

# ニューラルネットワークによる身体動作分析に関する研究 ラクロスのシューティングフォームを例として

2018SE072 小関翔也

指導教員：野呂昌満

## 1 はじめに

近年、AI技術をスポーツへ活用する動向が見られる。実際に、サッカーやラグビー、アメリカンフットボールなどのプロスポーツではICTを活用したデータによる指導が浸透し始めている。

一方で、カレッジスポーツではICTの導入が進んでおらず、ビデオカメラで撮影した動画を目視確認する程度にとどまっている。ICTの導入はスポーツそのものの専門知識だけでなく、ICTの高度な技術や知識が必要であることが原因で難しいものとなっている。さらに、ラクロスといった日本では比較的マイナースポーツではそもそもの人口が少ないため、その導入が阻害されている。

本研究では著者がラクロスのプレイヤーであったので、ラクロスにおけるICTの導入を目標とし、ラクロス選手のショットフォームを熟練度別に分類・評価するシステムを構築する。本研究の課題は、以下のとおりである。

1. データセットの作成
2. ニューラルネットワークの設計
3. VGG16の使用

## 2 既存研究ならびに関連技術

### 2.1 既存研究

金子 [1] らは、OpenPoseを用いてサッカー選手のシュートフォームから関節点を抽出し、その間接点を機械学習によって分類・評価を行った。

青木 [2] は OpenPose を用いて関節点を抽出し、抽出した関節点をニューラルネットワークに学習させ、バドミントンのショットの自動分類を行うシステムを設計した。

### 2.2 関連技術

CNN は畳み込みニューラルネットワーク (convolutional neural network) の略称である。CNN は多くのデータが局所特徴を持つという性質を利用したニューラルネットワークである。畳み込み層とプーリング層を交互に重ねることで作られている。

VGG16 は、深さが 16 層の畳み込みニューラルネットワークである。このニューラルネットワークは 100 万枚以上の画像を事前に学習させているものである。全結合層を取り除くことで、画像の特徴量を抽出することができ、事前学習済みの 1000 クラス以外の分類も可能である。

OpenPose は、カーネギーメロン大学の Cao [3] らが発表した DeepLearning を活用した姿勢推定ツールである。画像や動画を分析することで、目や鼻などの顔のパーツか

ら図 3 のように首、肩、肘、などの 18 点の間接点を抽出する。OpenPose は手や足などの人物のパーツを学習しているため、選手の一部が隠れていても、映っているボディパーツの識別により選手の特定が可能である。このことにより、不特定多数の人物の検出が可能となる。

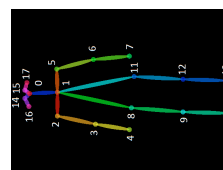


図 1 OpenPose により抽出された関節点

## 3 技術課題の解決

### 3.1 データセットの作成

本研究では南山大学男子ラクロス部のショットシーンを動画から切り抜いて使用している。上級生 (ラクロスの経験年数が 1 年以上) と 1 年生 (ラクロスの経験年数が 1 年未満) の 2 クラスの分類を目標とし、データセットを作成した。ラクロスのショットには様々な種類があるが、本研究ではスタンドシュートの打つ前のフォームの画像を使用している。シュートの角度も横からのフォームと限定した。以下の画像の一番左の画像で、データセットを作成している。

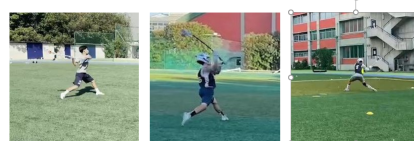


図 2 使用するデータ

### 3.2 ニューラルネットワークの設計

使用する画像は横に長く、木や別の人間など、分類に関係ないものも多く映り込んでいる。画像を判定するために余計な部分をなくすため、入力する画像を 200 × 200 の正方形とした。

ニューラルネットワークは実験した中で、最も分類できていた畳み込み層 3 層、プーリング層 3 層のものを記述する。

200 × 200 の RGB 画像を入力する、畳み込み層 3 層、プーリング層 3 層とのネットワークを設計した。前述した

データセットを用いて、上級生と1年生の2クラスを分類する。

### 3.3 VGG16の使用

自作したニューラルネットワークだけでなく、1000万枚以上の画像を事前学習しているVGG16を用いての分類も行う。224×224のRGB画像を入力する。畳み込み層が13層、全結合層が3層で構成されている。本研究では、自作のデータセットで分類を行っていくので、全結合層を取り除き、特徴量を得て、それを別の全結合層で分類を行っていく。

## 4 実験

### 4.1 プロトタイプの実装環境

プロトタイプの実装環境を以下に示す。

表1 プロトタイプ

コンポーネント	名前	バージョン
OS	Windows10 Pro	20H2
使用言語	python	3.9.7
深層学習フレームワーク	Keras	2.7.0
	tensorflow	2.7.0
描画ライブラリ	Metplotlib	3.4.3

### 4.2 ニューラルネットワークの実現

作成したデータセットを用いて分類を行っていく。学習回数を60回とし、バッチサイズは32とした。以下の図は損失関数と正答率を表したである。青色の線がトレーニングセットに対する結果で、オレンジ色の線がテストセットに対する結果である。

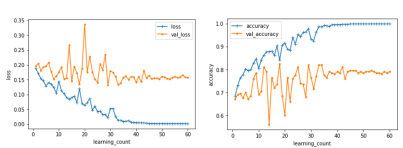


図3 実験結果

最終的には、損失値がおおよそ15%、正答率がおおよそ75%に収束していった。

### 4.3 VGG16を用いての分類

VGG16を用いて分類を行う。224×224の画像を入力しepochを30として学習を行った。以下の図が実験の結果である。

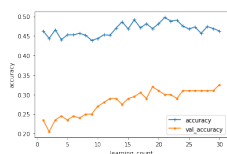


図4 実験結果

自作のニューラルネットワークよりも精度が悪く、トレーニングセットに対する正答率が50%、テストセットに対するセットに対する正答率が30%であった。

## 5 考察

自作のニューラルネットワークを使用した分類では、学習回数が30回をこえてからグラフの変化がなくなっており、過学習が起きている。VGG16を使用した分類でも、分類がうまくいっていない。原因として、

1. データの不足
  2. データの質
  3. ニューラルネットワークの構成
- の3点が考えられる。

## 6 おわりに

本研究では、CNNを用いてラクロスのショットフォームの分類を行うことを目的とした。今後は、本研究で見つかった課題を解決し、ショットフォームを熟練度別に分類できるようにしていきたい。前述の考察を踏まえ、今後の課題は以下のとおりである。

1. 学習モデルの精度向上
2. データの見直し

VGG16を使用したときよりも、自作のニューラルネットワークを使用したときのほうが、精度が高くなっている。VGG16を使用するよりも自作のニューラルネットワークを使うことが適している可能性がある。学習モデルの層の深さや数を考えていく必要がある。本研究の結果には、データの量や質が大きく関係していると考えられる。本研究のデータを集める基準として、「ショットする前のフォーム」という曖昧な条件を作ってしまった。データを収集する基準として数値を用いることでデータのラベル付けにより明確な意義を持たせることができる。データの質を高めていくという点において、本研究ではできなかったOpenPoseを用いて関節点を抽出することなどの前処理を行っていくことで分類の精度が上がる可能性がある。以上を踏まえ、データの質を向上させていく。

## 参考文献

- [1] 金子和樹, 中村拓馬, 矢入郁子, 平田均, OpenPoseを用いたサッカー熟練度の分類, (2020).
- [2] 青木義満, スポーツ競技を対象とした画像AI技術, 34巻4号, p. 497-502.
- [3] Cao, Zhe and Simon, Tomas and Wei, Shih-En and Sheikh, Yaser : Realtime multi-person 2d pose estimation using part affinity fields . Proceedings of the IEEE conference on computer vision and pattern recognition, (2017).