

3DCNN を用いたサッカー動画におけるファウル検出

2018SE005 藤澤信太郎

指導教員：宮澤元

1 はじめに

近年の機械学習の進展に伴い、機械学習による画像分類技術を用いたサッカーのファウルの判定の試みがなされている。しかし、サッカーのファウル判定はプレーの一連の流れから判断する必要があるため、一連のプレーの解析が求められる。

一連のプレーの解析に役立つのが 3-Dimensional Convolutional Neural Network (3DCNN) による動画認識技術である。3DCNN とは画像の縦横軸に加えて時間軸にも畳み込み計算を行うことにより、時間変化を考慮した機械学習を可能としたもので、動画から特徴表現を得ることができる [1]。

本研究の目的は、サッカーの試合映像からファウルの有無を判定するシステムを構築し、それを用いて学習データの違いが与えた影響を検出することである。対象とするファウルは、サッカーの試合の中で最も多い「足をかけた場合のファウル」とし、過去の試合映像からファウルの疑いがあったシーンを抽出して機械学習させることにより、ファウルの有無を自動判定する。

本研究では、動画から切り分けた画像に対して下半身検出を前処理として行ったデータを、3DCNN を用いて機械学習することでサッカーのファウルを判定する方法を提案する。下半身検出という工夫を加えることで、下半身に注目した学習を行うことができる。

提案手法の効果を検討するために実験を行う。前処理として下半身検出を行った学習用データと下半身検出を行わない学習用データ、それぞれを 3DCNN の入力として学習し、2 種類のモデルと重みを作成する。次に、それぞれのモデルと重みを使ったテストを行い、得られた判定精度を比べることで、下半身検出が与えた違いを検出し、より一層の判定精度の向上に向けた考察を行う。

研究課題は以下の 3 点である。

- 3DCNN を用いたサッカーのファウルを自動判定するシステムを構築する
- 作成したシステムの判定精度を確認する評価実験を行う
- 判定精度の向上に向けた考察を行う

2 研究の背景

本節では、研究の背景としてサッカーにおける機械学習の活用と 3DCNN について述べる。

2.1 サッカーにおける機械学習の活用

スポーツの多くは競技結果を始め、多くの要素が数値として算出されるので、膨大なデータを取り扱うことを得意とし、人間では発見が難しい特徴を抽出することに長けた機械学習との親和性が高い [2]。

しかし、サッカーのファウル判定での機械学習の活用はうまくいっておらず、機械学習による画像分類技術を用いたファウルの判定の試みがなされている。近藤らは機械学習による画像分類技術を用いたファウルの判定で、学習画像データの増しによる影響を論じている [3]。しかし、この手法では時間軸方向を意識しない手法になるので、サッカーのファウル検出には向かない。

2.2 3DCNN

3DCNN とは、Shuiwang らによって提案された、Convolutional Neural Network (CNN) における畳み込み層を 3 次元に拡張したものである [1]。畳み込み層を 3 次元に拡張することにより、画像の縦横軸方向に加え、時間軸方向に対しての畳み込み計算も可能になり、時間軸を持つ動画に対しての時間変化を捉えた特徴の抽出を可能にした。

3 提案手法

本研究では、3DCNN を用いたサッカーにおけるファウル検出システムを提案する。

提案手法の手順は、まず入力する学習データの動画に対し、動画の前処理を行い、動画の前処理で得られた画像群で 3DCNN を用いた学習を行い、判定で用いるためのモデルと重みを作成する。

3.1 動画の前処理

1 つ目の前処理として入力した動画を 0.1 秒毎の画像に切り分ける。3DCNN での学習データの inputs は、一定時間内の連続につながった画像群になるので、この処理を行う必要がある。

2 つ目の前処理として下半身検出を行う。あらかじめ準備しておいた機械学習の結果データを参照し、入力した画像から下半身を検出する。

3.2 3DCNN を用いた学習

本研究で提案するネットワーク構造を図 1 に示す。サッカーのファウルという複雑な分類を行うため、3 次元畳み込み層とプーリング層を 4 層繰り返す、複雑な分類に対応する。

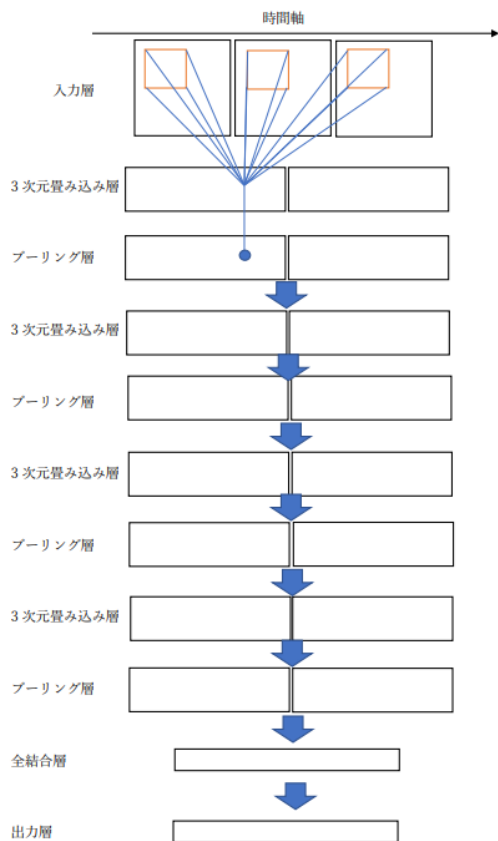


図1 提案手法のネットワーク構造

4 実験

提案手法を OpenCV, keras などのライブラリを用いて実装し、サッカーのファウルの判定実験を行った。下半身検出あり、下半身検出なし、それぞれの学習用データを 3DCNN の入力として学習し、2 種類のモデルと重みを作成し、それらを使ったテストを行い、得られた判定精度を比べる。

本研究で用いるデータは、動画サイト YouTube に公開されている過去の試合のプレー動画とする。データ数は、下半身検出あり、下半身検出なしの学習用データがそれぞれ 100 本、テストデータが 10 本とする。

4.1 実験結果

バッチサイズは 20、エポック数は 200 で学習を行った。

下半身検出ありの学習結果は、最終的な損失関数値はおよそ 0.693、精度は 50.0 % となった。また、エポック数に対する損失関数値と精度の推移を図 2 に示す。今回の学習ではエポック数に応じた精度の向上が見られず、正確なモデルの作成はできなかったと言える。

下半身検出なしの学習結果は、最終的な損失関数値はおよそ 4.45、精度は 50.0 % となった。また、エポック数に対する損失関数値と精度の推移を図 3 に示す。今回の学習では損失関数値が後半から上昇し、予測と実際の値のずれが大きくなってしまっており、正確なモデルの作成はできな

かったと言える。

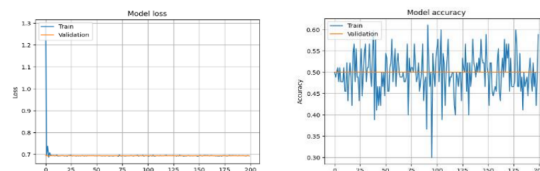


図2 下半身検出ありの学習結果

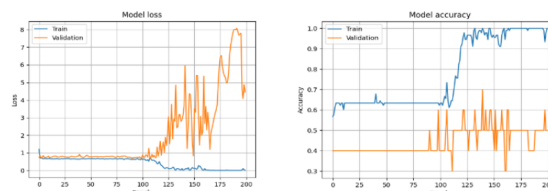


図3 下半身検出なしの学習結果

テストデータによる判定精度については、下半身検出ありの場合が 50.0 %、下半身検出なしの場合は 20.0 % となった。

4.2 考察

今回の実験から、下半身検出は判定精度の向上に繋がったと言える。しかし今回の結果はデータ数の少なさ、モデルの正確性といった問題を解決することができなかった点を考慮する必要がある。

5 おわりに

本研究では、3DCNN を用いてサッカーのファウルを判定する方法を提案し、それをもとに実験を行い、サッカーのファウル検出に対する下半身検出の有無が与える影響を検出した。

サッカーのファウルという複雑な問題に対してのデータ数の少なさ、モデルの正確性といった問題を解決することができなかったが、今回の結果から下半身検出はファウル検出に良い影響を与えたと考えられる。

今後の課題は、足や靴などより細かい体の部位がサッカーのファウル検出に与える影響を検出し、より精度の高いファウル判定を実現することである。

参考文献

- [1] Ji Shuiwang et al: “3d convolutional neural networks for human action recognition”, Pattern Analysis and Machine Intelligence, vol.35 (2013).
- [2] Laboro: 試合, トレーニング, 観戦まで. 進むスポーツ業界の AI 活用, <https://laboro.ai/activity/column/laboro/sports-ai/>
- [3] 近藤 涼月 他: “画像によるサッカーのファウル判別における水増し手法の影響”, 電子情報通信学会 (2019).