を平均損失、横軸が学習の反復回数(Iteration number)となっており、平均損

失が 0 へと収束していくのが分かる。平均損失は 0 に収束するほど学習は良好であると言えるが、一定の値以下になると過学習状態となる。過学習状態のモデルでは未知データに対する検出ができない。今回の学習では反復回数 71012 回とした。

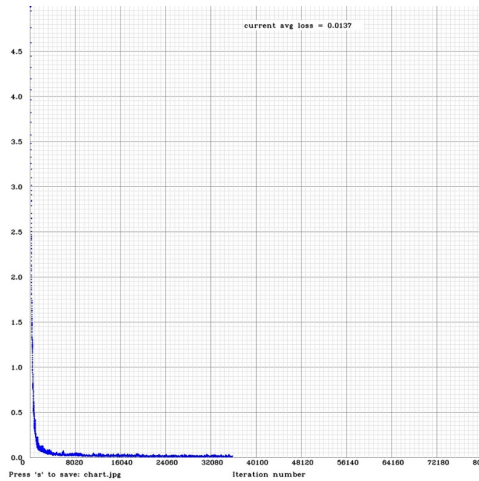


図 5 平均損失の変化
 Figure 5 Change in average loss

4. 検出実験

4.1 実験概要

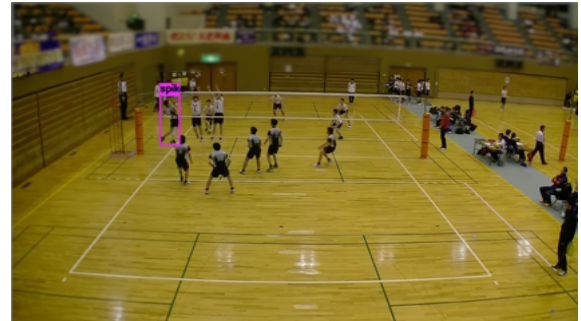
本稿ではバレーボール競技の試合撮影を行い、その試合動画を基に YOLO で学習させるためのデータセットを作成し、3 種類の動作を学習させた。未知データであるバレーボール競技 1 セット分の試合動画に対し、動作学習モデルでの選手 6 人の動作検出を行った。また動作検出で得られたデータを用いて各選手の動作割合を求め、分析を行った。

4.2 検出結果と考察

動画からの動作検出において各動作の成功したフレーム例、失敗したフレーム例を図 6~8 に示す。

図 6~8 の(b)中の枠で囲んだ部分が検出できなかった。スパイクの検出結果である図 6 の(a)と(b)を比較すると、動作検出に失敗している画像はコート上右側からスパイクを打っているのがわかる。同様にアンダーハンドでのプレーの検出結果である図 7 の(a)と(b)を比較すると、動作検出に失敗している画像は選手の体勢が大きく崩れ、しゃがんでアンダーハンドレシーブを行っている様に見える。またオーバーハンドでのプレーの検出結果である図 8 の(a)と(b)を比較すると、動作検出に失敗している画像は選手が後ろ向きにプレーしている。スパイクの学習に用いた画像は手前コート左側から打っていたものが多かった。このことから、図 6(a)の手前コート右側からのスパイクを打っている様な

画像が足りなかったことがわかる。データの偏りを改善することで検出が可能になると考える。また図 7(b)や図 8(b)で示している検出失敗例は、プレー中の選手の体勢を想定していなかったことが原因であると考えられる。また、学習に用いた動作の座標データをより正確性の高いものにするこで、検出精度は上がると考えられる。

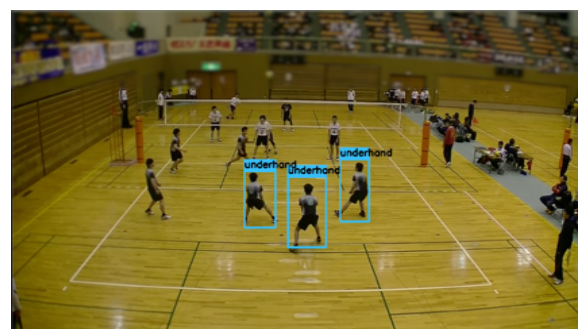


(a) 成功
 (a) Success



(b) 失敗
 (b) Failure

図 6 スパイクの検出結果
 Figure 6 Detection result of spike

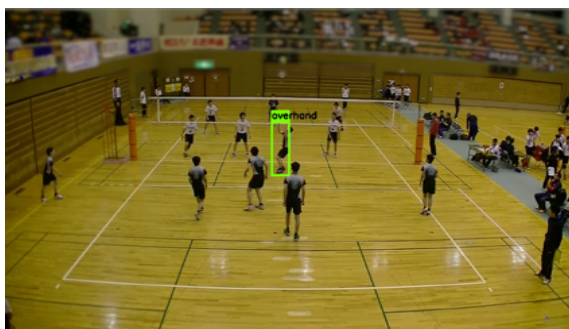


(a) 成功
 (a) Success



(b) 失敗
 (b) Failure

図7 アンダーハンドでのプレーの検出結果
 Figure 7 Detection result of play in underhand



(a) 成功
 (a) Success



(b) 失敗
 (b) Failure

図8 オーバーハンドでのプレーの検出結果
 Figure 8 Detection result of play in overhand

表2 選手の動作数と学習モデルによる検出数
 Table 2 Number of actions of players
 and number of detections by learning model

	選手1	選手2	選手3	選手4	選手5	選手6
動作数	31	47	18	18	23	33
検出数	20	37	14	13	16	24

動作検出結果を表2に示す。動作数は実際に選手が行った各動作を人間が計測した回数、検出数は選手が行った各動作を学習モデルが検出した動作回数である。動作数、検

出数共にスパイク、アンダーハンドでのプレー、オーバーハンドでのプレーの3つの動作数の合計となっている。

今回の実験では、学習モデルでの動作検出が出来ていても動作予測の確率が40%以下であるもの、動作検出が1フレーム分のみで極端に短い動作と判断された場合のものは検出数に含めていない。

検出結果の内、正例で正と検出されたものをTP、負と検出されたものをTNとし、負例で正と検出されたものをFP、負と検出されたものをFNとして検出性能の指標を式(1)に示す。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

表2の結果に(1)式を適用して学習モデルのOverall Accuracy(全体の精度)を求める。Overall Accuracyは0.72であった。この結果から学習の際に画像を増幅し、また反復学習を適正のタイミングで終えられたことがOverall Accuracyを高めた要因だと考えられる。

各選手の動作割合を求める。YOLOにて検出されたフレームから1セット中に選手が構えた時間の割合と選手が構えてから実際にボールコンタクトに成功したときの時間の割合を求め、比較する。インターバル時間を除いた1セット分時間は12分34秒であった。動作割合を求める式を(2)~(3)式に、求めた動作割合を表3に示す。

$$\text{構えの割合} = \frac{\text{構え時間}}{\text{1セット時間}} \quad (2)$$

$$\text{ボールコンタクトの割合} = \frac{\text{構えからボールコンタクトに繋がった時の時間}}{\text{1セット時間}} \quad (3)$$

表3 各選手の動作割合

Table 3 Percentage of each player's action

×10 ⁻³	選手1	選手2	選手3	選手4	選手5	選手6
構えの割合	19.8	51.7	14.0	10.0	14.6	17.2
ボールコンタクトの割合	21.2	10.6	9.3	13.3	15.9	26.5

構えの割合は1セットあたり選手がどれだけ構えに時間を使っていたかを示したものである。ボールコンタクトの割合は1セットあたり選手が構えてからボールコンタクトに繋がったときの動作時間である。選手1~6の構えの割合を比較すると、他の選手より選手2の構えの割合が多いのがわかる。しかし、選手2のボールコンタクトの割合は他

の選手より突出して多いとは言えない。この結果から選手がいくら構えたかといって、必ずしもボールコンタクトに繋がるとは限らないことを示していると考えられる。しかしこれは、バレーボールが6人で1つのボールを追う競技であり、相対的にボールに接触する回数が少なくなることがこの考察に繋がる要因である。

表3の結果のように、動作検出により機械的に算出できるようになれば、今まで利用・分析されていなかったデータや、手動でやっていた分析が自動化できる可能性がある。本稿では基本的な動作であるスパイク、オーバーハンド、アンダーハンドを検出対象としたが、ボールの動きや得失点、相手選手やチームの動作を変数に入れることも考えられる。

5. おわりに

今後、さらに検出精度を上げるために、データセットの見直しを行っていく。画像内の学習させたい動作範囲である座標データを、より正確な値に変えることで、体育館の床など余分な背景を学習せず、誤検出が減少すると考えられる。また、背景色や撮影角度を変えたデータセットも用意して学習させていくことで、試合会場の照明や床の色などの違いに対応できるモデルを構築することを目指す。

本稿で構築したモデルでは4.2節で述べた通り、想定していなかった選手の動作には対応していない。学習の際、4.2節で述べた検出失敗例の様な画像の種類を増やすことで対応できると考えられる。本稿では3種類の動作の検出を行ったが、バレーボール競技における動作は他にも存在する。ブロックやトスフェイントなど試合中に多用する動作も学習に組み込んでいく。さらに、予備動作も学習に追加することで、どの選手がどのタイミングで動き始めたかを判断する材料となる。タイミングの判断ができることで、動画を見る人の知識や判断力に依存せずにチーム内でプレーの指摘や共有が可能となる。以上のように、検出する動作の種類を増やすことで、指導に用いるための情報や試合分析の幅が増加し、指導の質が向上する。

また今後は、動作認識だけでなく、動作認識と数字認識を組み合わせることで各選手の背番号を識別し、動作の解析を行うモデルを構築していく。動画内で試合中の背番号を認識し追いつけ、選手の位置や移動距離などを数値化する。これにより、試合中の移動距離や動作回数、ジャンプの高さなど多様な情報を基に各選手の試合中の疲れの解析を行っていきたい。さらにボールの動きや相手選手の動作検出も分析に含めることで、本研究の最終目標として、誰でも手軽にプロのデータアナリストと同等の分析結果を手に入れられるシステムを目指したい。

参考文献

- [1] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. arXiv:1506.02640[cs.CV].
- [2] 浅井憲一, 大神訓章. バレーボールの左右移動に依る熟練の要因に関する運動的考察. 山形大学紀要 教育科学. 2001, vol. 12, p. 47-58.
- [3] 玉木徹, 牛山幸彦, 八坂剛史. スポーツ選手の技能向上のための動画処理とその実用化. 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解. 2005, vol. 105, p. 13-18