

アメリカンフットボールにおける オフライン選手の動作認識

Motion Recognition of an Offensive Line Player in American Football

西村 宏武 岡 夏樹 早川 博章 廣田 敦士
Hirotake NISHIMURA Natsuki OKA Hirofumi HAYAKAWA Atsushi HIROTA

京都工芸繊維大学 大学院工芸科学研究科
Graduate School of Science and Technology, Kyoto Institute of Technology

Toward play recognition of American football, this paper proposes a motion recognition model of a single offensive line player using a 3D convolutional neural network. The model successfully learned to recognize the motions of OL players by simultaneously learning both a simple situation of only one offensive line (OL) player and a more complicated situation of an OL player confronting a defensive line (DL) player. For the real setting, it is necessary to recognize the motion of OL players wearing an arbitrary uniform. In our evaluation experiment, even though the OL player wore a uniform not included in the learning data, the model achieved high recognition accuracy of 98.6% in the situation of only OL player, and 95.2% in the situation of an OL player confronting a DL player.

1. はじめに

アメリカンフットボールは格闘技と球技の要素に加え、深い戦略性を併せ持つスポーツである。繰り広げられるパワーとテクニックの応酬に世界中の多くのファンが魅了されている [藤田 14]。選手はそれぞれ与えられた役割を果たし、緻密な作戦を遂行させ、プレイの成功を目指す。そのため、選手個々の体力や技術だけでなく、プレイ・戦術も非常に重要となるスポーツである。1回のプレイに参加する選手はオフェンスチーム、ディフェンスチームともに11人であり、オフェンスチームは多様なフォーメーションから多様なプレイを繰り出す。選手は様々なプレイを理解・記憶したうえで試合に臨む。試合中のオフェンスプレイは、一般的にオフェンスコーディネーターが決定し、選手に伝達する。

試合の戦術を立てるために、アメリカンフットボールでは練習に加えてスカウティングを行う。スカウティングは、相手を分析し戦術を組み立てる上で柱となる重要な作業である。分析対象チームの戦況対応、フォーメーション別の得意なプレイを細かく分類し分析することで、自チームの作戦を立てる。分析対象チームの遂行した作戦の記録が必要となるが、これは非常に時間のかかる作業である。

作戦記録の作業時間を短縮するため、コンピュータに作戦記録の作業を任せることは有効である。本研究の目的は、図1に示すオフラインにおけるポジションの1つであるオフライン選手の動作認識とする。オフラインはプレイを成功させるために最も重要なポジションであり、プレイを細かく分類する際に大きな参考となる [Bass 12]。また、オフライン選手に対峙するディフェンスライン選手を加えた、より複雑な状況でのオフライン選手の動作認識も行う。以下、表中ではオフラインをOL、ディフェンスラインをDLと略記する。

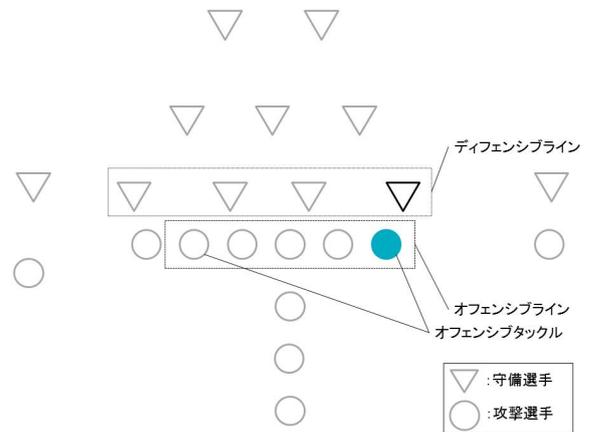


図 1: オフェンシブタックルの位置

2. 実験

2.1 データ

認識対象の動作を図2~7に示す*1。オフライン選手に対して、各図中の上一列は正面5.5メートル、下一列は横5.5メートルから撮影したものである。また、オフライン選手目線で右側(各図中上一列では左側)をフィールドの外側と定義した。動画データは固定ビデオカメラから60fpsで撮影した。本研究で動作を記録した選手7人の内、6人の各動作を学習データとし、1人の各動作をテストデータとした。オフライン選手の場合、ディフェンスライン選手に対峙する場合の2通りをそれぞれ撮影した。

連絡先: 〒606-8585 京都市左京区松ヶ崎橋上町1
京都工芸繊維大学 大学院工芸科学研究科 情報工学専攻
インタラクティブ知能研究室, nishimura@ii.is.kit.ac.jp

*1 オフラインは各図中の上一列では奥、下一列では左の選手。ディフェンスラインは上一列では手前、下一列では右の選手。



図 2: ドライブブロック



図 3: カットオフブロック



図 4: インサイドリーチブロック



図 5: パスプロテクション



図 6: アウトサイドリーチブロック



図 7: スライドブロック

各動作の内容を以下に示す。

- 図 2: ドライブブロック. ランプレイの内, 主にオフENSイブライン選手に対してフィールドの内側のプレイで使われる. オフENSイブライン選手はブロック対象の正面に踏み出しブロックする.
- 図 3: カットオフブロック. ランプレイの内, 主にオフENSイブライン選手に対してフィールドの外側のプレイで使われる. オフENSイブライン選手は, 基本的にブロック対象の外肩を目指して脚を外に踏み出しブロックする.
- 図 4: インサイドリーチブロック. ランプレイの内, 主にオフENSイブライン選手に対してフィールドの内側のプレイで使われる. オフENSイブライン選手は, 基本的に正面のディフェンシブライン選手から内側に向かって離れ, 一線奥にいるブロック対象へブロックする.
- 図 5: パスプロテクション. 主にパスプレイで使われる. オフENSイブライン選手は, 基本的には上体を地面と鉛直になるように起こし, パスを投げる攻撃選手を守る.
- 図 6: アウトサイドリーチブロック. ランプレイのうち, オフENSイブライン選手に対してフィールドの外側に走るプレイで使われる. オフENSイブライン選手は, 基本的に正面のディフェンシブライン選手から外側に向かって離れ, 一線奥にいるブロック対象をブロックする.
- 図 7: スライドブロック. ラン・パスプレイともに使われ, 基本的にオフENSイブライン選手にとって内側に1歩詰め, その後は後方に下がる.

学習データは, オフENSイブライン選手を正面から撮影した動画データを扱う. 各データの先頭を0フレーム目, 1つ過去のフレームを2フレーム目とすると, 抽出した画像列は0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30フレーム目の計16フレームを1つのデータとして扱った. 学習データ不足を解決するため, 図8に示すように, 動画データから動作開始後30~0フレーム, 動作開始後31~1フレーム, というように抽出箇所を1フレームずつずらしデータを抽出した. 1つの動画データから動作開始30~59フレーム目を抽出開始点として, 30列の学習・テストデータを作成した. 1つの動作は1秒程度で完結するため, 動作開始から59フレーム目(凡そ1秒後)を1つの動画データの最後尾とした. 動作開始後30フレームからデータを作成した理由は, 動作認識のために入力画像列の長さを十分に取る必要があることと, 動画を1フレームずつ見たとき, 動作開始から0.5秒(30フレーム)後までが著者の経験上動作を認識するために十分だからである. 加えて, 複数異なる動作を部分的に組み合わせた動作を認識するためにも, 動作の開始から終了までを1つのデータとして捉えるのではなく, 動作の時間的な部分(本研究では0.5秒間ごと)でも認識する必要があるためである. 各画像列は元サイズ1000×1000ピクセル, これを128×128ピクセルに圧縮した. 各々の学習・テストデータ数を表1に示す.

表 1: 学習・テストデータ数

	学習データ数	テストデータ数
OL 選手のみ	4680	900
OL 選手と DL 選手	4770	630

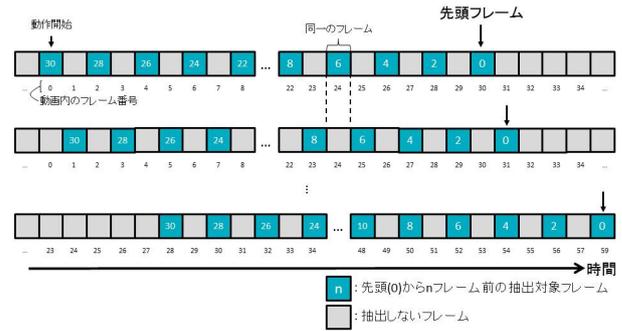


図 8: 動画データの抽出方法

2.2 学習

学習モデルは三次元畳み込みニューラルネットワークを用いた. モデルの詳細を表2に示す. conv は畳み込み層, pool はプーリング層, fc は全結合層である. 最初の畳み込み層 conv1 では, パッチサイズを4@4×4(時間方向に4, 縦横に4×4)とし, ストライドを2@2×2とした. これにより conv1 ではフレーム間を畳み込むことによって時間上の差異を抽出するようにした. pool1~conv4 はパッチサイズを1@3×3, ストライドを1@2×2とした. 全結合ユニットをfc1では256とし, fc6では識別するカテゴリ数6と一致させた. 過学習を避けるために [Srivastava 14], fc5ではドロップアウト率を0.500として学習した. 学習率は 1.00×10^{-5} とした. また, 活性化関数は conv1, conv2, conv3, conv4, fc1では Rectified Linear Unit(ReLU, ランプ関数)を用い [LeCun 15], 出力層 fc6では softmax 関数を用いた.

表 2: 学習モデル

layer	patch	stride	map	activation function
input	-	-	16@128×128×3	-
conv1	4@4×4	2@2×2	7@63×63×48	ReLU
pool1	1@3×3	1@2×2	7@31×31×48	-
conv2	1@3×3	1@2×2	7@15×15×192	ReLU
pool2	1@3×3	1@2×2	7@7×7×192	-
conv3	1@3×3	1@2×2	7@3×3×384	ReLU
conv4	1@3×3	1@2×2	7@1×1×1024	ReLU
fc5	-	-	256	ReLU
fc6	-	-	6	softmax

入力は128×128×3の16フレーム分の画像列, 出力は6通りの各動作のカテゴリ尤度とした. また, 誤差関数は交差エントロピー誤差を用いた. 5通りのモデルの学習方法を以下に示す.

- O : オフENSイブライン選手のみでの動作を学習したモデル.
- D : ディフェンシブライン選手に対峙した状況でのオフENSイブライン選手の動作を学習したモデル.
- OtoD : モデルOの学習後, ディフェンシブライン選手に対峙した状況でのオフENSイブライン選手の動作を学習したモデル.
- DtoO : モデルDの学習後, オフENSイブライン選手のみでの動作を学習したモデル.

MIX : オフェンシブライン選手のための動作、ディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作をシャッフルして同時に学習したモデル。

3. 結果と考察

各モデルの認識精度を表3に示す。モデルOではオフェンシブライン選手のための動作で認識精度が97.0%を超えた一方、ディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作認識では43.5%となった。モデルDはディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作で認識精度が97.0%を超えた一方、オフェンシブライン選手のための動作認識で57.0%となった。モデルOの学習後に他方の状況を学習したモデルOtoDでは、オフェンシブライン選手のための動作認識精度が61.8%、ディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作認識精度が94.3%となった。モデルDの学習後に他方の状況を学習したモデルDtoOでは、オフェンシブライン選手のための動作認識精度が98.7%、ディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作認識精度が63.5%となった。モデルMIXでは、オフェンシブライン選手のための動作認識精度が98.6%、ディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作認識精度が95.2%となった。

モデルOとモデルDtoOでは、オフェンシブライン選手のための動作を、またモデルDとモデルOtoDではディフェンシブライン選手に対峙した状況でのオフェンシブライン選手の動作を、それぞれ高い精度で認識することができた。モデルMIXはオフェンシブライン選手のための状況と、ディフェンシブライン選手に対峙した状況の、2つの状況下で実際にオフェンシブライン選手の動作を共通して取り出すことができる、とは言いきれない。そこで、実験的な根拠を付けるためにアテンションメカニズム[Vinyals 16]を導入し、2つの状況で共通してオフェンシブライン選手の動作を取り出すことができているかの分析を計画している。

表 3: 認識精度 [%]

	OL 選手のみ	DL 選手に対峙した OL 選手
O	97.3	43.5
D	57.0	97.6
OtoD	61.8	94.3
DtoO	98.7	63.5
MIX	98.6	95.2

4. まとめ

アメリカンフットボールのプレイ認識に先立って、三次元畳込みニューラルネットワークを用いたオフェンシブライン選手の動作認識を行った。ランダムに混合して同時に学習することで、2つの状況下での動作認識ができた。また、スカウティングでは任意のユニフォームを着ている選手から動作を認識する必要があるが、本研究では学習データに含まれていないユニフォーム及び選手でも高い認識精度(オフェンシブライン選手のための状況で98.6%、ディフェンシブライン選手に対峙した状況で95.2%)を出すことができた。

実際の試合では、各プレイ毎に全ての選手の位置が変わることが多いため、学習には様々な角度から撮影したデータを用いる必要がある。さらには、ディフェンシブライン選手が独自の

動きをする場合や、他のディフェンス選手、オフェンス選手がいる場合におけるオフェンシブライン選手の動作認識が必要になる。この点では、ディフェンス選手目線でオフェンシブライン選手を追った、カメラを固定せずに撮影した動画データを用いて学習する必要がある。また一方で、フィールド内に22人いる選手の中からオフェンシブライン選手を見つけることも課題となる。本研究ではディフェンシブライン選手へ対峙する・しない状況共に95%を超える高い精度を出すことができたが、その2つの状況に影響されずに、オフェンシブライン選手を注目して動作を取り出しているのかを今後の課題とする。

参考文献

- [Bass 12] Bass, T.: *FOOTBALL SKILLS & DRILLS*, Human Kinetics, Champaign, Illinois (2012)
- [LeCun 15] LeCun, Y., Bengio, Y., and Geoffrey Hinton, : Deep learning, *Nature*, Vol. 521, No. 7553, pp. 436–444 (2015)
- [Srivastava 14] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research*, Vol. 15, pp. 1929–1958 (2014)
- [Vinyals 16] Vinyals, O., Toshev, A., Bengio, S., and Erhan, D.: Show and Tell: Lessons learned from the 2015 MSCOCO Image Captioning Challenge, *CoRR* (2016)
- [山西 15] 山西 宏平: 深層学習に基づく行動に着目したシーン抽出手法, Master's thesis, 大阪大学基礎工学部情報科学科 (2015)
- [藤田 14] 藤田 智: アメリカンフットボール最強の戦術論, メイツ出版, 東京都 (2014)