

ディープラーニングを用いたスポーツ動作分析支援へのチャレンジ

赤木亮太¹, 城所収二²

1: 芝浦工業大学 システム理工学部 生命科学科

2: 国立スポーツ科学センター スポーツ科学部

1. 緒言

スポーツ科学分野では、競技力向上や障害予防に貢献すべく、歩行や走行、跳躍、ピッチングやバッティング等、様々な場面、動作を対象として、バイオメカニクス的手法を用いた動作分析が実施されている。その際、ビデオ画像やモーションキャプチャシステムなどの各種センサを用いて、必要な情報を自動または手動で取得し、3次元座標を推定している。これらの分析には多大な時間やコストを要するだけでなく、被験者にも負担を強いることが多いのが現実である。それ故、正確性を維持しつつ、作業時間を短縮することは、動作分析を対象とする研究の推進、ひいては、スポーツ科学分野において非常に価値あることである。

今回我々は、このような動作分析に、人間が自然に行うタスクをコンピュータに学習させる機械学習の手法のひとつであるディープラーニングを適用することを考えた。近年盛んに取りあげられているディープラーニングは、人工知能の急速な発展を支える技術であり、様々な分野への実用化が進んでいる。その中には画像認識も含まれており、ディープラーニングの動作分析に対する親和性は高いことが推察される。また、将来的なビッグデータの分析等への進展性も期待される技術である。本研究では、個人差はあるものの、機械的に見れば似通っていると考えられる野球のバッティング動作を例に取り、ディープラーニングを用いた新たな動作分析手法の確立可能性について検討することを目的とした。

2. 方法

2.1 被験者

小学生～大学生の野球部員 72 名（年齢：14±4 歳，身長：157±17 cm，体重：49±17 kg，競技歴：5±4 年）を対象とした。本研究は国立スポーツ科学センターの倫理審査委員会の承認を得たうえで実施した。実験に先立ち、被験者に対して、研究の目的や実験手順、研究への参加が任意であることを説明した。併せて、被験者本人から、あるいは、被験者が未成年者の場合はその保護者（代諾者）から、実験参加の同意を書面により得た。

2.2 実験手順

実験は国立スポーツ科学センターの屋内実験場あるいは各チームの屋外グラウンドにて実施した。ホームベースから 15～17 m 離れた地点よりボールが投射されるよう、スプリング式の簡易型ピッチングマシン（L60111J, Louisville Slugger 社製）を設置した。本研究では、ディープラーニングを用いて、野球のバッティング動作時のバットヘッド及びグリップエンドの中心点の抽出にチャレンジした。それ故、被験者が普段通りのスイングを行える投球速度（70～110 km/h）を設定した際のバッティング動作を分析対象とした。各被験者には、24～52 スイングのバッティングを行わせた。その際、グリップエンドに近い位置を把持するスイング、グリップエンドからこぶし 1 個分余らせるスイングの双方を織り交ぜながら実施した。各被験者には、自身のスイングを崩さない程度の全力で、かつ、スイングの意識づけを統一させるためにセンター返しを心掛けるよう指示した。

バッティング動作中のバットの挙動を、電氣的に同期した 2 台の高速度カメラ（Phantom V311 及び

V310, Vision Research 社製), あるいは 20 台の赤外線カメラを利用したモーションキャプチャシステム (VICON MX シリーズ, VICON 社製) を用いて記録した. 高速度カメラについては, 撮影速度 1000 fps, 露光時間 0.4 ms (屋外測定時は 0.2 ms), 画面解像度 1280×800 px で記録し, モーションキャプチャシステムについては, 撮影速度 500 fps で記録した. ただし本報告では, 2 種類の機器で測定したデータの内, 既にデータ処理が完了している高速度カメラ映像に対する手法を説明する.

各カメラは, 2 台のカメラの光軸がおおよそ 90° になることと, インパクト局面のボールとバットが打者によって隠れないよう, 右打者ならば 1 台 (カメラ 1) を一塁ベース方向に, もう 1 台 (カメラ 2) を打者の右斜め後方に設置した (図 1). 左打者の場合には, カメラを右打者の配置からセンターラインを基準に反転させた位置に設置した. 撮影した画像の画角はおおよそ横 3.2 m \times 縦 2 m であり, 1 ピクセルは約 2.5 mm に相当する.

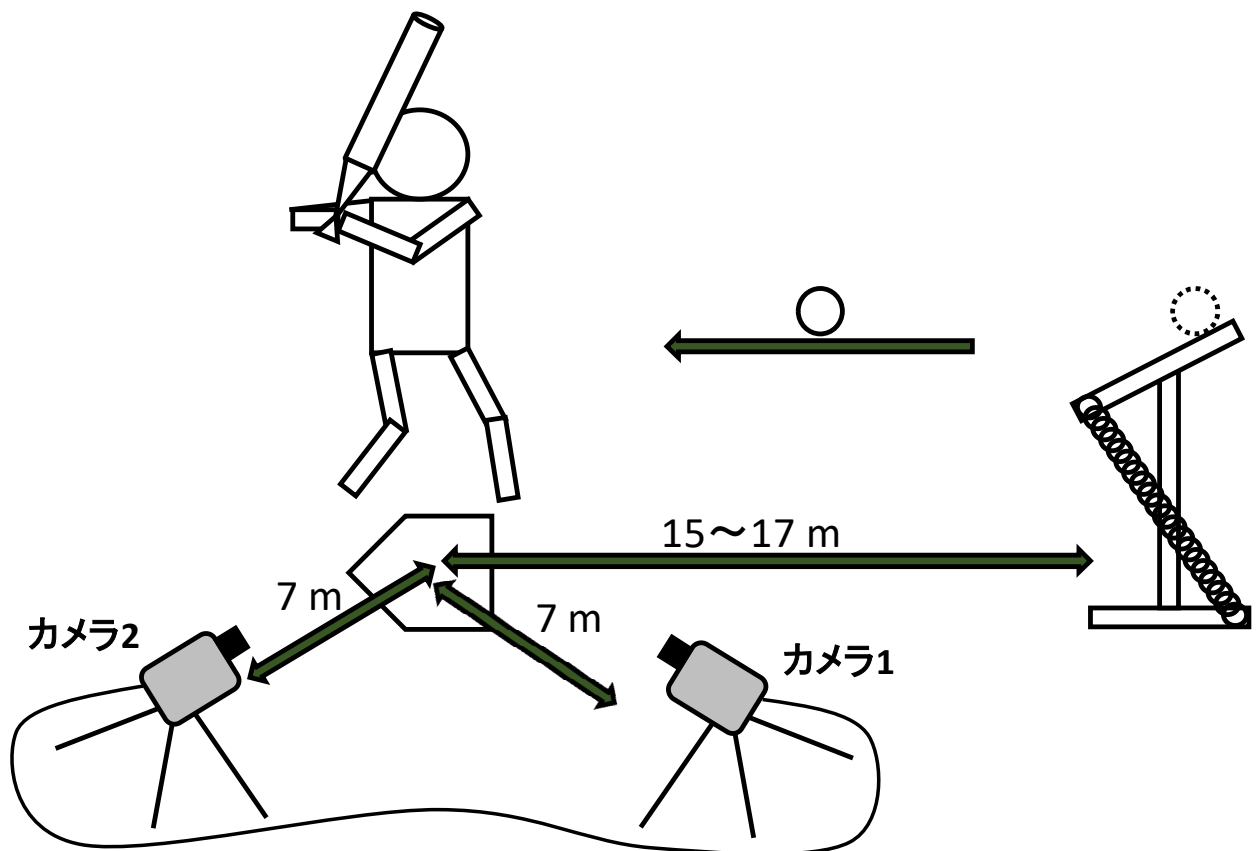


図 1. カメラ設置の概略図

試技を計測する前に, 3 次元空間を構築するためのキャリブレーションを行った. キャリブレーションには, 20 cm 間隔で紅白に色分けされた 2 m のポールを用いた. ポール上の色が変わる位置 (10 ヶ所) をコントロールポイントとし, このポールを撮影画角内の地面 12 ヶ所へ鉛直に立て, 計 120 点を撮影した. キャリブレーションにおける標準誤差は, 全ての実験日で xyz の座標について 2 mm 以内に収まっていた.

2.3 データ処理

ホームベースの後ろ端を原点とし、水平面上の投手方向へ向かうベクトルをY軸、鉛直上向きをZ軸、Y軸とZ軸を外積させたベクトルをX軸として静止座標系を定義した。本研究における打者の表記は、全ての打者を右打者とみなせるように変換して記述した。撮影されたバットの挙動は、画像解析ソフトウェア（Frame-DIAS V, DKH社製）を用いて分析した。デジタルイズを行う分析区間は、“インパクトの瞬間を含む手前401フレーム（0.40 s）”から“インパクト後21フレーム（0.02 s）”までとした。デジタルイズポイントはバットヘッドの中心点及びグリップエンドの中心点とし、これら2点の位置を視覚的に判別することにより、マニュアルデジタルイズを実施した。各デジタルイズ点の3次元座標を、DLT法を用いて取得した。ここで、マニュアルデジタルイズによる分析精度を確保するために、1) デジタルイズの軌跡が滑らかであること、2) 既知の長さのバット長に対する算出されたバット長の差が、全ての区間で ± 5 mm未満に収まること、3) 既知の長さのバット長に対する算出されたバット長の差が、インパクトの直前は ± 3 mm未満に収まること、の3つが満たされない場合には再度デジタルイズを実施した。取得したバットの3次元座標は、4次のバターワース型ローパスフィルタを用いて平滑化した。インパクト直前のバット座標に関しては、インパクト前20フレームの座標値を基準に点対称に回転させたデータを、インパクト後の座標値として仮想的に追加したうえで平滑化を行った。インパクト直前のバット座標の生データを極力歪めないこととノイズが抑えられることの両者が最大限満たされるよう、遮断周波数は35 Hzとした。デジタルイズの残像の典型例は図2の通りである。

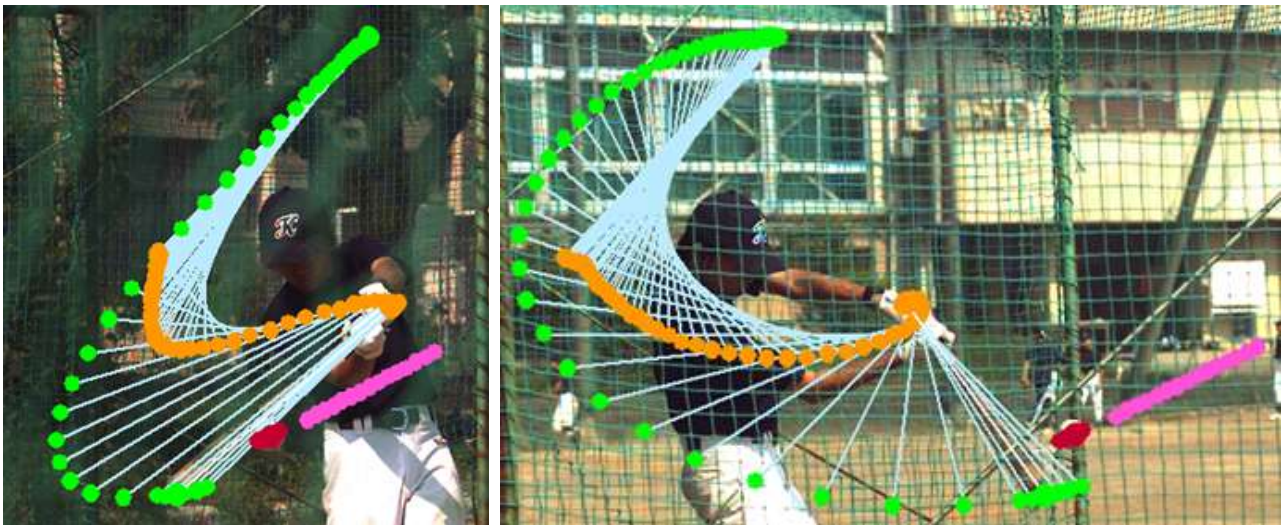


図2. デジタルイズの残像の典型例（左：カメラ1，右：カメラ2）

2.4 ディープラーニング

本研究では、ディープラーニングによる物体検出モデルとして You Only Look Once (YOLO) ^[1]を用いた。その際、1) YOLOが Common Objects in Context (COCO) ^[2]に用意されているデータセットを学習する手法（モデル1）、2) YOLOが実験で得られた動画を学習する手法（モデル2）、の2つの手法を採用した。そして、両者の手法で得られた結果を比較、検証することにより、ディープラーニングの打撃動作分析への適用可能性について検討した。前述のデータ処理が既に終わっていた168試技分のデータの内、158試技分を学習に、残りの10試技分を学習効果の検証に使用した。その際、各試技において、1) 全画像の内、バットヘッド及びグリップエンドの中心点を検出できた画像の割合、2) 両モデル、各々

において検出できた画像における，マニュアルデジタイズで求めた座標との誤差，3) 両モデル共に検出できた画像における，マニュアルデジタイズで求めた座標との誤差，の3つを算出した．なお，2)及び3)における誤差の表記は，実空間での誤差ではなく，カメラ座標上でのピクセル数となる．

2.5 統計分析

全ての測定項目を平均 ± 標準偏差で示す．YOLO の学習効果を検証するために，カメラ1，カメラ2，それぞれにおいて，以下の項目におけるモデル1とモデル2の違いを対応のあるt検定を用いて検討した：1) 各試技に含まれる全画像の内，バットヘッド及びグリップエンドの中心点を検出できた画像の割合，2) 各試技における，両モデル共に検出できた画像を対象とした，バットヘッド及びグリップエンドの中心点の，デジタイズデータからの誤差（x座標，y座標ともに）．危険率5%未満をもって統計的に有意とした．

3. 結果及び考察

3.1 カメラ1

モデル2において検出できた画像の割合は，いずれの試技においても，モデル1のそれよりも増加した（モデル1：52%±21%，モデル2：96%±5%）．それ故，両者の間に有意な差が認められた．

各試技における，バットヘッド及びグリップエンドの中心点の，デジタイズデータからの誤差を，それぞれ表1，表2に示す．バットヘッドに関しては，試技1のy座標，試技7及び試技9のx座標（表1の灰色部分）を除き，いずれもモデル2の誤差の方が小さい値を示した．一方，グリップエンドに関しては，x座標，y座標共に，いずれの試技においても，モデル2の方が低値となった．

表1 バットヘッドの中心点のデジタイズデータからの誤差（カメラ1）

	モデル1		モデル2	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]
試技1	10.7 ± 6.1	9.9 ± 8.9	9.3 ± 11.1	14.5 ± 11.8
試技2	7.2 ± 7.5	12.4 ± 11.6	4.6 ± 4.1	3.5 ± 4.1
試技3	8.0 ± 5.1	13.2 ± 13.9	3.9 ± 4.6	3.1 ± 2.2
試技4	5.8 ± 5.4	17.6 ± 18.4	4.9 ± 7.7	2.3 ± 2.0
試技5	14.2 ± 17.3	19.8 ± 4.7	6.7 ± 4.9	2.9 ± 3.0
試技6	10.0 ± 6.7	14.8 ± 8.6	8.0 ± 5.6	5.3 ± 8.9
試技7	3.9 ± 3.5	17.7 ± 7.8	4.5 ± 5.5	4.4 ± 8.7
試技8	8.1 ± 4.3	16.1 ± 7.0	6.8 ± 10.3	8.5 ± 18.5
試技9	4.6 ± 4.8	11.0 ± 4.2	9.6 ± 10.8	6.1 ± 9.6
試技10	9.1 ± 7.1	14.5 ± 6.2	5.0 ± 4.1	2.9 ± 2.7

表2 グリップエンドの中心点のデジタイズデータからの誤差 (カメラ1)

	モデル1		モデル2	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]
試技1	35.2 ± 21.2	58.9 ± 21.7	5.8 ± 5.0	4.2 ± 2.8
試技2	44.3 ± 32.5	27.4 ± 24.2	6.4 ± 4.1	4.1 ± 3.0
試技3	56.8 ± 10.3	88.1 ± 18.7	8.2 ± 5.8	10.3 ± 6.3
試技4	45.9 ± 28.1	35.5 ± 23.7	6.4 ± 5.1	4.9 ± 3.6
試技5	100.0 ± 37.4	24.3 ± 17.8	12.1 ± 7.7	5.7 ± 2.6
試技6	41.0 ± 11.8	22.0 ± 9.8	11.9 ± 9.7	4.7 ± 3.9
試技7	35.1 ± 10.6	8.3 ± 5.4	5.8 ± 5.4	3.4 ± 3.2
試技8	74.5 ± 23.9	34.7 ± 23.9	11.8 ± 11.4	5.5 ± 3.8
試技9	103.9 ± 29.4	46.9 ± 23.8	8.5 ± 6.9	6.7 ± 4.3
試技10	75.5 ± 46.6	60.7 ± 38.6	3.7 ± 3.3	6.9 ± 5.0

表3及び表4に、各試技において、モデル1・モデル2の双方で検出できた画像に限定して算出した、バットヘッド及びグリップエンドの中心点のデジタイズデータからの誤差を示す。バットヘッドに関しては、試技1のy座標及び試技6のx座標(表3の灰色部分)を除き、モデル2の誤差はモデル1のそれよりも有意に低い値を示した。グリップエンドに関しては、いずれの試技においても、x・y座標の双方で、モデル2の誤差はモデル1の誤差よりも有意に低い値となった。

表3 両モデルで検出された画像でのバットヘッドの中心点のデジタイズデータからの誤差 (カメラ1)

	モデル1		モデル2		有意差(モデル間)	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]	x座標	y座標
試技1	10.8 ± 6.1	10.0 ± 8.9	7.5 ± 3.6	19.7 ± 12.0	1>2	1<2
試技2	7.2 ± 7.5	12.4 ± 11.6	4.1 ± 2.4	3.6 ± 4.2	1>2	1>2
試技3	7.8 ± 5.1	13.6 ± 14.0	3.9 ± 2.8	1.9 ± 1.4	1>2	1>2
試技4	5.8 ± 5.4	17.6 ± 18.4	3.7 ± 2.6	1.9 ± 1.6	1>2	1>2
試技5	14.2 ± 17.3	19.8 ± 4.7	6.1 ± 3.4	2.1 ± 1.9	1>2	1>2
試技6	10.0 ± 6.7	14.8 ± 8.6	10.7 ± 8.2	3.5 ± 3.2	—	1>2
試技7	3.9 ± 3.5	17.7 ± 7.8	2.6 ± 2.9	1.7 ± 2.1	1>2	1>2
試技8	8.3 ± 4.3	16.9 ± 6.5	4.8 ± 4.6	4.3 ± 2.9	1>2	1>2
試技9	4.6 ± 4.8	11.0 ± 4.2	2.9 ± 1.5	2.3 ± 2.5	1>2	1>2
試技10	9.1 ± 7.1	14.5 ± 6.3	4.7 ± 4.0	3.0 ± 2.8	1>2	1>2

表4 両モデルで検出された画像でのグリップエンドの中心点のデジタイズデータからの誤差(カメラ1)

	モデル1		モデル2		有意差(モデル間)	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]	x座標	y座標
試技1	35.1 ± 21.3	58.8 ± 21.7	5.9 ± 3.4	3.7 ± 2.2	1>2	1>2
試技2	44.3 ± 32.5	27.4 ± 24.2	6.4 ± 3.5	4.3 ± 3.2	1>2	1>2
試技3	57.5 ± 9.8	88.4 ± 19.0	9.2 ± 2.6	7.6 ± 3.0	1>2	1>2
試技4	45.9 ± 28.1	35.5 ± 23.7	5.7 ± 4.4	5.2 ± 3.9	1>2	1>2
試技5	100.0 ± 37.4	24.3 ± 17.8	12.9 ± 8.3	6.2 ± 2.4	1>2	1>2
試技6	41.0 ± 11.8	22.0 ± 9.8	10.8 ± 9.4	6.2 ± 4.1	1>2	1>2
試技7	35.1 ± 10.6	8.3 ± 5.4	3.6 ± 2.9	3.0 ± 2.8	1>2	1>2
試技8	73.0 ± 23.9	34.7 ± 24.7	11.5 ± 6.2	5.0 ± 2.0	1>2	1>2
試技9	103.9 ± 29.4	46.9 ± 23.8	7.3 ± 7.0	7.7 ± 3.2	1>2	1>2
試技10	76.4 ± 47.2	61.4 ± 39.2	3.8 ± 3.3	7.0 ± 5.3	1>2	1>2

表1～表4の結果から、全般的に、実験で得られた動画を学習することにより分析精度は向上したと判断できる。しかしながら、試技1に関しては、実験で得られた動画を学習することで誤差が増大してしまったことが示唆される。試技1の動画の特徴としては、バットと同系色の背景（窓枠）がバット近くにあったことが挙げられる。また、全体的に暗い動画であった。以上のことから、バットと背景との同化が精度を悪化させる大きな原因になったのではないかと推察される。

3.2 カメラ2

カメラ1と同様に、いずれの試技においても、モデル2を用いて検出できた画像の割合は、モデル1と比較して増加した（モデル1：58%±29%，モデル2：92%±20%）。それ故、両者の間には有意な差が認められた。

各試技における、バットヘッド及びグリップエンドの中心点のデジタイズデータからの誤差は、グリップエンドの試技7のx座標を除き、いずれの条件でもモデル2の方が小さい値を示した（表5, 表6）。

表5 バットヘッドの中心点のデジタイズデータからの誤差（カメラ2）

	モデル1		モデル2	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]
試技1	8.5 ± 7.4	33.8 ± 39.5	3.3 ± 3.1	3.4 ± 2.7
試技2	13.8 ± 10.3	27.0 ± 28.6	2.5 ± 2.6	2.5 ± 2.2
試技3	10.1 ± 7.3	21.9 ± 31.8	2.8 ± 3.0	2.4 ± 2.0
試技4	12.2 ± 6.0	6.1 ± 4.4	3.4 ± 2.9	2.4 ± 2.0
試技5	20.2 ± 19.7	11.5 ± 4.8	9.2 ± 6.9	7.6 ± 6.3
試技6	42.6 ± 22.9	27.0 ± 11.2	37.2 ± 37.0	20.8 ± 20.1
試技7	122.8 ± 13.3	87.6 ± 3.5	8.9 ± 7.8	5.1 ± 8.1
試技8	34.2 ± 26.6	13.6 ± 6.0	8.6 ± 8.2	4.5 ± 2.8
試技9	9.3 ± 5.4	11.7 ± 5.1	3.1 ± 2.5	2.3 ± 1.9
試技10	13.9 ± 9.4	11.8 ± 9.6	3.5 ± 2.5	3.0 ± 2.7

表6 グリップエンドの中心点のデジタイズデータからの誤差（カメラ2）

	モデル1		モデル2	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]
試技1	30.0 ± 25.2	45.5 ± 25.6	8.5 ± 7.1	15.9 ± 10.6
試技2	25.4 ± 27.9	21.9 ± 14.8	4.0 ± 3.8	3.9 ± 3.2
試技3	26.2 ± 24.7	33.9 ± 15.0	10.4 ± 9.1	13.6 ± 9.9
試技4	29.1 ± 27.0	21.2 ± 17.4	3.9 ± 2.8	3.1 ± 2.3
試技5	21.1 ± 25.2	17.7 ± 16.8	2.9 ± 2.3	1.8 ± 1.4
試技6	18.5 ± 33.3	15.0 ± 8.3	3.8 ± 3.9	2.5 ± 1.5
試技7	12.7 ± 7.5	31.3 ± 2.3	14.4 ± 13.3	3.7 ± 3.0
試技8	24.6 ± 7.1	14.3 ± 13.4	5.0 ± 5.0	5.6 ± 3.9
試技9	17.0 ± 16.9	33.6 ± 18.7	3.0 ± 2.9	4.8 ± 2.9
試技10	19.3 ± 20.1	34.0 ± 22.6	3.2 ± 4.0	5.8 ± 3.9

モデル1・モデル2の双方で検出できた画像に限定して、バットヘッド（表7）及びグリップエンド（表8）の中心点のデジタイズデータからの誤差を算出した結果、バットヘッドの試技6の両座標及び試技8のx座標（表7の灰色部分）、グリップエンドの試技7のx座標（表8の灰色部分）を除いて、モデル2の誤差はモデル1の誤差よりも有意に低い値となった。

表7 両モデルで検出された画像でのバットヘッドの中心点のデジタイズデータからの誤差（カメラ2）

	モデル1		モデル2		有意差(モデル間)	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]	x座標	y座標
試技1	8.6 ± 7.5	35.2 ± 39.8	2.9 ± 2.9	3.8 ± 2.8	1>2	1>2
試技2	14.1 ± 10.3	27.7 ± 28.7	2.0 ± 2.0	2.5 ± 2.3	1>2	1>2
試技3	10.1 ± 7.3	22.3 ± 32.1	2.1 ± 2.1	2.0 ± 1.5	1>2	1>2
試技4	12.2 ± 6.0	6.1 ± 4.4	2.9 ± 2.0	2.3 ± 1.8	1>2	1>2
試技5	20.2 ± 19.7	11.5 ± 4.8	9.9 ± 7.6	7.5 ± 6.6	1>2	1>2
試技6	42.6 ± 22.9	27.0 ± 11.2	52.9 ± 33.5	31.4 ± 17.9	1<2	1<2
試技7	122.8 ± 13.3	87.6 ± 3.5	1.7 ± 0.2	5.4 ± 1.0	1>2	1>2
試技8	25.5 ± 22.2	12.4 ± 6.0	16.0 ± 6.2	2.3 ± 1.7	—	1>2
試技9	9.3 ± 5.4	11.7 ± 5.1	2.5 ± 1.7	2.2 ± 1.5	1>2	1>2
試技10	13.9 ± 9.4	11.6 ± 9.4	3.3 ± 2.3	2.7 ± 2.2	1>2	1>2

表8 両モデルで検出された画像でのグリップエンドの中心点のデジタイズデータからの誤差（カメラ2）

	モデル1		モデル2		有意差(モデル間)	
	x座標 [px]	y座標 [px]	x座標 [px]	y座標 [px]	x座標	y座標
試技1	29.5 ± 25.6	45.4 ± 26.1	8.4 ± 7.0	19.2 ± 9.3	1>2	1>2
試技2	25.6 ± 28.2	22.3 ± 14.8	3.8 ± 3.2	4.2 ± 3.0	1>2	1>2
試技3	25.9 ± 24.8	34.0 ± 15.0	11.6 ± 9.5	16.3 ± 9.5	1>2	1>2
試技4	29.1 ± 27.0	21.2 ± 17.4	3.9 ± 2.7	3.4 ± 2.4	1>2	1>2
試技5	21.1 ± 25.2	17.7 ± 16.8	3.4 ± 2.6	1.6 ± 1.5	1>2	1>2
試技6	18.5 ± 33.3	15.0 ± 8.3	3.6 ± 3.2	2.3 ± 1.3	1>2	1>2
試技7	12.7 ± 7.5	31.3 ± 2.3	4.4 ± 6.0	1.6 ± 1.3	—	1>2
試技8	25.1 ± 8.0	17.7 ± 13.1	10.2 ± 5.6	1.5 ± 1.0	1>2	1>2
試技9	17.0 ± 16.9	33.6 ± 18.7	2.3 ± 1.9	6.1 ± 2.6	1>2	1>2
試技10	19.2 ± 20.2	34.2 ± 22.6	3.1 ± 4.0	6.1 ± 3.1	1>2	1>2

カメラ2についても、全般的に、実験で得られた動画の学習を通じて、分析精度の向上は達成できたことが示唆される。ただし、試技6に関しては、実験で得られた動画を学習することで誤差が増大してしまった。試技6の動画の特徴として、最初バットが人間の頭に隠れて分断されており、その後、バットヘッドが人間の頭の後ろに完全に隠れて見えなくなる局面を経て、再び登場することが挙げられる。同様の特徴を持つ動画はもう1試技あったが、その動画と比較して、試技6の動画はバットヘッド部分が暗く、また、バットが隠れている部分が大きかった。これらのことが、精度を悪化させる要因になったことが推察される。

3.3 本報告における研究の限界

3.1及び3.2で述べた通り、本研究を通じて、バッティング動作の撮影条件が、自動検出精度を悪化させる要因になっていることが示唆されている。今後、ディープラーニングによる座標値の自動検出を前提に動画の撮影を行うにあたり、バットの色・背景・カメラのポジション・動画の撮影条件（フレームレート・露光時間）等、検出精度向上に寄与し得る要因を洗い出す必要がある。しかしながら、本研究

の168試技の中には、国立スポーツ科学センターの屋内実験場、各チームの屋外グラウンドの双方で撮影された試技が含まれていることを考慮すると、当該手法の汎用性については十分に期待できるものだろう。

今回の報告では、学習に158試技分、効果検証に10試技分の動画を使用する形となった。これは、学習効果を高めるために、多くの動画を使用する必要性が生じ、その結果、効果検証に回せる動画が限られてしまったことに起因する。また、分析精度が向上したモデル2においても、マニュアルデジタイズデータとの違いが数px以上残されていることも事実である。現在分析中の動画が多数残されているため、今後、再度学習に使用できる動画を増やし、効果検証にも多くの動画を採用することにより、当該手法の精度向上が果たされるのか、研究を進めていきたい。

4. 結言

本研究では、野球のバッティング動作を例に取り、ディープラーニングを用いた新たな動作分析手法の確立可能性について検討することを目的とした。本研究では、ディープラーニングによる物体検出モデルとしてYOLOを用い、1)YOLOがCOCOに用意されているデータセットを学習する手法、2)YOLOが実験で得られた動画を学習する手法、の2つに手法により得られたデータを比較した。その結果、実験で得られた動画を学習教材として活用することにより、バットヘッド及びグリップエンドの中心点の自動抽出精度の向上が確認された。今後、学習動画及び効果検証動画を増やすことで、今回残された課題についても解消され、ディープラーニングを用いた動作分析手法がバッティングに対して有用であることを提言できることが期待される。

5. 謝辞

本研究を遂行するにあたり、公益財団法人ミズノスポーツ振興財団の皆様には多大なご支援をいただきました。また、ディープラーニングを実施する際には、株式会社ガラパゴスのご協力を賜りました。ここに感謝申し上げます。

参考文献

[1] Redmon J., Divvala S., Girshick R., Farhadi A.: You only look once: Unified, real-time object detection. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. 779-788, 2016.

[2] <http://cocodataset.org/#home>