

深層学習による剣道の有効打突判定のための簡易支援システムの可能性

坂本 一磨 (公立小松大学 生産システム科学科, kazuma.sakamoto@komatsu-u.ac.jp)

岩田 伊織 (公立小松大学 大学院サステイナブルシステム科学研究科, 22211005@komatsu-u.ac.jp)

上田 芳弘 (公立小松大学 生産システム科学科, yoshihiro.ueda@komatsu-u.ac.jp)

酒井 忍 (公立小松大学 生産システム科学科, shinobu.sakai@komatsu-u.ac.jp)

坂下 ちえり (元公立小松大学 生産システム科学科, 19111033@komatsu-u.ac.jp)

Possibility of a simplified support system for judging valid strikes in kendo using deep learning

Kazuma Sakamoto (Department of Production Systems Engineering and Sciences, Komatsu University, Japan)

Iori Iwata (Graduate School of Sustainable System Science, Komatsu University, Japan)

Yoshihiro Ueda (Department of Production Systems Engineering and Sciences, Komatsu University, Japan)

Shinobu Sakai (Department of Production Systems Engineering and Sciences, Komatsu University, Japan)

Chieri Sakashita (Department of Production Systems Engineering and Sciences, Komatsu University, Japan)

要約

様々なスポーツにおいてビデオ判定が導入されており、サッカーのVAR、バレーボールのチャレンジシステム等が適用されている。武道の分野では、柔道で2007年からビデオ判定を本格導入している。また、剣道においては、2011年4月に「映像判読訴請規定」として導入され、映像による判読が実施された。剣道を対象とした映像による判読方法は、カメラを3台使用し、審判主任を含む訴請委員3名から5名によって判読を行った。しかし、実装は3大会のみに留まっており、実用化には至っていない。既存研究では、センサやマイクロホンの使用や、直上からの撮影により、聴覚情報と視覚情報から高精度な判定を試みている。しかし、これらの機材を使用できる環境は限られており、地方の小規模な大会やカレッジスポーツ等での導入は難しい。そこで、本研究では、剣道を対象とし、実際の審判員を想定した位置から撮影したカメラ映像を入力として、ResNet-18を用いて学習及び評価を行う。そして、複数方向からのカメラ映像を用いて有効打突の判定を行い、提案手法の実現可能性を検証する。

Abstract

Video judging has been introduced in a variety of sports, including VAR in soccer and challenge systems in volleyball. In the field of martial arts, video judging was introduced on a full scale in judo in 2007. In kendo, three cameras were used to make video judgments, and judgments were made by three to five judges, including the chief referee. However, it has yet to be put into practical use, having only been introduced in three tournaments. Existing studies have attempted to make highly accurate judgments based on auditory and visual information by using sensors and microphones or by shooting from directly above. However, the locations where these devices can be used are limited, making it difficult to introduce them to small-scale local tournaments and university sports. Therefore, in this research, we conduct learning and evaluation of kendo using ResNet-18 with camera images taken from the actual referee's position as input. Then, we then verify the feasibility of the proposed method by obtaining effective hits using camera images from multiple directions.

キーワード

剣道, ResNet, 画像認識, 深層学習, ビデオカメラ

1. はじめに

2021年に行われた東京オリンピック、パラリンピック競技大会では、日本選手の活躍により、過去最高のメダル数を獲得した。この結果からさらなる競技力向上が期待されており、文部科学省スポーツ庁(2022)により、2022年度から2026年度までの5年間で取り組む第3期「スポーツ基本計画」が策定された。その計画の重点施策の一つとして、デジタル技術を活用した新たなスポーツ機会や新たなビジネスの創出などのDX(Digital Transformation)の推進が掲げられている。このような背景に伴い、スポーツとAI(Artificial Intelligence)を組み合わせて、戦術分析や自動採点等が行われている。自動採点の方法の一つとして、ビデオ判定が導入されており、サッ

カーのVAR(Video Assistant Referee)、バレーボールのチャレンジシステム等、様々な分野で適用されている。武道の分野に着目すると、2007年から柔道でのビデオ判定が、本格導入されており、市村・中村(2017)の研究においても信頼性も含めて言及されている。しかし、他の武道では、導入が進んでいない。剣道においては、韓国で2011年4月に「映像判読訴請規定」として導入されており、映像による判読が実施され、加藤(2012)によって分析が行われている。しかし、日本においても本格的な導入には、至っていないのが現状である。

剣道の勝敗を左右する有効打突の判定は、全日本剣道連盟(2019)より、競技者が打突を試みた際、打突部位である「面」、「籠手」、「胴」と「突き」の4箇所(以下、有効打突部位)の打突に関して、審判に有効打突として認定することにより、判定される。しかし、この判定は、審判の主観的な判断に依存する部分が多いため、客観的に評価するシステムが現時点では

存在しない。

既存研究である渡邊隆 (2020) は、ハイスピードカメラでの撮影画像と指向性マイクロホンにて取得した打突音から打突動作の有効性を評価している。しかし、本論文で述べられているとおり、評価において、様々なデバイスの使用や、試合を直上から撮影し、撮影された映像を使用しているため、機器を試合会場の直上に設置する必要がある。これにより、システムが導入可能な環境は限られる課題がある。また、飯田・後藤 (2014) では、距離画像センサである Kinect を使用することによって、身体運動を定量化し、竹刀の振り動作の自動判別アルゴリズムを構築することを試みている。しかし、環境によって大きく判定成功率が変化することや人物検出の範囲が 0.5 m から 4.5 m 以内という制約がある。

そこで、本研究では、様々な環境で使用することを想定し、どのような環境下でも導入が容易であるビデオカメラ、タブレット端末やスマートフォンといった動画撮影が可能な機器のみを用いた有効打突の判定を試みる。既存研究と本研究での撮影機器等の比較を表 1 に示す。表 1 に示すとおり、Kinect やハイスピードカメラといった特別な機器を使用せず、設置場所は審判員の配置を考慮して、撮影することを想定している。これにより、既存のビデオカメラやタブレット端末等を使用することができるため、金銭的成本や設置コスト抑えることが可能である。また、地方の小規模な大会やカレッジスポーツ等でも導入が容易に行うことができるため、競技力向上に寄与できると考えられる。

検証実験では、実際の審判員を想定した位置から撮影したカメラ映像を用いた剣道の勝敗判定支援システムを提案し、実現可能性を検証する。

2. 研究概要

2.1 研究目的

本研究は、次に示す 2 つの研究課題を設定し、検証実験を通して解決することで、カメラ映像を用いた剣道の勝敗判定支援システムを提案する。また、既存研究と比較して、撮影のための設営、使用が比較的容易なビデオカメラを使用することを検討し、様々な場面でのシステムの導入を試みる。

2.1.1 研究課題 1：カメラ映像から打突の自動判定が可能かの検証

現状の剣道に関する研究では、飯田・後藤 (2014) が Kinect を用いて、剣道の基本的な技自動判定システムの構築が行われている。しかし、これらの手法は撮影環境に大きく左右される点や技の判別は可能でも実際にその技が有効か否かの判

別が難しいことが現状である。

本研究では、一般的に販売されているビデオカメラやタブレット端末を使用し、環境に左右されることがないシステムの構築を目指す。これにより、所有しているビデオカメラを用いることが可能である。また、プロのみではなくアマチュアスポーツ等に活用が可能なため、競技力向上にも寄与できると考えられる。また、ビデオカメラを用いることで、距離による制約が緩和されるため、離れた箇所から撮影した場合においても精度が担保可能である。

以上のように、一般的なカメラで撮影した映像を入力として、深層学習を使用し、打突の判定を自動化するシステムを開発する。そして、検証実験を通して精度を評価し、研究課題 1 を証明する。

2.1.2 研究課題 2：複数台のカメラを使用することで打突の判定精度が向上するかの検証

単視点での撮影では、オクルージョンや死角によって、打突判定が難しいパターンが存在する。また、竹刀や選手の腕が背景の色と同化するという場合も打突判定が困難になることが予測される。剣道試合審判規則 (2019) おいても審判員は、主審 1 名と副審 2 名を原則として、有効打突および、その他の判定について同等の権限を有するということが明記されている。

本研究では、審判員の配置を考慮して、複数台のカメラでの撮影を試みる。そして、各視点での打突判定結果に対し、多数決を用いて、最終的な判定を行うシステムを構築する。これにより、単視点では、難しい判定においても正確に判定可能であるかを検証して評価し、研究課題 2 を証明する。

2.2 提案手法

本研究の処理フローを図 1 に示す。本システムでは、カメラから撮影した映像を入力とし、画像生成機能、データセット作成機能、学習・評価機能により、入力画像に対する正解率を出力する。また、使用した深層学習のネットワークは、Kaiming et al. (2016) が提案した ResNet-18 ネットワークに入力画像を対応できるように修正したシステムである。

2.2.1 画像生成機能

本機能では、映像から学習及び評価に用いる画像を生成する。生成方法としては、まず、複数台のカメラで時刻同期を行い、撮影した映像を入力として、フレーム画像に分割にする。そして、それらの画像を 480 px × 480 px でトリミングする。時刻の同期方法は、梅原ら (2021) の研究を参考に複数カメラで撮影した映像の最初と最後のフレームを抜き出し、フレーム

表 1：既存研究と本研究での撮影機器等の比較

	機器	機器台数	設置場所
渡邊 (2020)	ハイスピードカメラ +マイクロホン	2	直上、側方
飯田・後藤 (2014)	Kinect	1	側方
本研究	ビデオカメラ等	3	側方 (審判位置)

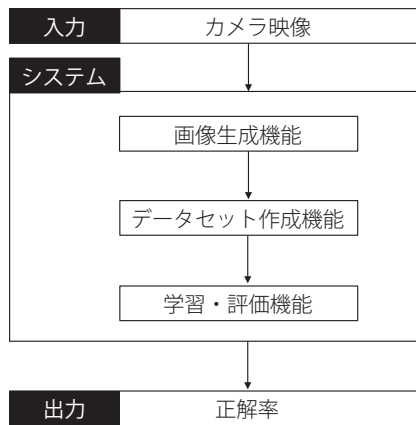


図1：処理フロー

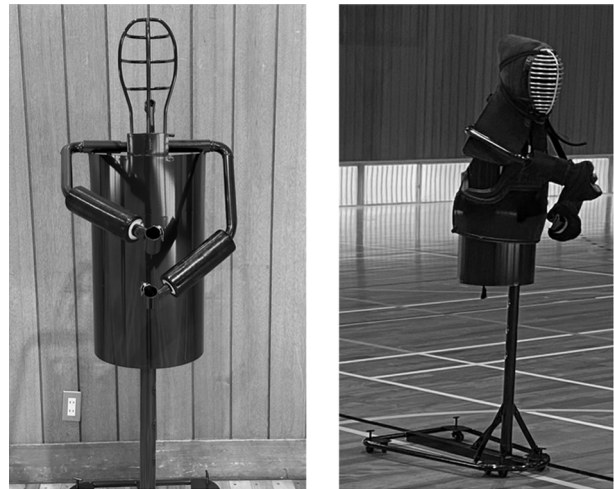


図2：打ち込み台

数を合わせて、フレームごとに時刻同期を行った。

2.2.2 データセット作成機能

本機能では、画像から学習と評価に使用するデータセットを作成する。データセットとは、打突、その他の画像をまとめたデータ群である。作成方法としては、まず、各画像に対し、打突の瞬間とその他の2つに分類する。そして、それらを学習データと検証データに分割する。本研究では、打突の瞬間は目視で判断し、検証データ数は、学習データ数の1割で設定する。

2.2.3 学習・評価機能

本機能では、データセット作成機能で作成したデータセットを用いて、学習及び評価を行う。本研究では、ResNet-18を用いて学習を行い、検証データの評価を行う。使用したネットワークは、Kaiming et al. (2016)が提案したResNet-18ネットワークであり、畳み込み層20、プーリング層は最初の層がMaxpooling、最後の層はAveragepoolingとした。また、モデルの判定結果は、学習が完了したモデルに対して、検証データを入力し、精度を評価した。精度評価は、打突かその他の割合を算出し、値が大きい方を結果として使用して、その値で評価をした。

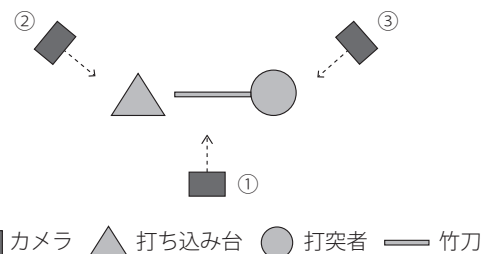


図3：カメラの配置図

3. 実験概要

3.1 実験環境

本研究の撮影場所は、公立小松大学粟津キャンパスの体育館とし、打ち込み対象は高さ1.6mとして、図2に示す人形式打ち込み台を使用する。照明は、体育館設置の照明を全点灯して使用した。カメラアプリは、Apple社提供の純正のカメラアプリを使用し、カメラアプリを実行するハードウェアは、表1に記載されているビデオカメラとして、タブレット2台(iPad Air (第4世代))、スマートフォン1台(iPhone14 Pro)を使用した。本研究では、打突の瞬間を確実に捉えるため撮影する映像のフレームレートを240fps、解像度を1080p、倍率を1倍で設定した。打ち込み台と打者、カメラの配置を図3、撮影した映像の一部を図4に示す。カメラと打ち込み台の距離は、約3～4mとした。



図4：撮影例

3.2 実験方法

3.1節における実験環境で撮影した映像をフレーム画像に分割し、打突の瞬間における画像とその他にラベリングしたデータセットを作成する。そして、学習及び評価を行い、出力されたラベルと正解ラベルとの正解率を確認する。入力画像は、表2に示す条件で作成する。また、各条件のデータセットの枚数を表3、検証結果を表4に示す。以下、条件番号と表2、表3、表4の番号はそれぞれ対応している。

条件(1) (表2～表4 No. 1) では、打突の有無以外の条件を揃えて、撮影画像から打突の判定自体が可能なのかを検証する。

条件(2) (表2～表4 No. 2) では、撮影角度と距離、背景が

表2：入力画像条件

No.	条件
1	面打ちを対象とし、学習データと検証データを同日時のものを使用
2	面打ちを対象とし、学習データと検証データの撮影条件が異なるものを使用
3	条件(2)の手元と打突部位に注目したものを使用
4	条件(3)で不正解であった画像の他方向から撮影したものを使用
5	籠手打ちを対象とし、学習データと検証データを同日時のものを使用
6	籠手打ちを対象とし、条件(5)の打突部位が認識しづらい画像を除いたものを使用

表3：学習データセット

No.	データ枚数
1	学習データ：12,000枚、検証データ：1,200枚
2	学習データ：12,000枚、検証データ：1,200枚
3	学習データ：12,000枚、検証データ：1,200枚
4	学習データ：12,000枚(条件(3)の学習データと同様) 検証データ：条件(3)の不正解である画像から100枚
5	学習データ：4,825枚、検証データ：607枚
6	学習データ：3,217枚、検証データ：405枚

表4：実験結果

No.	打突時	その他	正解率(全体)
1	1.00	1.00	1.00
2	0.21	0.94	0.57
3	0.92	0.72	0.82
4	0.88	1.00	0.94
5	0.97	0.68	0.82
6	1.00	1.00	1.00

異なる場合でも精度が維持できるかを検証する。

条件(3) (表2～表4 No.3)では、背景による影響を確認するため、打突の瞬間を特徴付ける打者の手元と打突部位のみが映るようにトリミングした画像を使用し、背景や打者の差異を極力小さくする。そして、条件(2)の結果との比較によって、背景や打者の差異による影響を検証する。

条件(4) (表2～表4 No.4)では、単視点からの画像で正しく判定することが難しい画像を他方向からの複数視点での画像で判定し、実際の審判員と同様、3方向からの判定の多数決によって、最終的な打突の判定が可能か検証する。

条件(5) (表2～表4 No.5)では、籠手打ちを対象とし、面打ちと同様に打突自体の判定が可能か検証する。

条件(6) (表2～表4 No.6)では、条件(5)において、判定精度が低い角度からの画像を除いたデータを使用することで、死角による影響を検証する。

3.3 検証結果

検証結果(表4)に示すとおり、打突時のラベルは、打突時の画像の内、正しく打突と判断した割合、その他のラベルは、打突時以外についての割合を示している。正解率は、打突時とその他の両方に対する正解数の割合を表している。

条件(1) (表2～表4 No.1)において、面打ちを対象に学習データと検証データを同日時のデータを使用し、条件を揃えた場合、高精度で分類ができた。学習データと検証データがほぼ同じ条件である場合、打突の判定自体は可能であるとわかった。

条件(2) (表2～表4 No.2)において、撮影角度や距離、背景が異なるデータを使用して評価したところ、精度が大きく低下し、特に打突時の精度が約2割と非常に低い結果となった。

条件(3) (表2～表4 No.3)において、打者の手元と打突部位のみが映るようにトリミングした画像を使用した場合、精度は約8割まで向上した。

条件(4) (表2～表4 No.4)で一方向から正しく判定することが難しい画像において、他の2方向から撮影された画像による判定を行った場合、9割を超える高精度であった。

条件(5) (表2～表4 No.5)においては、籠手打ちを対象とし、面打ちと同様の撮影条件の場合、精度が高い傾向となった。

条件(6) (表2～表4 No.6)において、打ち込み台の斜め後ろ方向からの画像を除き評価した場合、精度は非常に高くなった。

3.4 考察

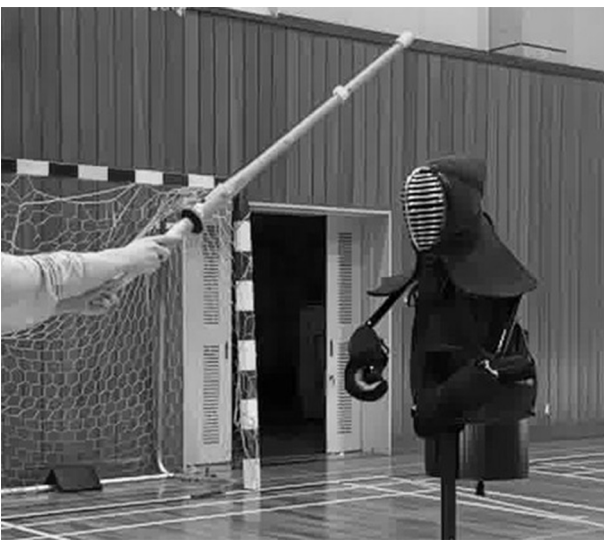
本提案システムは、学習データを取得する際の条件を実際の大会会場と同一の場所、撮影条件にすることで、大会でも使用できる可能性が示唆される。

条件(1)や条件(2)の結果より、角度や距離は大きく変化せず、背景のみが違うことによって、精度が大きく低下していることから背景による影響が大きいと考えられる。竹刀を認識する際に背景の色が竹刀と同じ場合に精度の低下がみられるため、背景の色が竹刀と同色でない角度からの映像を使用するなど、撮影角度を工夫し、背景と竹刀の色が同様とならないようにする必要がある。このことから背景や距離、打者等の条件の差異は判定精度に大きく影響すると考えられる。また、打ち込み台の斜め後ろからの画像が、特に精度が低かった。この要因として、打突部位が他方向よりも比較的認識しづらく、打突の瞬間が曖昧になることが原因であると考えられる。そのため、試合中、選手の位置が激しく動きのある剣道は、選手同士が交差することがあり、カメラの撮影位置の工夫や複数台のカメラの使用が必要となるため、提案手法での多数決での判定が有用であると考えられる。

条件(4)で誤判定であった画像を図5に示す。図5(a)では、竹刀と打突部位が比較的近いことや竹刀がぶれていることで、誤判定に繋がったと考えられる。図5(b)では、竹刀と打突部位が十分離れており、画像のぶれもないが、打突していると誤判定している。これは、条件(1)や条件(2)の時と同様に竹刀と壁の色が酷似していることや、模様による影響等



(a) 誤判定結果例1



(b) 誤判定結果例2

図5：条件(4)での誤判定の画像例

によって誤判定に繋がったと考えられる。

条件(6)の結果においては、打ち込み台の斜め後ろ方向からの画像の精度が非常に低く、他方向に比べて、打突部位が認識しづらいことが原因であると考えられる。また、誤判定の画像の内、打ち込み台の斜め後ろ方向からの画像は打突しているにもかかわらず、打突していないと判定した。この要因として、籠手は面よりも死角になりやすいことから面打ちを行った条件より、精度が低下する傾向があると考えられる。

以上の結果より、実際の審判員同様、一方向のみでは判定難しい場合でも他の2方向を考慮することで、打突の判定が可能であると考えられる。また、条件(4)と同様に3方向の内、2方向からの評価をもとに判定が可能であると考えられる。さらに、視点に着目すると、単視点では、打突の判定の精度が安定せず、トリミングなどの加工が必要になると考えられる。しかし、複数視点で撮影を行うことで、精度は向上する傾向があることがわかった。試合では、カメラを複数台使用し、複数視点の多数決による判定によって、高精度に有効打突を判定できることがわかった。

4. おわりに

本研究では、剣道の面打ち及び籠手打ちを対象とし、カメラで撮影した映像を使用して、様々な条件下での打突の判定が可能かを検証した。実験の結果、打突部位が死角となる角度からの判定が難しく、籠手打ちにおいては面打ちよりも死角になりうる角度が多いため、精度が低くなる傾向がみられた。しかし、他方向から撮影した映像による判定を含め、3方向の多数決で最終的な打突の判定を行うことで、高精度な結果となった。

以上より、本研究では、試合会場の直上ではなく、すぐに設置が可能な真横からの撮影により、打突判定が可能のため、既存研究より様々な環境での導入が可能である。また、安価なビデオカメラやタブレット端末のみを使用するため、カレッジスポーツやアマチュアスポーツにも導入が容易であり、競技力向上にも寄与できると考えられる。

引用文献

- Kaiming, H., Xiangyu, Z., Shaoqing, R. and Jian, S. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.
- 飯田大介・後藤淳 (2014). Kinectを用いた剣道の基本技自動判別システムの構築—非専門家にも優しい剣道支援システムへ—. 福井大学大学院工学研究科研究報告, Vol. 63, 5-13.
- 市村さやか・中村剛 (2017). 柔道競技における審判員の判定能力に関する運動学的研究. *スポーツ運動学研究*, Vol. 29, No. 0, 15-28.
- 加藤純一 (2012). 韓国における剣道試合の有効打突判定に関する一考察—韓国実業剣道連盟による映像判読訴請規定制定までの流れとその実施過程を踏まえて—. *武道学研究*, Vol. 45, No. 1, 1-21.
- 加藤純一 (2012). 韓国実業剣道連盟「映像判読訴請規定」に則して行使された訴請の内容とその判読結果の分析. *武道学研究*, Vol. 45.
- 文部科学省スポーツ庁 (2022). 第3期スポーツ基本計画. https://www.mext.go.jp/sports/content/000021299_20220316_2.pdf. (閲覧日：2022年11月1日)
- 梅原喜政・山本雄平・姜文淵・寺口敏生・田中成典・佐藤衛 (2021). 人物計測のためのビデオカメラ間およびビデオカメラとGNSSの時刻同期に関する研究. *写真測量とリモートセンシング*, Vol. 60, No. 3, 129-143.
- 渡邊隆 (2020). 画像・音声認識による剣道競技者追跡と有効打突判定条件の定量化. <https://kaken.nii.ac.jp/ja/file/KAKENHI-PROJECT-17K01708/17K01708seika.pdf>. (閲覧日：2022年11月1日)
- 全日本剣道連盟 (2019). 剣道試合・審判規則 剣道試合・審判細則. https://www.kendo.or.jp/wp/wp-content/themes/kendo/assets/library/pdf/kendo-shiai_regulations.pdf. (閲覧日：2022年11月1日)

(受稿：2023年4月28日 受理：2023年6月26日)