

フットサル動画分析のためのデータ収集システム

大村 昇平^{1,a)} 谷村 朋樹^{2,b)} 大越 匡^{2,c)} 中澤 仁^{1,d)}

概要: データを収集する機器の進化に伴い、スポーツの世界でも多様な分析を行うことが可能になっている。しかし、分析に必要なデバイスは高価であるため、データ収集・分析の恩恵を受けることができるのは一部のみに限られている。この問題を解決するために、多くの人に普及しているスマートフォンで撮影された動画から、データを収集・分析できるシステムを提案する。提案するシステムでは、物体検出手法を用いて選手とボールを検出し、その検出結果のコート上での位置を使用することでフットサルの試合を分析する。本研究では、システムの個別の技術精度と分析の精度という2軸で評価した。システム個別の評価項目は、選手とボールの検出精度、選手の分類の精度、分析の精度の評価項目は支配率の精度とし、実験を行った。選手とボールの検出精度は選手が mAP スコア 0.81、ボールは mAP スコア 0.437 となり、選手分類の精度は高い精度を出すことができたが、1 クラス 30 % と低い精度になった。また、支配率に関しては、目視で判断した場合と大きな差はなかった。精度が低くなったものに関して考察した結果、小さな物体を検出するために更なる研究が求められるとわかった。

1. はじめに

多くの分野と同じように、スポーツの分野でもデータを収集・分析する技術が価値を生み出すことが増えている。また、特にデータ収集の技術の発展により、単純な集計データだけでなく、サッカーのような複数人が多様な動きを見せるスポーツのトラッキングデータ等を分析に使用することができるようになった。[1,2] 例えば、Hoang M Le らの研究 [1] では、機械学習技術によってサッカーのトラッキングデータから守備側の守り方を学習して実際の動きを予測することで、得点の期待値を算出している。このように、データを活用することで、高度な分析をすることができるようになっている。

また、上で述べたような研究的な分析だけでなく、試合に勝つためにもデータを分析することが重要になっている。日本のサッカーのトップカテゴリーである Jリーグでは、DataStadium 社のサービスで試合のデータを取得し、各チームがそれを使用して、試合に望んでいる。また、練習の際にはウェアラブルデバイスを用いて心拍数や選手の動

き等のデータを収集し、選手のコンディションを管理しているチームも多い。自チームや相手チームを分析することで質の高い準備を行うことができるため、競技者やスタッフにとってもデータを活用することは重要である。このように、データを活用することで、分析をする際に分析者の感覚に頼っていた部分が改善され、客観的な視点から試合や練習を評価することが可能になってきており、各競技は大きく進歩している。

しかし、データを活用することで競技力が上がっているカテゴリーがある一方で、多くの競技団体ではデータを活用することができていないという問題がある。その大きな原因として、データ収集に使用する機器が極めて高額であることがあげられる。ここでデータ収集に使用する機器の代表的な例として、ウェアラブルデバイスとカメラについて紹介する。まず、ウェアラブルデバイスはチームの人数分のデバイスを用意し、管理しなければならないため、高価である。例えば、S and C Corporation のウェアラブルデバイス [3] は、サッカーの約 1 チームの人数分でおおよそ 150 万円円で販売されており、手軽に使用することは困難である。次に、カメラを用いたものだと、DataStadium 社が Jリーグに提供しているサービスでも使用されている TRACAB というサービスがある。[4] このサービスはミサイルを追尾するために使われるような高精度なカメラとシステムを応用したもので、スタジアムに設置された複数台のカメラで撮影された動画からデータを収集するものである。その他のカメラを用いたものも複数台の固定カメラが必要になる

¹ 慶應義塾大学環境情報学部
Faculty of Environment and Information Studies, Keio University, Fujisawa, Kanagawa 252-0882, Japan

² 慶應義塾大学大学院政策・メディア研究科
Graduate School of Media and Governance, Keio University, Fujisawa, Kanagawa 252-0882, Japan

a) shoheidesuno1024@keio.jp

b) tanimu@ht.sfc.keio.ac.jp

c) slash@sfc.keio.ac.jp

d) jin@sfc.keio.ac.jp

など、使用するためには優れた環境が必要となるため、多くの競技団体では使用することが困難である。

このように、データを収集・分析するための技術は進歩しているものの、それらが高価であるために多くの人がそれを利用できていない。本研究は、この問題を解決し、一部の恵まれた環境の競技団体だけでなく、多くの競技団体でデータの活用を可能にすることを目的とする。

その目的を達成する手段として、本研究では、多くの人に普及しているスマートフォン等のデバイスで撮影された動画から分析に必要なデータを収集できるシステムの構築に取り組んだ。本研究で作成したシステムでは、分析のために必要なデータとしてコート内の選手とボールの位置のデータを取得している。その理由は、フットサルにおいて試合中に起きている現象を理解するために最も基礎的な情報となるのがコート内の選手とボールの位置だと考えたためである。このシステムは、はじめに必要な情報をユーザーに入力させ、その情報を使って各フレームに対して自動で処理を行うという流れになっている。システムの構成は、主にユーザーによるコートの範囲とユニフォームごとの選手の選択、各フレームの選手とボールの位置の検出、ユーザーの選択結果と検出結果を用いることによるコート内の選手のチーム分類、真上から見た時の座標への正規化という4つのモジュールから構成される。

実験では、各モジュールの精度の実験とシステム全体の評価のための実験を行った。モジュールごとの精度の実験では、選手とボールそれぞれに対する物体検出モデルの精度の実験、検出結果のユニフォームごとの分類の精度の実験を行った。物体検出のモデルの実験には、慶應義塾大学フットサル倶楽部エルレイナの試合動画の前半部分を使用し、50フレームごとに切り出してコート内の選手とボールにラベルをつけた画像253枚を使用した。システム全体の評価のための実験では、支配率を算出しその精度を評価した。

本稿の構成は以下の通りである。まず、2章ではスポーツにおけるデータ収集の手法と有効性と問題点について、事例とともに詳しく述べる。次に、3章では、2章で指摘した問題点を解決するための手段としてスマートフォンで撮影された動画から選手とボールの位置を座標として取得できるシステムを提案する。そして、4章でシステムを評価するための実験を行い、それぞれの結果に基づいた考察と課題について述べる。最後に、5章で本論文をまとめる。

2. スポーツにおけるデータ活用

はじめに述べたように、スポーツの分野でもデータ活用の重要性が高まっている。具体的な事例として、野球におけるセイバーメトリクスが有名である。[5] セイバーメトリクスは、勝つために必要な要素を重視するために様々な集計データを用いて選手の評価を行うものである。また、人間

の目で数えることができるような単純な集計データを扱ったものだが、技術の進歩に伴ってより複雑で大量なデータを収集・分析することができるようになってきている。そのおかげで、サッカーやフットサルのような複数人が多様な動きを見せるスポーツのより本質的な分析が可能になってきている。この章では、複数人が多様な動きを見せるスポーツにおいて発展的な分析を行っている事例と、その分析を行うために必要なデータ収集の方法について詳しく述べる。

2.1 スポーツにおけるデータ分析

スポーツには離散的なデータが重要なものと、複数人の動きの連続的なデータが重要なものに分けることができる。例えば、野球は、それぞれの選手の動きの自由度が低く、選手の動きによる戦況の変化が少ない。野球のようなスポーツでは、出塁率などの離散的なデータが重要である。一方で、サッカーやバスケットボールなどのスポーツは、それぞれの選手の動きの自由度が高く、各選手の動きによって試合の戦況が大きく変わる。サッカーやバスケットボールは、コート内の選手のそれぞれの連続的な動きのデータが重要である。離散的なデータは人間が目で見えて数えることで集めることが容易である一方で、様々な動きを見せる選手を全て人の目で追いかけて動きを把握するのは困難である。しかし、技術の進歩に伴って選手をトラッキングすることができるようになり、複数人が多様な動きを見せるスポーツにおいて、高度な分析を行うことが可能になってきている。

例えば、サッカーの研究としては、トラッキングデータから守備パターンを予測してゴール期待値を予測する研究 [1]、高いインテンシティのプレーの影響についての研究 [6] 等があり、複雑な分析が可能になっていることがわかる。

また、バスケットボールでも、データコンペティションのプラットフォームである CodaLab でトラッキングデータからどのプレーがスクリーンか予測するというコンペティションが開かれた。[7] このように、複数人の多様な動きをデータとして取り扱えることが可能になっており、競技は日々進化している。

フットサルはマイナースポーツであり事例が少ないが、サッカーやバスケットボールと同じような複数人が多様な動きを見せるスポーツであり、これらのスポーツと同じような分析を行うことで価値を生むことができる。

これらのスポーツの分析を支えるデータ収集の技術とその問題点について、次で述べる。

2.2 データ収集の手法とその課題

上で述べたような分析を行うためには、人間の目で見ただけでなく、機械の力を借りて自動で大量のデータを収集できる必要がある。機械の力を借りたデータ収集の例とし

て、例えばサッカーではウェアラブル機器 [3] や高精度なカメラ [4] が使用されている。これらの方法は高精度である一方で、端末を揃えるための費用が効果であったり複数台の決まった角度から撮影するカメラが必要であったりするため、一部の資金的に恵まれているチームしか使用することができない。また、1 視点で撮影された動画を用いて、ユーザーが選手の領域を選択し、それをトラッキングする方法 [8] もあるが、これは 1 人 1 人選手を選択したり、トラッキングが外れた時に自分で選手の領域を選択し直す必要があるため、時間がかかる。さらに、このような手法ではボールのような小さな物体のトラッキングは極めて困難である。

このように、選手の複雑な動きをデータとして収集する手法は進歩している一方で、それらを利用するためのコストが高いため、利用できる人が限られているのが現状の課題である。

3. 提案手法

本研究では、コストの高さを原因としてデータを活用できる人が限られているという問題を解決するために、多くの人に普及しているスマートフォン等のデバイスで撮影された動画から低コストでデータを収集できるシステムを提案する。具体的には、スマートフォン等の広く普及しているデバイスで撮影された動画から選手とボールを検出し、その座標を用いて選手のチーム分けや座標の正規化を行う。本システムは、ユーザーによるコート上の範囲とユニフォームごとの選手の選択、各フレームの選手とボールの位置の検出、ユーザーの選択結果と検出結果を用いることによるコート内の選手のチーム分類、コートを上から見た際の座標への正規化の 4 つのモジュールで構成される。

ユーザーによるコート上の範囲選択では、ユーザーにコートの角をクリックしてもらい、その座標を保持する。次に、ユーザーによるユニフォームごとの選手を選択では、ユーザーにコート内に存在する選手をユニフォームごとに 1 人ずつ長方形で囲んでもらい、その長方形の中心近くの平均色を保持する。選手とボールの位置の検出には、深層学習を用いた物体検出モデルの 1 つである YOLOv3 を使用した。コート内の選手のチーム分類では、最初に保持しておいたそれぞれのユニフォームごとの色と、検出結果の中心部分の色の間の距離を比較することでチームを分類している。また、真上から見た際の座標への正規化を行う際は、変換前と変換後の対応する点の座標を用いて射影変換を行っている。

本章では、上で述べたシステムの各モジュールで行う処理について詳しく述べる。システムの概要については図 1 で示す。

3.1 コートとチームの指定

本システムでは、最初にユーザーにコート上の範囲とチームの分類を行ってもらい、コート上の範囲は、動画内に写る角の座標をクリックしていき、その内側をコート上の範囲として考える。コート上の範囲は、試合に関係のない領域のデータを削除する際と、コートを上から見た際の座標に変換する際に使用する。また、チームの分類は、選手が着ているユニフォームの色ごとに 1 人ずつ選手が写っている領域を長方形で選択してもらい、この時、フットサルでは、ホームチームのゴレイロ・フィールドプレイヤー、アウェイチームのゴレイロ・フィールドプレイヤーの合計 4 種類のユニフォームを着た選手がコート内で試合を行うため、4 人のプレイヤーが選択されることになる。選択された範囲のうち、高さと幅を両端からそれぞれ 3 分の 1 ずつ切り取った長方形の領域を中心部分とし、RGB 各色の平均色をとり、平均色として保持しておく。ここで保持する平均色は、検出結果をチームごとに分類する際に使用する。

3.2 選手とボールの検出

本システムでは、入力された動画に対してフレームごとに物体検出を行い、選手とボールを検出する。選手とボールの位置を検出する理由は、フットサルの試合中に起きている現象を理解するために最も基礎的な情報となるのがこれらの情報だと考えたためである。

選手とボールを検出するために使用した物体検出モデルは、YOLOv3 である。[9] YOLOv3 は画像を入力として、その中で物体が写っている領域を返すモデルである。YOLOv3 による検出結果は、図 2 の左の画像のようになる。また、YOLOv3 は、物体検出モデルの中で、精度と速度が共に優れているモデルである。

本研究では、選手とボールを検出するために汎用の重みを使用するだけでは不十分であったため、動画内の選手とボールの位置にラベルをつけたデータセットを用いてファインチューニングを行った重みを使用している。

3.3 チーム分類

チーム分類には、最初にユーザーが選択した選手の範囲の中の中心部分の平均色を用いる。選択された範囲の中心部分を使用する理由は、物体検出の結果が選手が写っている部分を長方形として返され、その領域を全て使用するとコート上の色等のユニフォームではない色が多く含まれてしまうためである。検出された物体のうち、選手として検出された領域のそれぞれの高さと幅を両端からそれぞれ 3 分の 1 ずつ切り取った長方形の領域の平均色を計算する。そして、全ての選手として検出された領域の色と、最初にユーザーが選択した 4 通りのユニフォームの色の類似度をユークリッド距離で計算し、最も近かったユニフォームの色をその選手の着ているユニフォームとして判断する。ここで

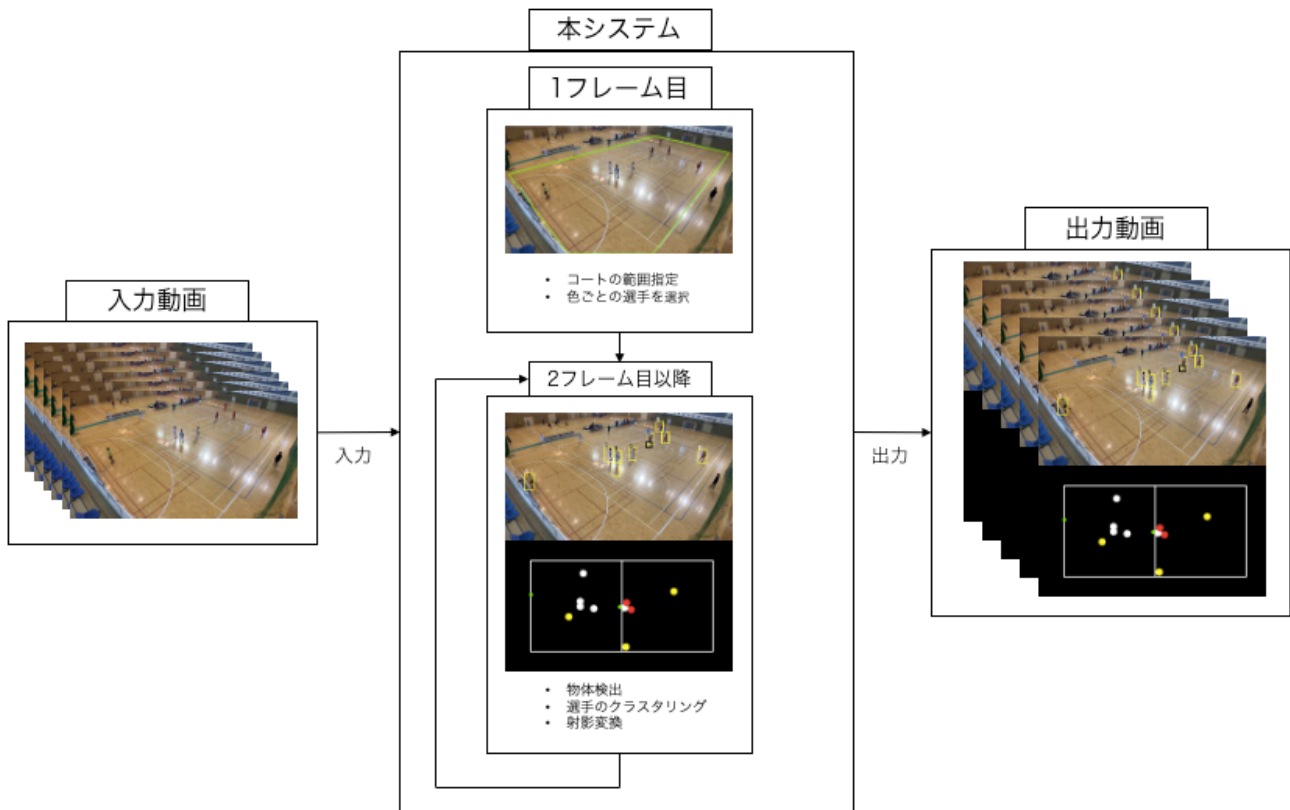


図 1 本システム全体の概要。本システムは、ユーザーが撮影した動画を入力とし、初めのフレームで必要な情報を選択してもらい、その後に各フレームに対して処理を行い、真上から見た際の座標に変換したものを出力としている。また、本システムは、ユーザーの入力、物体検出、ユニフォームごとの選手の分類、座標の変換という 4 つのモジュールで構成されている。

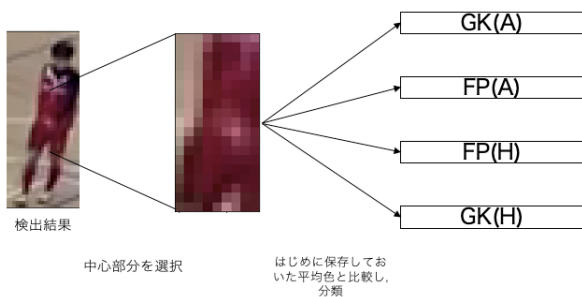


図 2 検出結果の分類の例。中心部分の色をはじめに選択した領域の平均色と比較して最も近いものに分類する。

行っていることの例を図 2 で示した。着ているユニフォームの色ごとに番号を割り当て、最終的な出力に使用する。

3.4 座標変換

検出された選手の座標を射影変換を用いてコート真上から見た時の座標へと正規化する。正規化が必要な理由は、スマートフォンで動画を撮影することを想定した時、会場やその時の条件等により撮影される角度が異なることが考えられるためである。撮影される角度によって取れるデータの形が変わらないようにコートの真上から見た座標に正規

化をする必要がある。射影変換を行うために、最初にユーザーが選択したコートの角の座標を用いる。図を挿入したい。対応する 4 点の座標を求めることができれば座標をコートの真上から見た時の座標へと変換するための変換行列を求めることができ、その変換行列を用いて検出結果の座標を変換することで正規化を行う。検出結果の座標として用いるのは検出結果の中央下の点とした。これは斜めから撮影された動画の場合地面と接している点が検出結果の長方形のうち中央下の点となるためである。また、この際コートの外に位置する検出結果に関しては試合を分析するために必要なデータではないので削除する。

4. 実験

はじめに本システムの実装環境について記す。本システムの使用言語は、Python3.7.3 である。フレームワークは、深層学習に関わる処理を PyTorch、画像の読み込み、加工、書き出し等の処理を OpenCV を用いた。

本研究では、提案手法の評価をするために個別のモジュールの評価とシステム全体の評価をそれぞれ行った。個別のモジュールでは、物体検出の精度、チーム分類の精度をそれぞれ評価した。物体検出の精度は、選手とボールの検出精



図 3 データセットの例

度を, 作成したデータセットに対するモデルの予測を物体検出の評価指標である mAP(mean Average Precision) で評価した. チーム分類の精度は, 1 秒毎に動画のフレームを抽出し, そのフレームの中で正確にチームを分類できている割合を調べた. システム全体の評価では, 支配率を算出し, 実際に目で見た際との比較を行った. ここで, 支配率の定義を一番ボールに近い選手がボールを支配していると考えたことにした. 本章ではこれらの実験についての詳細と, その結果から得られる考察について記述する.

4.1 データセット

使用したデータセットは慶應義塾大学フットサル倶楽部エルレイナの試合動画前半を 50 フレームごとに切り出し, コート内の選手とボールにラベルをつけたもの 253 枚である. また, このデータセットは, コート外の範囲を黒く塗り潰したものである. 黒く塗り潰したのは, 必要以上に小さい人間やボール等のコート内の物体を検出するのに必要な部分の情報をモデルが学習しないようにするためである. 253 枚のデータセットのうち 208 枚をトレーニング用, 残りの 45 枚をテスト用とし, YOLOv3 のファインチューニングに使用した. データセットの例を図 3 で示す.

4.2 結果

物体検出の精度は, 選手とボールの 2 つのクラスにわけ, それぞれ mAP スコアを用いて評価した. プレイヤーの検出では, mAP スコア 0.81 となり, 比較的高い精度での検出ができた. 一方, ボールの検出に関しては, mAP スコア 0.437 とプレイヤーに比べて精度が大きく下がった. ユニフォームごとの選手の分類は, 正しく検出できたものに関して, 実際のユニフォームの色ごとに分類できているかを目視で比較した. その結果, ホームチームの GK 以外は 100 %に近い精度で分類できていたが, ホームチームの GK のみ 30 %と低い結果となった. 支配率は, 目視では, 正確な距離を図ることは不可能であるため, 著者の経験上どちらのチームがボールを支配しているかを判断し, 目安となる支配率として算出した. 一方, システム上ではボールとの距離が一番近い選手がボールを保持していると判断して

表 1 検出精度

検出対象	mAP
人	0.810
ボール	0.437

表 2 クラスタリングの精度

クラスター	Accuracy
GK(A)	0.993
FP(A)	0.979
FP(H)	0.982
GK(H)	0.302

表 3 ホームチーム支配率

	支配率
目視	0.6
システム	0.605

算出した. その結果, 目安とした目視で判断した支配率がホームチームが 60 %, システムによる算出結果はホームチームが 60.5 %になった. 表に上記の各モジュールに対する実験の結果を示した.

4.3 考察

実験の結果を踏まえ, 評価できる点と改善すべき点にわけ, それぞれについて考察する.

まず, 評価できる点としては, チームの分類と座標変換の 2 つがある. チーム分類モジュールでは, 検出結果が正しいものに関しては概ね正しく分類することができていたため, YOLO の検出結果をユニフォームごとに分類するためには有効であると言える. また, 座標変換モジュールに関しても, 斜めから撮影された動画の選手の位置を真上からみた時の座標に変換することができたため, 分析に必要なデータ収集を行うシステムにとって価値があると言える.

一方で, 改善すべき点は, 選手とボールの検出精度である. 選手とボールの検出精度は, 表 1 の通りだが, 精度を下げている要因は誤検出の多さだった. 誤検出が多くなる理由としては, 使用した動画では選手やボールが小さく写っていることが考えられる. 特にボールは大きさだけでなく, 素早い動きをするためにぼやけてしまったり, 人の影に隠れてしまったりすることにより正しく検出できないことが多かった. 人間が動画にうつるボールの位置を探す際には, 数フレーム見えなくなっているフレームがあったとしても前後のフレームの情報から推測して見つけることができるため, ボールを検出するモデルにも 1 フレームだけではなく, 前後のフレームの情報を使用できるようにしたい. フットサルではなくテニスのボールの検出のための研究だが, 実際に複数フレームを使用した方がボールの検出精度が高いという研究 [10] があるため, 本研究にも有効だと考えている.

また, 本システムでは各フレームに対して画像処理を行い, フレームごとの座標を離散的に取得していたため, 分析を行うためのデータとしては不十分な可能性がある. 選手

ごとの動きやチームとしての特徴を分析するためには、試合を通してのトラッキングデータが必要な可能性があるため、物体検出だけでなくトラッキングの機能の追加も求められるかもしれない。

5. まとめ

スポーツにおけるデータ活用の重要性が増している一方で、その恩恵を受けることができない競技団体が多く存在するという問題の解決のために、多くの人に普及しているスマートフォン等のデバイスで撮影された動画からフットサルの試合のデータを収集するためのシステムの作成に取り組んだ。提案したシステムでは、動画の各フレームに対して画像処理を行い、選手とボールを検出し、その検出結果を処理することでチーム分類や座標の変換を行った。実験では、検出結果が正しい時にチーム分類や座標変換が有効であることがわかった一方で、選手とボールの検出精度に改善の必要があることがわかった。精度を落としている要因として、動画の中に写る物体が小さかったりぼやけたりしていることが考えられるためその点を踏まえて改善に取り組みたい。また、今後の展望として、トラッキングの機能の追加にも取り組んでいきたい。

謝辞

本研究は、JST, CREST, JPMJCR19A4 の支援を受けたものである。

参考文献

- [1] Hoang M Le, Peter Carr, Yisong Yue, and Patrick Lucey. Data-driven ghosting using deep imitation learning. 2017.
- [2] データによってサッカーはもっと輝く — football lab [フットボールラボ]. <https://www.football-lab.jp/>. (Accessed on 09/07/2020).
- [3] Gps トラッキングシステム : fieldwiz v2 -フィールドウィズ — S & C スポーツ科学計測テクノロジー スポーツパフォーマンス分析. <http://sandcplanning.com/solution/category/detail/?cd=11>. (Accessed on 09/07/2020).
- [4] Tracab optical tracking - chyronhego. <https://chyronhego.com/products/sports-tracking/tracab-optical-tracking/>. (Accessed on 09/07/2020).
- [5] Philip Beneventano, Paul D Berger, and Bruce D Weinberg. Predicting run production and run prevention in baseball: the impact of sabermetrics. *Int J Bus Humanit Technol*, Vol. 2, No. 4, pp. 67–75, 2012.
- [6] Valter Di Salvo, Warren Gregson, Greg Atkinson, P Tordoff, and Barry Drust. Analysis of high intensity activity in premier league soccer. *International journal of sports medicine*, Vol. 30, No. 03, pp. 205–212, 2009.
- [7] Codalab - competition. <https://competitions.codalab.org/competitions/23905>. (Accessed on 09/07/2020).
- [8] Sport tracking project. <http://www.bioirc.ac.rs/index.php/sport-tracking-project>. (Accessed on

- 09/07/2020).
- [9] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.
 - [10] Yu-Chuan Huang, I-No Liao, Ching-Hsuan Chen, Tsi-Uí ĩk, and Wen-Chih Peng. Tracknet: A deep learning network for tracking high-speed and tiny objects in sports applications. In *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–8. IEEE, 2019.