

サッカートラッキングデータを用いた 機械学習に基づくプレー認識手法の提案

今井 友揮^{1,a)} 内山 彰^{1,b)} 馬込 卓弥^{2,c)} 東野 輝夫^{1,d)}

概要: 近年、戦術分析や選手育成などの目的で、スポーツにおけるデータ解析が注目を集めている。このため、日本プロサッカーリーグ（Jリーグ）においては、専門の記録者が手作業でパス、ドリブルなどのボールタッチプレーの記録を行っているが、非常に負担が大きく、プロの中でもごく一部の試合に限られている。そこで本研究では、低コストなプレーのラベル付けを実現するため、ボールタッチプレーの認識手法を提案する。提案手法では、GPS やカメラにより各選手とボールの移動軌跡が記録されたトラッキングデータを用いる。まず、ボールの移動軌跡の変化に着目することで、各プレーが発生したタイミングを検出する。次に、検出したタイミングにおいて、選手とボールの位置関係から特徴量を抽出し、機械学習を用いてプレーの種類を認識する。提案手法の有効性を評価するために、実際に J リーグで記録された 20 試合分のトラッキングデータを用いて実験を行った。その結果、プレー検出では、適合率 86.7%、再現率 85.3%を達成できることが分かった。また、プレー認識においては、CRF, SVM, Random Forest の機械学習アルゴリズムを比較した結果、Random Forest が平均的に高い性能であり、F 値 54.7%となることが確認できた。特に、試合中に発生するプレーの約 90%を占めるパスとトラップについては、それぞれ 86.0%、85.9%の F 値を達成し、低コストなプレーのラベル付けが可能なが分かった。

1. はじめに

近年、スポーツにおけるプレーデータは情報技術の発展に伴い重要度を増してきている。例えば、サッカーにおいては試合に関する統計情報として、シュートやコーナーキックの回数が試合ごとに記録されている。日本プロサッカーリーグ（Jリーグ）においては、戦術分析に活用するため、パス、ドリブル、ブロック等のボールタッチプレーについて、回数や成否を記録している。これらのプレーデータは、現状では訓練を受けた記録者が試合映像を目視で確認しながら記録付けを行っているが、熟練した記録者であっても 1 試合あたり 10 時間程度を要し、負担が大きい。したがって、プレーデータの記録は J リーグの中でも 1 部リーグの試合に限られている。これに対し、ごく一部のプロ選手だけでなく、高校生などの試合においてもプレーデータが記録できるようになれば、戦術分析のみならず、ビッグデータ解析による選手の早期育成やケガのリスク低減などの様々な活用が期待される。このためには、低コストでプ

レーデータを記録する方法が求められる。

これに対して、映像を用いたスポーツ情報解析が注目を集めている [1]。近年、発展が目覚ましい深層学習を用いたプレー認識手法が提案されているが、十分な性能が得られているとはいえないのが現状である [2], [3], [4]。このため、映像だけでなく様々なセンサ情報を組み合わせることが精度向上のためには重要と考えられる。ウェアラブルセンサの加速度、角速度データ、および選手の位置情報を組み合わせ、ラグビーにおけるプレー認識を行う手法も考案されている [5]。しかし、対象はタックルとスクラムの 2 種類に限定されており、様々なスポーツにおけるプレー認識においては依然として課題が残されている。

そこで、本研究ではサッカーを対象として、選手とボールの位置情報に着目した低コストなプレー認識手法を提案する。認識の対象は、パス、トラップ等といった、実際に J1 リーグの全試合で記録されているボールタッチを伴うプレーである。選手とボールの位置情報は GPS やカメラ等により得られることを想定する。特に、Catapult Sports 社の OptimEye S5[6] などの GPS 内蔵型ウェアラブルセンサはスポーツ向けに普及が進んでおり、adidas 社の micoach smart ball[7] のようなセンサを内蔵したボールも販売されている。こういったセンサはカメラと異なり、専門家による設置位置などの調整が不要なため、高校のサッカー部な

¹ 大阪大学大学院情報科学研究科

² 大阪大学大学院医学系研究科

a) t-imai@ist.osaka-u.ac.jp

b) uchiyama@ist.osaka-u.ac.jp

c) t-magome@anat2.med.osaka-u.ac.jp

d) higashino@ist.osaka-u.ac.jp

どでも容易に運用可能と考えられる。

提案手法では、まず、プレーが発生した瞬間にボールの軌道もしくは移動速度が一定以上変化するという点に着目し、トラッキングデータ内においてボールタッチプレーが発生したタイミングをルールベースで検出する。次に、検出したタイミングごとに選手とボールの移動軌跡から特徴量を抽出し、機械学習を用いてプレーの種類を認識する。機械学習には、条件付き確率場 (CRF)、サポートベクターマシン (SVM)、Random Forest (RF) の 3 種類を適用した。

本研究の位置づけは、以下の通りである。

- (1) サッカーを対象として実際に J リーグで記録されている 10 種類のボールタッチプレーを低コストに認識する手法を提案している。
- (2) CRF, SVM, RF の 3 種類の機械学習アルゴリズムを比較し、実際の J リーグの試合で記録されたデータで評価を行っている。

性能評価のため、実際に J リーグの試合で記録されたトラッキングデータを用いて実験を行った。その結果、プレー検出手法では適合率が 86.7%、再現率が 85.3% となり、プレー認識においては、平均的には RF が最も F 値が高く、54.7% であった。一方、ドリブルについては CRF が最も F 値が高くなり、プレーの時間的な前後関係を考慮することが有効な場合もあることが分かった。特に、試合中に発生するプレーの約 90% を占めるパスとトラップについては、RF によりそれぞれ F 値 86.0%、85.9% となり、低コストなプレーのラベル付けが可能なことを確認した。

2. 関連研究

2.1 スポーツプレー認識

スポーツにおけるトラッキングデータは様々な活用が検討されている。文献 [8] では、バスケットボールを対象に、機械学習を用いることで、複数選手の動きからボール保持選手を推定する手法を提案している。また、文献 [9], [10] では、トラッキングデータを用いて、発生したチームアクティビティを識別する手法を提案している。これらの手法では、少数のチームアクティビティを高精度に認識できるが、サッカーやラグビーにおけるタックルといった、選手同士が干渉することで発生するプレーに関しては認識ができない。文献 [11] では、サッカーにおいてプレー認識を行うために、本研究と同様に選手とボールの位置データの特徴量として用いた機械学習による手法を提案しているが、評価に用いるデータ量が 10 分程度しかなく、十分な検討がされていない。

近年では深層学習の発展に伴い、映像を利用したプレー認識も提案されている。文献 [12] では、アメリカンフットボールを対象に、オフENSE、ディフェンス、フィールドゴールといったプレーを認識している。また、文献 [13] で

は、アメリカンフットボールにおけるショートパス、ロングパス、パントなど 7 種類のプレーを認識するために、試合映像から抽出したカメラの動きのパラメータを利用した手法を提案している。文献 [3] では、フットサルにおけるパス、シュート、ドリブルを対象として、複数台のカメラによる映像を用いて畳み込みニューラルネットワークによるプレー認識を提案している。しかし、これらの映像を利用したプレー認識手法だけでは十分な認識精度が達成できておらず、映像以外の情報を併用することが必要不可欠である。

一方、ウェアラブルセンサデータに基づくプレー認識に関する研究も存在する。文献 [14] では、サッカーにおいて足首に装着したウェアラブルセンサの加速度、角速度データを用いてボールインパクトの瞬間を検出することで、キックの発生を認識している。文献 [5] では、ウェアラブルセンサの加速度、角速度、および選手の位置情報を組み合わせ、ラグビーにおけるタックル、スクラムといった 2 種類のプレーを認識する手法を提案している。しかし、これらの手法では、認識対象が大きなプレーに限られており、詳細なプレー認識は検討されていない。これに対して、本研究では、実際に J リーグで記録されている 10 種類のボールタッチプレーを対象にプレー認識を行う手法を提案する。

2.2 スポーツ情報解析

近年、映像を用いたスポーツ情報解析が注目を集めている [1]。その多くは選手やボールの位置を追跡するトラッキングが中心であるが [15]、例えば、ラグビー、テニス、サッカーなどの様々なスポーツ映像に対して、選手とボールの位置を追跡したり [15]、ホイッスル音を認識してプレーシーンの区切りを自動ラベリングするといった試みが成されている [16]。また、文献 [17] では、J リーグの試合で計測されたトラッキングデータ、プレーのラベルおよび中継映像を用いて、ゴールが起こった直前に発生したプレー群に着目し、チャンス、ピンチを作り出した要因を導出する研究を行っている。文献 [18] では、トラッキングデータとパスが発生した時刻のリストから、パスの質を“良い、普通、悪い”の 3 段階に自動で評価する手法を提案している。しかし、これらのスポーツ情報解析では、全て手動でプレーのラベル付けを行う必要があり、分析データの準備に大きな負担がかかるという問題がある。本研究でスポーツにおけるプレーラベル付けの自動化が実現できれば、このようなスポーツ情報解析に活用することができる。

3. 提案手法

3.1 概要

提案手法の概要を図 1 に示す。まず、ボールタッチプレー発生時にはボールの軌道や速度が変化することに着目

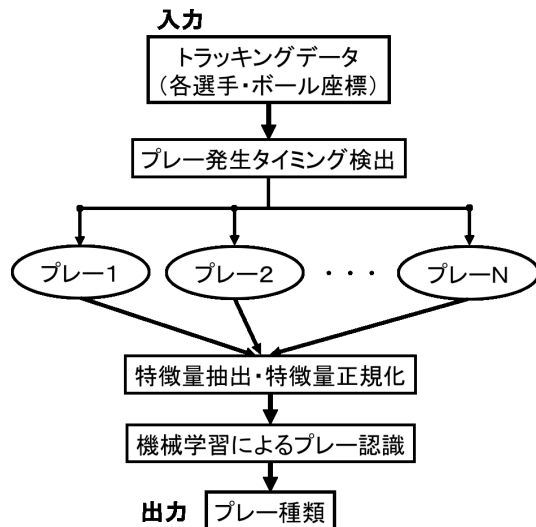


図 1 提案手法の概要

表 1 対象プレーのクラス

クラス名	記号	内容
パス	pa	味方選手の足元にボールを出す
スルーパス	tp	相手最終ラインの裏のスペースにボールを出す
クロス	cr	敵陣サイドからペナルティエリア内の味方へボールを出す
シュート	sh	相手ゴールを狙ってボールを出す
フィード	fe	キーパーが保持しているボールを前方へ出す
クリア	cl	守備側の選手が自陣でボールを大きく前方へ出す
トラップ	tr	受けたボールを足元に収める
ドリブル	dr	ボールを運んで相手選手をかわそうする
キャッチ	ca	キーパーがボールを抱える, または手で掴む
ブロック	bl	相手選手が出したボールに体の一部を当てて進路を変える

し, 各プレーが発生したタイミングをトラッキングデータより検出する. 次に, 検出した各タイミングにおいて, トラッキングデータから得られるボールと各選手の空間的関係より特徴量を抽出し, 機械学習によりプレーを認識する. 機械学習には, CRF, SVM, RF の 3 種類を用いる. 対象となるプレーは表 1 に示す通りである.

3.2 プレー発生検出

プレー発生検出のため, タイムウィンドウ W に対して, ウィンドウの開始時 $t-W$ と終了時 t におけるボール座標をそれぞれ始点, 終点としたベクトル v_b^{t-W} , v_b^t を考える. このとき, 以下の式 (1), (2) のどちらかを満たす場合にプレーであるとして検出する.

$$\frac{v_b^{t-W} \cdot v_b^t}{|v_b^{t-W}| |v_b^t|} \leq \cos \theta_{TH} \quad (1)$$

$$|v_b^{t-W}| - |v_b^t| \geq V_{TH} \quad (2)$$

ここで, θ_{TH} , V_{TH} はそれぞれ角度変化, 移動距離の閾値である.

また, プレーでないにも関わらずボールのベクトルが変化するシーンの一部について, 誤ってプレーとして検出しないよう工夫を施した. 例えば, ゴールキーパーがボール

表 2 特徴量の候補

種別	値
位置	b_0, b_{-1}, b_1
	t_i における両チーム最終ラインの位置
	b_0 から 3m 以内の距離に位置する選手の数
距離	$D(p_0, p_1), D(p_0, p_{0+s}), D(p_0, p_{-1})$
	$D(q_0, q_1), D(q_0, q_{-1})$
	$D(b_0, b_1), D(b_0, b_{0+s}), D(b_0, b_{-1})$
	$D(p_1, b_1), D(p_{-1}, b_{-1}), D(p_{0+s}, b_{0+s})$
	$D(p_0, p'_0)$
	$D(gk_0, b_0), D(gk'_0, b_0)$
方向	ボールの移動方向
	ゴール接近度
時間	$t_i - t_{i-1}, t_{i+1} - t_i$
状況	t_i, t_{i-1}, t_{i+1} それぞれにおけるボール保持チーム
	2 値分類の結果

をキャッチした後, 一定時間ボールを保持したまま移動し, 最終的にフィードを行うシーンを想定する. このとき, 検出したいプレー発生タイミングは, キャッチが行われた瞬間とフィードが行われた瞬間のみであるが, ボールを保持したまま移動した区間においてもボールのベクトルが変化するため, プレーでないにも関わらずプレーとして誤検出される可能性がある. そのため, ボールの最も近くに存在する選手をウィンドウ内におけるボール保持選手として仮定し, ボール保持選手が同一である区間に発生したボールベクトルの変化に関してはプレーとして検出しないようにした.

3.3 特徴量選択

検出された各プレー発生時刻 t_i において, 選手, ボールの座標を基に空間的な特徴量候補を定義する. 特徴量候補を表 2 に示す. t_i に対して, 直前のプレーが発生した時刻を t_{i-1} , 直後のプレーが発生した時刻を t_{i+1} とする. 以降では, プレー発生時刻においてボールと最も距離が近い選手をプレーの実行選手とみなす. t_i に発生したプレーの実行選手を p , t_{i+1} に発生したプレーの実行選手を q と表記する. t_i における p の位置を p_0 とし, この選手の t_{i-1} 時点における位置を p_{-1} , t_{i+1} 時点における位置を p_1 とする. p' は p に対して最もボールに近い相手チーム選手の位置であり, b はボールの位置を表す. また, 両チームのゴールキーパー選手の位置をそれぞれ gk, gk' と表す. 最終ラインの位置は, チームの選手中, ゴールキーパーを除いて最も後方 (自陣側) に位置する選手の位置を意味する. 位置は, x, y 座標がピッチ中央を原点とした座標, z 座標が地面を基準とした高さを表す 3 次元座標 (x, y, z) で定義される.

ボールや選手間の距離は最も基本的な特徴量と考えられる. このため, 2 つの対象 O_1, O_2 間の距離を $D(O_1, O_2)$ で表す. 対象 O は, 選手, ボール, 最終ライン, ゴールキー

パーのうちいずれかである。また、ボールの移動方向も重要な特徴量であるため、 b_0, b_{0+1} の 2 点を通る直線が x 軸と成す角度により定義した。さらに、シュートなどのプレーはボールがゴール方向に向かう傾向が強いため、これを反映する特徴量として、ゴール接近度を定義する。ゴール接近度は、 b_0, b_{0+1} の 2 点を通る直線に対して、 b_{0+1} に近い側のゴールラインとの交点の y 座標で定義される。これらに加えて、プレー発生から一定時間後の選手やボールの位置は瞬間的な移動速度を反映するため、 t_i から s 秒後の時刻 $t_i + s$ における p および b の位置 p_{0+s} および b_{0+s} もそれぞれ特徴量として用いた。 s の値は、プレーによって異なると考えられるため、経験的に 0.3, 0.6, 1.0 の 3 種類を用いた。

また、認識対象プレーは“ボール移動系”と“ボール保持系”の 2 種類に大別できるため、決定木で得られた 2 値分類の結果もプレー認識の特徴量候補とした。“ボール移動系”とは、プレー発生後にボールがプレー選手から離れる傾向のあるプレー群を指し、パス、スルーパス、クロス、シュート、フィード、クリア、ブロックの 7 種類が含まれる。“ボール保持系”とは、ボールを保持するプレー群であり、トラップ、ドリブル、キャッチの 3 種類のプレーで構成される。決定木による 2 値分類に用いた特徴量には、 $D(p_i, b_i)$ と、 p と q が同一かどうかを表すバイナリ値の 2 つを用いた。

表 2 で示した以上の特徴量候補を各プレー発生タイミングごとに抽出し、特徴ベクトルを生成する。このとき、特徴ベクトルの各成分はそれぞれスケールが異なり等価に扱うことができないため、各成分の最小値が 0、最大値が 1 となるように正規化を行った [19]。

提案手法では、特徴量の一部として“直前/直後のプレー発生時”における情報も使用している。この理由は、サッカーにおいては時系列前後のボール、選手の位置状況によって発生しやすいプレーが異なると考えられるためである。例えば、ボールが移動する系統のプレーに関しては発生タイミング以降のボールの位置情報に特徴が大きく表れるが、ボールを保持する系統のプレーに関しては発生タイミング以前のボールの位置情報に特徴が大きく表れると考えられる。しかし、直前もしくは直後のプレーが存在しない場合には、これらの特徴量が存在しないため、本研究の認識対象からは除外するものとした。除外したプレーに対しては、SVM などのプレーの前後関係を利用しない機械学習アルゴリズムにより、個別に認識手法を設計する必要がある。

特徴量選択のため、ランダムフォレストを用いて認識における各特徴量の重要度を算出し、変数減少法 [20] を適用した [21]。本研究では、ランダムフォレストを用いて正解率の低下が 1% 以上となった時点で特徴量の除去を停止するものとした。最終的に、表 3 で示す 33 個の特徴量を以

表 3 最終的に用いた特徴量

種別	値
位置	b_0, b_{-1}, b_1
	t_i における両チーム最終ラインの位置
距離	$D(p_0, p_1), D(p_0, p_{-1})$
	$D(q_0, q_1), D(q_0, q_{-1})$
	$D(b_0, b_1), D(b_0, b_{0+s}), D(b_0, b_{-1})$
	$D(p_1, b_1), D(p_{-1}, b_{-1}), D(p_{0+s}, b_{0+s})$
	$D(gk_0, b_0), D(gk'_0, b_0)$
方向	ゴール接近度
時間	$t_i - t_{i-1}, t_{i+1} - t_i$
状況	2 値分類の結果

降では用いることとした。

3.4 プレー認識

3.4.1 Conditional Random Field

CRF は観測に加えてラベルの順序関係を考慮できる機械学習アルゴリズムであり、品詞付与 [22] や固有表現抽出 [22] といった系列ラベリング問題で高い精度を示すことが知られている。サッカーにおいて発生するプレーには、時系列的に相関が見られる場合があるため、CRF によってプレーの前後関係を考慮した分類を行うことが有効と考えられる。例えば、キャッチというラベルが付与される場合、直後に発生するプレーはフィードである確率が非常に高い。一方で、トラップというラベルが付与される場合、直後に発生するプレーがトラップやドリブルである確率は極めて低い。本研究では、観測系列 x をトラッキングデータより抽出した特徴量系列、ラベル系列 y をプレー種類とした鎖状の Chain CRF モデルを用いる。3.2 節で検出した各プレー発生タイミングの系列に対して、最適なラベル系列 $\hat{y}$ を求めることで、プレーの前後関係を考慮した分類を行う。

観測系列 $x = \{x_1, \dots, x_n\}$ が与えられた場合にラベル系列（出力） $y = \{l_1, \dots, l_n\}$ となる条件付き確率は、以下のようにならされる。

$$P(y|x) = \frac{1}{Z_x} \exp\left(\sum_{i=1}^n \sum_k \lambda_k f_k(l_{i-1}, l_i, x)\right) \quad (3)$$

ここで、 Z_x は全系列を考慮した場合に確率の和が 1 となるようにするための正規化項であり、次の式で表される。

$$Z_x = \sum_{\hat{y} \in Y(x)} \exp\left(\sum_{i=1}^n \sum_k \lambda_k f_k(\hat{l}_{i-1}, \hat{l}_i, x)\right) \quad (4)$$

$f_k(l_{i-1}, l_i, x)$ は任意の素性関数であり、 i 番目、 $i-1$ 番目の出力ラベルと入力系列 x に依存する。素性関数とは、連続するラベル l_{i-1} 、 l_i および観測された素性 x が特定の 3 つの組である場合に 1、そうでない場合に 0 を返す関数である。 λ_k は素性関数の重みを設定するパラメータであり、学習により決定される。 $Y(x)$ は入力系列 x に対する出力

ラベル系列の集合を指す。入力系列 x が与えられたときの最適出力 $\hat{y}$ は以下で定義され、最適化には Viterbi アルゴリズムが用いられる [23], [24]。

$$\hat{y} = \operatorname{argmax}_{y \in Y(x)} P(y|x) \quad (5)$$

CRF を適用するため、本研究ではプレーの流れが開始してから流れが終了するまでの一連の区間を一つの系列とし、この区間内に検出した各プレーを認識対象とする。プレーの流れ開始とは、キックオフもしくはセットプレー（スローイン、フリーキック、コーナーキック、ゴールキック）が発生することを指し、プレーの流れ終了とは、ボールアウト、ファール発生、前半/後半終了のいずれかを指す。本研究では、実装に Python のライブラリである PyStruct[25] を利用し、学習に Block-Coordinate Frank-Wolfe アルゴリズム [26] を用いた。

3.4.2 Support Vector Machine

SVM は、Vapnik[27] が考案した Optimal Separating Hyperplane を起源とする超平面による特徴空間の分割法であり、2 値分類問題を解決するための最も優秀な学習モデルの一つとして知られている [28]。学習データセット $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_J, y_J)$ を 2 クラスに分類すると想定する。ここで、 $\mathbf{x}_i \in R^N$ および $y_i \in \{-1, +1\}$ は特徴ベクトルおよびクラスラベルである。この 2 クラスを超平面 $\mathbf{w} * \mathbf{x}_i + b$ で分割可能であると仮定し、データの分散に関して事前知識が得られていないものとする。この 2 クラスを分割する超平面と超平面に最も近いデータとの距離をマージンと呼び、マージンが最大となるとときに識別超平面は最適となる。最適超平面の $\mathbf{w}$ および b は次式で表される最小化問題を解くことで得られる。

$$\min \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{subject to } y_i(\mathbf{w} * \mathbf{x}_i + b) \geq 1, \forall i = 1, \dots, n \quad (6)$$

上式に対し Lagrange の未定乗数法を用いると、次式の識別関数が得られる。

$$f(\mathbf{x}) = \operatorname{sign} \left(\sum_{i=1}^n \lambda_i y_i \mathbf{x}_i * \mathbf{x} + b \right) \quad (7)$$

ここで、学習データの大部分の Lagrange 未定乗数 λ_i が 0 となり、 $\lambda_i > 0$ となる $\mathbf{x}_i$ のみで識別関数が決定され、このような $\mathbf{x}_i$ をサポートベクトルと呼ぶ。また、分離不可能な場合、Lagrange 未定乗数を $0 \leq \lambda_i \leq C (i = 1, \dots, n)$ に修正する。ここで、 C は誤識別に対するペナルティである [29]。

本研究では、Python のライブラリである scikit-learn[30] の OneVsRestClassifier を用いて、SVM による多クラス分類を実装した。これは、 K 個のクラスの分類を考える際に、ある特定のクラスに分類されるか、もしくは他の $K - 1$ 個のクラスのいずれかに分類されるかという 2 値分類問題を

解く分類器を合計 K 個用いる手法である。ハイパーパラメータの決定にはグリッドサーチを行い、全体の F 値が最も高くなる組み合わせ (rbf kernel, $C = 1000$) を決定した。

3.4.3 Random Forest

Random Forest (RF) では、複数の決定木を用いて識別を行う [31]。個々の決定木は高い識別性能を持つわけではないが、それぞれの結果を補うアンサンブル学習によって高い分類性能を得られる [32]。本研究では、実装に scikit-learn[30] の RandomForestClassifier を用いた。SVM と同様に、最適なハイパーパラメータの組み合わせ（決定木数 100、深さ 40）をグリッドサーチにより決定した。

4. 性能評価

4.1 評価環境

評価データとして、J1 リーグ公式戦で実際に計測された 20 試合分のトラッキングデータを用いた。トラッキングデータは TRACAB[33] と呼ばれる ChyronHego 社製のカメラシステムを用いて計測されたものであり、ピッチ全域を捉えた映像に対して画像認識技術を用いることでピッチ上の全選手とボールの移動軌跡を算出する [34]。正解データには訓練された専門家により記録されたラベルを用いた。

評価に利用したトラッキングデータには映像およびトラッキングアルゴリズムに起因する位置誤差が含まれており、特にボールアウト、キックオフ、セットプレー発生後はボールのトラッキングが大きく外れることが頻発する。このため、キックオフとセットプレー発生後の 10 秒間、ファール発生からフリーキックにより再開後 3 秒間、ボールアウト発生後およびゴール発生後からプレー再開までの期間は除外した。将来的には、トラッキングアルゴリズムの改良や、選手およびボールに直接取り付けられたセンサによりトラッキングを行うといった方法によりこのような問題は起こらないと考えている。

正解データのプレー発生時刻は試合映像を確認しながら人手で記録されたものであり、実際の発生時刻と比べて僅かにずれた時間が記録されている場合がある。このため、経験的にタイムウィンドウ W を 0.2 秒と定め、プレーが検出されたウィンドウの前後 2 ウィンドウ (0.4 秒) のずれは正解として許容するようにした。また、式 (1)、式 (2) のパラメータは $\theta_{TH} = 0.62 \text{radian}$, $V_{TH} = 0.006 \text{m}$ とした。これらの閾値は、プレー検出の F 値が最も高くなるように設定した。

プレー認識の評価では、16 試合分を学習データ、4 試合分をテストデータとした交差検証を行った。なお、学習時にクラスごとのサンプル数の偏りが及ぼす影響を抑制するため、各クラスのサンプル数の比率の逆数に応じた重みを与えている*1。

*1 実装に用いた Pystruct および scikit-learn の *class_weight* により重みを設定した。

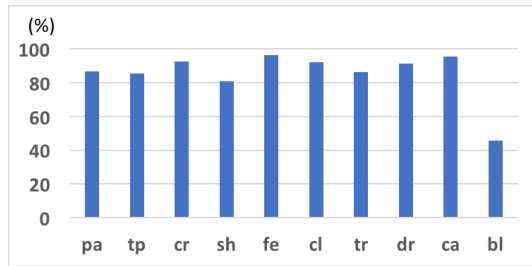


図 2 プレークラスごとの再現率

表 4 各学習手法を用いた場合における適合率

(%)	pa	tp	cr	sh	fe	cl	tr	dr	ca	bl
CRF	85.7	36.8	54.7	31.0	53.9	18.1	86.1	26.9	48.3	2.8
SVM	90.3	22.8	55.9	24.0	60.0	14.4	87.1	17.6	64.5	0.4
RF	89.2	71.0	74.1	44.4	68.7	32.7	83.7	42.4	60.6	1.3

4.2 評価結果

4.2.1 プレー検出

プレー検出の性能評価を行った結果、適合率 86.7%、再現率 85.3%、F 値 86.0% となった。さらなる分析のため、クラス別のプレー検出性能を図 2 に示す。再現率の低いプレーの特徴として、ボールの軌道や移動速度が殆ど変化しないプレーが大半を占めた。このようなプレーはトラッキングデータのみでは十分な特徴が表れないため、本質的に検出は困難であると考えられる。また、ウィンドウ幅 W は 0.2 秒に設定したが、実際には 0.2 秒以内に複数のプレーが発生する場合もごくまれに存在した。特に、ブロックは相手選手がボールを蹴った直後に発生しやすいため、再現率が 45.7% と極端に低くなっている。これに対しては、プレーごとに適切なウィンドウ幅を設定するといった工夫が必要となる。

適合率については、ボールが選手付近でバウンドする場合に誤検出が多く見られた。これは、バウンド直後にボールの移動速度が大きく変化するためであり、選手付近でのバウンドをトラッキングデータのみから判別することは非常に難しいことを示している。

また、適合率、再現率を共に低下させている主要因として、トラッキングデータの誤差が挙げられる。特に、今回は映像から導出されたトラッキングデータを用いているため、選手がボール付近に密集する状況では、位置誤差が大きくなりやすい。実際、シュートの再現率が 80% 以下となっていることから、トラッキングデータ誤差の及ぼす影響が大きいことが推測できる。従って、トラッキングアルゴリズムの改良や GPS の利用などによりトラッキングデータの精度を向上させることができれば、提案手法の性能向上が期待できる。

4.2.2 プレー認識

プレー認識の適合率、再現率、F 値をそれぞれ表 4、表 5、表 6 に示す。表 6 より、F 値に着目すると、大半のクラスでは RF を用いた場合に最も高い性能を示すことがわかった。ただし、ドリブルについては CRF の F 値が RF

表 5 各学習手法を用いた場合における再現率

(%)	pa	tp	cr	sh	fe	cl	tr	dr	ca	bl
CRF	84.9	39.4	78.4	50.0	78.7	30.0	80.5	55.2	84.3	16.1
SVM	81.7	64.8	68.0	33.3	63.9	60.0	79.5	46.6	78.4	3.2
RF	83.0	31.0	64.9	44.4	75.4	36.0	88.3	24.1	84.3	16.1

表 6 各学習手法を用いた場合における F 値

(%)	pa	tp	cr	sh	fe	cl	tr	dr	ca	bl
CRF	85.3	38.1	64.4	38.3	64.0	22.6	83.2	36.2	61.4	4.8
SVM	85.8	33.7	61.4	27.9	61.9	23.2	83.1	25.6	70.8	0.7
RF	86.0	43.2	69.2	44.4	71.9	34.3	85.9	30.7	70.5	2.4

表 7 RF における混同行列 (提案手法)

推定値 \ 真値	pa	tp	cr	sh	fe	cl	tr	dr	ca	bl
pa	2162	4	14	7	14	23	130	0	1	251
tp	30	22	1	2	0	0	3	0	0	13
cr	16	0	63	3	0	2	3	0	0	10
sh	7	0	2	16	0	2	7	0	0	2
fe	7	0	0	0	46	1	3	0	2	2
cl	13	1	1	4	0	18	6	0	1	6
tr	115	1	0	3	2	3	1685	14	3	83
dr	2	0	1	0	0	0	39	14	0	2
ca	1	0	1	0	0	1	3	0	43	2
bl	11	1	0	0	0	2	12	0	0	5
not play	59	2	2	1	5	3	122	5	21	8

を上回っている。この理由として、ドリブルの直後に発生するプレーはクロスまたはシュートである可能性が他のプレーと比較して遥かに大きいことが挙げられる。このような前後関係の結びつきが大きいプレーに関しては、系列を考慮可能な CRF を用いることが有用と考えられる。しかし、キャッチ、フィードのように前後関係の結びつきが大きいにも関わらず、CRF が RF の結果を大きく下回るプレーも存在する。この原因は、キャッチとして誤認識された場合に、高確率で直後のプレーもフィードとしてしまうことにあると考えられる。このことから、CRF により前後関係の結びつきを考慮することが常に有効とは限らないことが分かった。RF における各クラスの F 値に着目すると、試合中に発生するプレーのほとんどを占めるパス、トラップに関してはそれぞれ 86.0%、85.9% と高い値を達成している。この結果より、頻繁に発生するパス、トラップのほとんどを自動抽出できるため、手動で記録しているプレーのラベル付けを補助するという観点では、有用と考えられる。

プレー認識の詳細な分析のため、RF の混同行列を表 7 に示す。表 7 中の “not play” は、プレーでないにも関わらず誤って検出したものを指す。

混同行列より、スループスをパスとして誤認識するケースが多く見受けられる。これは、スループスとパスの定義の境界が曖昧であることが原因であると考えられる。スループスの基本的な定義は相手最終ラインの裏のスペースを狙ったパスであるが、映像を見ると、スループスの定義を満たしているにも関わらずパスという正解ラベルが付与されているプレーが多く存在した。また、トラップとドリ

ブルも相互に取り違えることが多い。トラップもドリブルもボールを運ぶプレーであるが、ドリブルは“相手選手を抜こうとするプレー”として定義されており、トラップとの区別が難しい。これらのプレーについては、定義そのものが曖昧なため、トラッキングデータのみで区別することは困難である。

一方で、プレーの定義が明確に異なるにも関わらず、パスとトラップを互いに混同する場合も多い。このような例として非常に近距離の味方に向けて出したパスをトラップとして誤認識する場合があった。また、トラップに失敗しボールが保持選手から大きく離れてしまった状況において、トラップをパスとして誤認識していた。今後これらの認識精度を向上させるためには、選手やボールに取り付けたセンサの計測データを用いる等、位置以外の情報も加えた認識を行う必要があると考えられる。

5. まとめ

本研究では、サッカーにおける選手とボールのトラッキングデータを用いてプレーを自動認識する手法を提案した。性能評価では、10種類のボールタッチプレーについて、プレー検出を用いて適合率86.7%、再現率85.3%となることが分かった。さらに、検出されたプレーを対象として機械学習によりプレー認識を行った結果、Random Forest (RF) が平均的に高い性能を示すことが確認された。特に、試合中に頻繁に発生するパスとトラップについては、それぞれF値86.0%、85.9%と高い精度で認識できることが分かった。

今後の課題として、様々なチームのデータを収集し、戦術の違いや対戦相手の違いによる影響を評価することが挙げられる。また、RFとCRFを組み合わせることで、さらなる精度向上を図ることや、ウェアラブルセンサによって得られる加速度などを併用することも検討している。

謝辞

本研究は、平成27年度～平成29年度スポーツ庁受託事業「スポーツ研究イノベーション拠点形成プロジェクト (SRIP)」における成果である。本研究で使用したトラッキングデータ、正解ラベルとして用いたプレーデータおよび試合映像を提供して頂いたデータスタジアム株式会社に感謝の意を表す。

参考文献

- [1] 誉田雅彰, 池永剛: センシング技術を用いたスポーツ情報解析, 情報処理, Vol. 57, No. 8, pp. 738–743 (2016).
- [2] Tora, M. R., Chen, J. and Little, J. J.: Classification of Puck Possession Events in Ice Hockey, *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 147–154 (online), DOI: 10.1109/CVPRW.2017.24 (2017).
- [3] Tsunoda, T., Komori, Y., Matsugu, M. and Harada, T.: Football Action Recognition Using Hierarchical LSTM, *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 155–163 (online), DOI: 10.1109/CVPRW.2017.25 (2017).
- [4] Mora, S. V. and Knottenbelt, W. J.: Deep Learning for Domain-Specific Action Recognition in Tennis, *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 170–178 (online), DOI: 10.1109/CVPRW.2017.27 (2017).
- [5] Kautz, T., Groh, B. and Eskofier, B.: Sensor fusion for multi-player activity recognition in game sports (2015).
- [6] Catapult: OptimEye S5 — Catapult Sports, Catapult (online), available from <https://www.catapultsports.com/products/optimeye-s5> (accessed 2018-01-24).
- [7] adidas Japan: miCoach smartball, adidas Japant (online), available from <https://shop.adidas.jp/micoach/smartball/> (accessed 2018-01-24).
- [8] Wei, X., Sha, L., Lucey, P., Carr, P., Sridharan, S. and Matthews, I.: Predicting Ball Ownership in Basketball from a Monocular View Using Only Player Trajectories, *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 780–787 (online), DOI: 10.1109/ICCVW.2015.106 (2015).
- [9] Direkçöglü, C. and O'Connor, N. E.: Team Activity Recognition in Sports, *Computer Vision – ECCV 2012* (Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y. and Schmid, C., eds.), Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 69–83 (2012).
- [10] Perše, M., Kristan, M., Perš, J., Mušič, G., Vučković, G. and Kovačič, S.: Analysis of multi-agent activity using petri nets, *Pattern Recognition*, Vol. 43, No. 4, pp. 1491–1501 (online), DOI: <https://doi.org/10.1016/j.patcog.2009.11.011> (2010).
- [11] Richly, K., Bothe, M., Rohloff, T. and Schwarz, C.: Recognizing Compound Events in Spatio-Temporal Football Data, *International Conference on Internet of Things and Big Data* (2016).
- [12] Chen, S., Feng, Z., Lu, Q., Mahasseni, B., Fiez, T., Fern, A. and Todorovic, S.: Play type recognition in real-world football video, *IEEE Winter Conference on Applications of Computer Vision*, pp. 652–659 (online), DOI: 10.1109/WACV.2014.6836040 (2014).
- [13] Lazarescu, M. and Venkatesh, S.: Using camera motion to identify types of American football plays, *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, Vol. 2, pp. II–181–4 vol.2 (2003).
- [14] Kim, W. and Kim, M.: Soccer kick detection using a wearable sensor, *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1207–1209 (online), DOI: 10.1109/ICTC.2016.7763408 (2016).
- [15] Maksai, A., Wang, X. and Fua, P.: What Players do with the Ball: A Physically Constrained Interaction Modeling, *arXiv preprint arXiv:1511.06181* (2015).
- [16] Corporation, T. D. S.: ディープラーニング技術: スポーツおよび接客作業の映像解析, Toshiba Digital Solutions Corporation (オンライン), 入手先 http://www.toshiba-sol.co.jp/tech/sat/case/1712_1.htm (参照 2018-01-24).
- [17] Kaneki, Y., Misumi, Y. and Miyamoto, M.: Analysis of Japanese Football Games by the Tracking Data and Sport Live Event Annotations, Vol. 496, pp. 39–48 (2017).

- [18] Horton, M., Gudmundsson, J., Chawla, S. and Estephan, J.: Automated Classification of Passing in Football, Vol. 9078, pp. 319–330 (2015).
- [19] Davatzikos, C., Ruparel, K., Fan, Y., Shen, D., Acharyya, M., Loughhead, J., Gur, R. and Langleben, D.: Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection, *NeuroImage*, Vol. 28, No. 3, pp. 663–668 (online), DOI: 10.1016/j.neuroimage.2005.08.009 (2005).
- [20] Guyon, I. and Elisseeff, A.: An Introduction to Variable and Feature Selection, *J. Mach. Learn. Res.*, Vol. 3, pp. 1157–1182 (online), available from <http://dl.acm.org/citation.cfm?id=944919.944968> (2003).
- [21] Chen, Y.-W. and Lin, C.-J.: *Combining SVMs with Various Feature Selection Strategies*, pp. 315–324 (online), DOI: 10.1007/978-3-540-35488-8_13, Springer Berlin Heidelberg (2006).
- [22] Lafferty, J. D., McCallum, A. and Pereira, F. C. N.: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, San Francisco, CA, USA, Morgan Kaufmann Publishers Inc., pp. 282–289 (online), available from <http://dl.acm.org/citation.cfm?id=645530.655813> (2001).
- [23] 工藤拓, 山本薫, 松本裕治: Conditional Random Fields を用いた日本語形態素解析, 情報処理学会研究報告 自然言語処理 (NL), Vol. 2004, No. 47, pp. 89–96 (オンライン), 入手先 <https://ci.nii.ac.jp/naid/110002911717/> (2004).
- [24] 薬師貴之, 太田学, 高須淳宏: CRF を用いた学術論文 OCR テキストからの自動書誌要素抽出, 情報処理学会論文誌データベース (TOD), Vol. 2, No. 2, pp. 126–136 (オンライン), 入手先 <https://ci.nii.ac.jp/naid/110007990044/> (2009).
- [25] Mueller, A.: PyStruct - Structured Learning in Python, (online), available from <https://pystruct.github.io> (accessed 2018-01-24).
- [26] Lacoste-Julien, S., Jaggi, M., Schmidt, M. and Pletscher, P.: Block-coordinate Frank-Wolfe Optimization for Structural SVMs, *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28*, ICML'13, JMLR.org, pp. I–53–I–61 (online), available from <http://dl.acm.org/citation.cfm?id=3042817.3042825> (2013).
- [27] Vapnik, V.: *Statistical Learning Theory*, John Wiley & Sons (1999).
- [28] 武藤亮介, 嶋田和孝, 遠藤勉: SVM と逐次学習を併用した HOG 特徴による手形状認識手法, 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理, Vol. 109, No. 471, pp. 459–464 (オンライン), 入手先 <https://ci.nii.ac.jp/naid/110008002632/> (2010).
- [29] 村尾和哉, 寺田努: 加速度センサの定常性判定による動作認識手法, 情報処理学会論文誌, Vol. 52, No. 6, pp. 1968–1979 (オンライン), 入手先 <https://ci.nii.ac.jp/naid/110008508026/> (2011).
- [30] Pedregosa, F.: scikit-learn: machine learning in Python, (online), available from <http://scikit-learn.org/stable/> (accessed 2018-01-24).
- [31] Breiman, L.: Random Forests, *Mach. Learn.*, Vol. 45, No. 1, pp. 5–32 (online), DOI: 10.1023/A:1010933404324 (2001).
- [32] 波部 斉: ランダムフォレスト, 技術報告 31, 近畿大学理工学部情報学科 (2012).
- [33] Corporation, C.: ChyronHego: TRACAB Optical Tracking, ChyronHego Corporation (online), available from <http://chyronhego.com/sports-data/tracab> (accessed 2018-01-24).
- [34] 加藤健太: サッカーにおけるデータ分析とチーム強化, 電子情報通信学会 通信ソサイエティマガジン, Vol. 10, No. 1, pp. 29–34 (オンライン), DOI: 10.1587/bplus.10.29 (2016).