

# 俯瞰・主観映像の統合によるスキルの解明： フットサルにおけるイベント単位の視線配分可視化

中野 裕基<sup>1</sup> 松岡 大樹<sup>1</sup> 和薬 崇伸<sup>1</sup> 牧野桂一朗<sup>1</sup> 丸西 遼<sup>1</sup> 佐々木 博史<sup>1</sup>  
廣江 葵<sup>1</sup> 林 悠太<sup>2</sup> 佐藤 亮<sup>2</sup> 夏川 浩明<sup>1,a)</sup>

**概要：**フットサルにおいて、状況把握のための視覚探索（スキヤニング）は技術と同等に重要なスキルであるが、実環境のトレーニング課題における定量的な検証は十分に確立されていない。本研究では、フットサルの代表的なパス練習である「ロンド」を対象に、パスおよびレシーブの成功・失敗といったイベント単位で選手の視線配分を定量化することを目的とする。手法として、体育館上部からの俯瞰映像によりイベント（pass, receive, miss）を特定し、ウェアラブルアイトラッカー（Tobii Glasses 3）で取得した視線データと統合する解析パイプラインを構築した。注視対象の特定には、物体検出モデル YOLOv11 を用いて「ball（ボール）」「defender（守備側選手）」「keepers（自分以外の保持側 3 選手）」などを主観映像から自動検出し、イベント直前 3 秒間の視線配分を算出した。12 名の選手を対象とした分析の結果、パス成功時には味方と守備者を一定の割合で注視していたのに対し、パスミス時には守備側選手（defender）への注視割合が増加し、ボール（ball）への注視が減少する傾向が確認された。これにより、失敗局面において守備プレッシャー等の要因が視線行動に変化を与える可能性が示唆された。本手法は、実環境下でのスポーツスキルの可視化および練習設計への還元寄予するものである。

## 1. はじめに

チームスポーツにおける意思決定プロセスの根幹をなすのが、環境から得られる視覚情報の知覚である。特に、狭小なフィールドと極めて短い時間的猶予を特徴とするフットサルにおいて、その重要性は顕著である。選手には、高度なボールコントロールを実行しながら、同時に敵味方の配置やプレッシャーの強度、パスコースといった複雑な状況を瞬時に把握することが求められる。そのため、プレー中に視線をどこへ向け、いかなるタイミングで周囲を探索（スキヤン）するかという視覚戦略は、フィジカルやテクニックと並び立つ極めて重要なスキルとして位置づけられている。

これまでの研究において、ボール受球前のスキヤニング行動が、その後のターンやパスの精度向上に寄与することが実証されている [1]。また、実戦を想定した視線計測（アイ・トラッキング）研究によれば、熟練選手は 1 秒未満の短時間で探索を行い、ボールの状態に応じてその持続時間を調整していることや、探索対象が味方よりも相手選手に偏重する傾向などが報告されている [2]。一方で、シュート

局面においてはボールを視覚的なアンカー（係留点）として中心視しつつ、周辺視を用いて守備者やゴレイロの挙動を監視する戦略が有効であるとされる [3]。さらに、熟練者と初心者の比較からは、長期的な競技経験が「見る」スキルを形成し、ボール操作中であっても周辺への注意配分を維持可能にすることが示唆されている [4]。

しかしながら、既存の知見の多くは、公式試合の分析や実験室的な統制課題から得られたものであり、日常的に行われるトレーニング（例：4 対 2 のロンド）の最中に、視線行動がどのように組織化されているかについては不明な点が多い。特に、動的かつ複雑な実環境下において、選手が「何を」「どの程度の時間」注視しているのか、その視線配分が熟練度やポジション、あるいはプレーの成否とどのように関連しているのかを定量的に検証した事例は乏しい。近年、研究のフィールド志向が高まっているとはいえ、ウェアラブルデバイスを用いた自然なプレー環境下でのデータ蓄積は依然として黎明期にあり [5], [6]、トレーニングの設計や指導現場へ直接的に還元しうる客観的エビデンスの確立が急務となっている。

そこで、本研究の目的は、フットサルのロンド練習において、保持側選手がパスやレシーブを行う直前にどの対象へ視線を配分しているかを定量化し、パスやレシーブの成

<sup>1</sup> 大阪成蹊大学データサイエンス学部

<sup>2</sup> 大阪成蹊大学経営学部

<sup>a)</sup> natsukawa@osaka-seikei.ac.jp

功 (pass, recei) と失敗 (miss\_P, miss\_R) で視線行動がどのように異なるかを明らかにすることにある。そのために、パス回しのシーンに特化した物体検出モデルを作成し、検出結果と計測した視線時系列データを統合し、注視対象を自動的に算出する解析パイプラインを構築した。その上で俯瞰映像から作成したイベントデータに基づきイベント直前 3 秒区間の注視対象割合を算出する解析手順を構築する。さらに、セッション間のばらつきを抑えつつ選手間比較を可能にするため、2 セッション単位での重み付き集計を導入し、全 12 選手を対象に ball, defender, keepers (自分以外の保持側 3 選手), other (未検出, 背景, 自分自身など) に統合した指標として可視化・比較することを狙いとする。

## 2. データ取得実験

本研究では、フットサルプレイヤーの視線計測を行うため、ロンド実施中の視線データに加え、選手視点と俯瞰視点の映像を収集した。ロンドはフットサルで広く用いられるボール保持トレーニングであり、通常は様々な人数構成で行われる。本実験においては、図 1 の通り保持側 4 名、守備側 2 名の形式で実施した。ルールとして、保持側は守備側の周囲でボールを回して保持を続け、守備側はボール奪取を目指してプレスをかける。また、プレイヤー識別の観点から、保持側には黒・ピンク・緑・オレンジのウェアを着用させ、守備側はゼッケン無しの通常ウェアとした。

### 2.1 被験者

実験協力者は、O 大学フットサル部に所属する男性 18 名 (保持側選手 12 名, 守備側選手 6 名) である。このうち、視線データの解析対象となる保持側選手 12 名の詳細を表 1 に示す。ポジション構成はアラ 8 名, ゴレイロ 2 名, フィクソ 1 名, ピヴォ 1 名であった。本実験は所属機関の倫理審査委員会の承認を得ており、全参加者に対し文書および口頭で研究内容を説明した上で、参加の同意を得て実施した。

表 1 被験者情報

	平均	最小	最大	標準偏差
年齢	19.8	18	22	1.53
サッカー経験	10.2	0	15	4.4
フットサル経験	5.3	1	14	4.1

### 2.2 システム

本研究では、Tobii 社製ウェアラブルアイトラッカー (Tobii Glasses 3) を用いて、視線データおよび選手の一人称映像を取得した。視線データのサンプリング周波数は 100 Hz, 一人称映像は 25 Hz (1920 × 1080 pixels) で記録された。また同時に、Sony α7R IV を用いて、体育館施設の



図 1 ロンド実施時の俯瞰カメラ映像のラベリングツール。2 人の守備側選手を取り囲む形で保持側選手がボールを取られないようにボール回しを続けるロンドの映像に対してパスやトラップ、パスミスなど全てのイベントとその発生時間を記録するツール。

2 階部分より俯瞰映像の撮影を 60 Hz (1920 × 1080 pixels) で行った。

### 2.3 手順

ロンドの開始および終了は、外部合図として笛の音により統制し、全ての試行において同一の手続きを実施した。本実験では 4 名のアイトラッカーと俯瞰カメラによる同時計測を行っているため、データ分析においては、この笛の音を基準としてデータの切り出し (トリミング) を行うことで、各デバイス間の同期を確保した。取得データのうち、アイトラッカーは 4 選手の視線と一人称映像の分析に用い、俯瞰カメラの映像はパス成功率などのパフォーマンス指標を算出するために用いた。

実験の手順は以下の通りである。

- (1) 4 人の被験者は 4 色 (ピンク, オレンジ, 黒, 緑) のゼッケンを着用。
- (2) 保持側選手 4 人全員にアイトラッカーを装着する。
- (3) 装着後キャリブレーションを実施し録画を開始する。俯瞰映像も同時に録画を開始する。
- (4) 笛の合図で約 3 分間のロンドを行う。
- (5) 終了後録画を停止する

上述の手順を 2 セットずつ行い、3 チーム × 4 名 × 2 セットののべ 24 回分 (12 名分) の視線データを取得した (うち欠損 4 のため、以降は 20 回分のデータを解析)。

### 3. 分析データの準備

計 20 回分 (12 名分) の視線計測より、被験者がロンド練習時に何をどのように見てプレーをしていたかについて視線データより算出するための解析パイプラインを作成する。アイトラッカーは眼鏡型で一人称映像を記録したシーンカメラとシーンカメラの映像の座標系における視線データが記録されている。また、シーンカメラと視線データはどちらも動的に変化するものであり、何を見ていたか導出するためにはシーンカメラ中に注視対象がいつどこで写っているかを検出し、検出した対象を視線で捉えていたかを判定する必要がある。そのため今回の実験環境の対象に特化した物体検出のモデルを構築し、視線計測と物体検出の統合解析を行った。以下に、物体検出モデルの構築と評価、視線計測と物体検出の統合解析について記す。

#### 3.1 画像認識モデルの作成

物体検出モデルの構築には Roboflow の学習パイプライン [7] を用い、ベースモデルとして YOLOv11[8] を採用した。YOLOv11 は Ultralytics 社により開発された YOLO シリーズの最新モデルであり、C3k2 ブロックなどの CNN ベースのアーキテクチャを採用することで、特徴抽出能力と推論速度の向上が図られている。Roboflow では用途に応じて複数のモデルサイズが選択可能であるが、本研究では検出精度を最優先するため、「Accurate」モデル (YOLOv11 相当) を選択した。本モデルは、対象クラスと信頼度に加えてバウンディングボックスを出力する。本研究では、各フレームにおける検出バウンディングボックスと視線座標を時系列的に統合し、観察対象の関心領域 (AOI: Area of Interest) を推定するために本モデルを使用した。

学習データの作成にあたっては、実験時にアイトラッカーのシーンカメラで撮影された映像から画像を切り出し、計 400 枚の初期データを用意した。さらにデータの多様性を高めるため、各学習サンプルに対して回転 ( $-15^\circ \sim +15^\circ$ )、クロップ ( $0\% \sim +20\%$  Zoom)、および最大 2.5 px のぼかし処理をランダムに適用するデータ拡張を行い、合計 1975 枚の学習データを構築した。学習プロセスでは、COCO データセット [9] の事前学習重みを初期値として利用し、追加学習 (ファインチューニング) を行った。なお、検出対象となるクラスには、「ball (ボール)」、「defender (守備側選手)」、およびウェアの色に応じた保持側選手として「orange」、「green」、「black」、「pink」の計 6 個の検出ラベル (det\_class) を設定した。

構築したモデルの精度については、物体検出の代表的な評価指標である mAP@50 が 86.2%、Precision が 93.9%、Recall が 84.5% となった。また、実際の推論結果を確認したところ、観察対象を安定的かつ高精度に検出できている

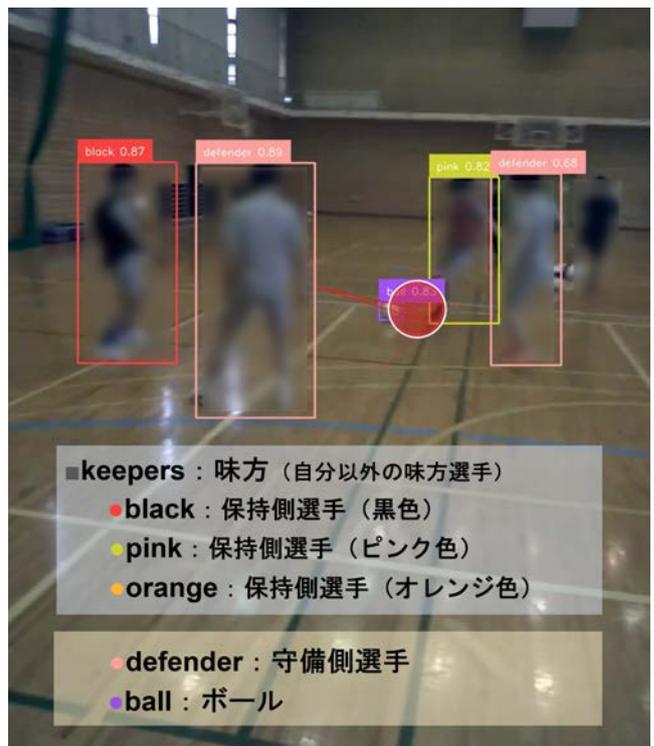


図 2 視線計測と物体検出を統合した後の画像：赤色の円が視線を表し、バウンディングボックスの色とそれに対応する名前を図中に記した。

ことが確認された。

#### 3.2 イベントデータの作成

ロンド実験においては、俯瞰カメラ映像を用い、図 1 の専用のラベリングツールによってイベントの抽出を行った。各選手がいつパスをしたか、いつパスミスをしたかというイベントの抽出結果を用いて、後のイベントに関連した視覚探索特徴を算出する。ラベリングツールでは、上から時計回りの正方形 (ひし形) 配置に基づき、1~4 の番号を選手に付与した。イベントラベルは「pass (パス)」、「recei (レシーブ、トラップの意)」、「miss P (パスミス)」、「miss R (レシーブミス)」の 4 種類とし、3 分間のロンドセッションにおいて、各選手がいつどの動作を行ったかを記録した。本研究では、3 分間のロンドセッション 6 回分 (計 3 セッション) から得られたイベントデータを分析対象とした。

#### 3.3 統合解析

本研究では、俯瞰映像から作成したイベントデータと、アイトラッカーで取得した視線時系列データを対応付け、イベント直前区間における注視対象割合を算出する統合解析パイプラインを定義した。解析単位は「イベント当事者 (当該イベントを実行した保持側選手)」とし、対応表 (イベント番号 ↔ 選手) に基づき、選手ごとに対象イベント集合を抽出する。

時刻同期は、イベント CSV の時刻 (HH:MM:SS.sss) を

秒へ変換し、セッション内の最初のイベント時刻を 0 とする相対時刻  $t_{ev}$  を定義することで行う。各イベントについて、直前 3 秒の窓区間  $[t_{ev} - 3, t_{ev}]$  に含まれる視線フレームを抽出し、窓内のカテゴリ別フレーム数を集計する。視線時系列はサンプリング間隔が一定でない可能性があるため、時刻列  $time\_sec$  の差分中央値を  $\Delta t$  とし、期待サンプル数  $N_{exp} = \lfloor 3/\Delta t \rfloor$  を求める。窓内の有効フレーム数  $N_{win}$  が閾値  $N_{thr} = \lfloor N_{exp} \cdot r \rfloor$  ( $r$  は最小有効比率) を満たさないイベントはノイズとして除外し、本稿の図示結果は  $r = 0.8$  を用いた。

検出ラベル (det\_class) は解釈可能性を高めるため、ball (ボール), defender (守備者), keepers (自分以外の保持側 3 選手), other (未検出, 背景, 自分自身など) へ統合する。以降のセッション間比較では、2 セッションを 1 単位 (1-2, 3-4, 5-6) として同一人物を  $player\_id$  で表現し、2 セッション分の有効イベント窓を結合した集合  $W$  上で注視割合を算出する。カテゴリ  $c$  の注視割合  $p_c$  は、各イベント窓  $k \in W$  におけるカテゴリ別フレーム数を  $n_{c,k}$  とし

$$p_c = \frac{\sum_{k \in W} n_{c,k}}{\sum_{k \in W} \sum_c n_{c,k}}$$

で定義し、 $\sum_c p_c = 1$  となるよう正規化する。本処理により、「全 12 選手 ( $player\_id$ ) × イベント種別 × 注視カテゴリ」の比較が可能となる。

#### 4. 分析結果と考察

keepers 統合後の 4 カテゴリ (ball, defender, keepers, other) について、全 12 選手 (2 セッション統合で作成した  $player\_id$ ) の注視割合をイベント種別ごとに算出した。ただし、本研究のデータでは other の割合が大きく、ball / defender / keepers の差が図上で判別しにくい。そこで、正規化 (全カテゴリで 100%) は維持したまま、グラフ表示において other のみを非表示とした。従って、図 3-5 の積み上げ棒グラフは ball / defender / keepers のみを示し、棒の総和は  $100\% - other$  となる。見やすさのため、縦軸は 0-40% に制限した。

図 3-5 より、全  $player\_id$  の平均と標準偏差を表 2 に示す。

表 2 全  $player\_id$  の平均と標準偏差

	ball 平均 ± 標準偏差	keepers 平均 ± 標準偏差	defender 平均 ± 標準偏差
pass	7.7% ± 3.8%	9.6% ± 3.5%	8.1% ± 2.0%
miss_P	5.5% ± 1.3%	7.5% ± 2.4%	7.8% ± 2.9%
recei	6.2% ± 2.9%	8.5% ± 2.8%	7.5% ± 3.3%

図 6 に、全イベント窓 (全選手・全窓) を統合したカテゴリ別分布の箱ひげ図を示す。この箱ひげ図も、分布の比

較を明確にするため other を非表示とした。

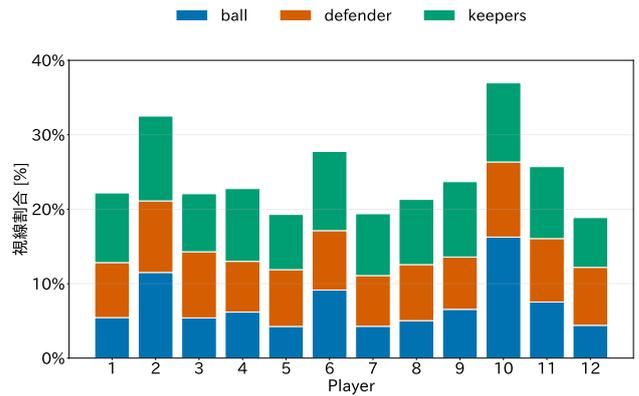


図 3 全 12 選手 : event\_type=pass 直前 3 秒注視割合 (積み上げ, other 非表示, 縦軸 0-40%)

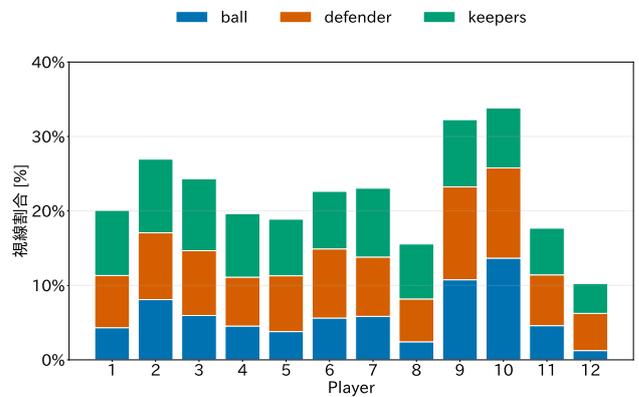


図 4 全 12 選手 : event\_type=miss\_P 直前 3 秒注視割合 (積み上げ, other 非表示, 縦軸 0-40%)

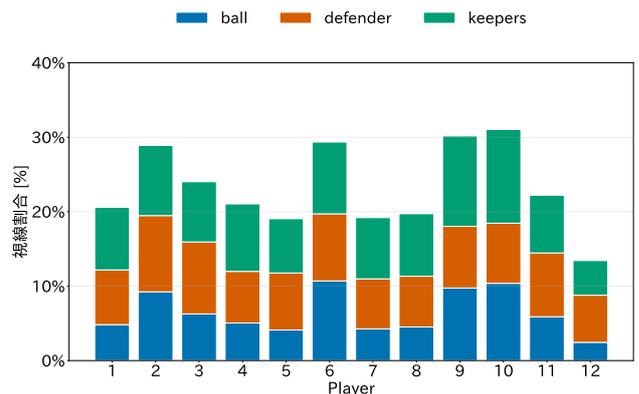


図 5 全 12 選手 : event\_type=recei 直前 3 秒注視割合 (積み上げ, other 非表示, 縦軸 0-40%)

図 3-6 より、成功局面 (pass, recei) では keepers と defender が近い水準で推移し、ball も一定量確保される傾向が確認できる。箱ひげ図の中央値と比較すると、pass では keepers 9.0% と defender 7.0% が ball 5.7% を上回り、recei でも keepers 8.2% と defender 6.8% が ball 4.8% を

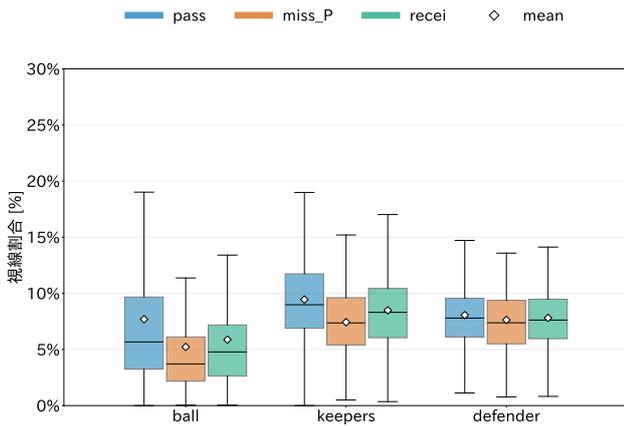


図 6 イベント別注視割合の分布 (箱ひげ, other 非表示): 全 player 統合 (miss\_R 除外)

上回った。この配分は、keepers と defender を継続的に監視しつつ、ボール処理に必要な ball 注視を確保する戦略と整合し得る。

失敗局面 (miss\_P) では、ball の中央値が 3.7% と低下し、defender (7.1%) と keepers (7.3%) が相対的に高い水準を占めた。守備圧や判断負荷の上昇により、パスコース探索や相手位置の確認が増加し、ボールへの注視が相対的に減少する可能性が考えられる。ただし miss\_P はイベント窓数が少なく (全窓統合で  $n = 70$ )、推定値が分散しやすい条件にあるため、傾向の一般化には追加データが必要となる。

また、図 3-5 では *player\_id* 間のばらつきが残り、特に ball の個人差が大きい。ボール注視が高い個体はトラップや次動作の安定化を優先しやすい一方、keepers 注視が高い個体は受け手探索や周辺配置の把握を優先する戦略を取り得る。これらの差は熟練度や役割 (ポジション)、立ち位置の幾何 (視野に入る相対配置) に依存すると考えられるが、本稿の集計のみから因果を断定しない。

最後に、本稿では図の判読性のため other を非表示としたが、other の割合自体は大きい。全イベント窓統合の平均は pass で  $74.8 \pm 11.0\%$  ( $n = 863$ )、miss\_P で  $79.7 \pm 10.2\%$  ( $n = 70$ )、recei で  $77.8 \pm 9.2\%$  ( $n = 273$ ) となる。本研究の統合処理は各フレームで視線中心に最も近い物体を 1 つだけ採用するため、other の増大は、(i) 明示的ターゲットを注視しない局面 (周辺視を用いた把握を含む)、(ii) 遮蔽やブレ等で検出が成立しない局面、(iii) 自分自身 (自分色) が視野に入り込む局面、が合算された混合指標と解釈できる。従って、other を分解可能にする追加ラベル設計 (未検出/背景/自分自身の分離) と検出モデルの改良、注視領域定義の再検討が今後の課題となる。

## 5. 終わりに

本研究では、フットサルのロンド練習を対象に、俯瞰映像から作成したイベントデータとアイトラッカーで取得し

た視線時系列データを統合し、イベント直前 3 秒区間における注視対象割合を算出する解析手順を構築した。アイトラッカーの主観映像から物体検出モデルにより注視対象を検出し、注視対象は ball, defender, keepers (自分以外の保持側 3 選手), other (未検出, 背景, 自分自身など) へ統合した。注視対象を 2 セッション単位で重み付き平均することで、全 12 選手の比較が可能となった。

結果として、pass, miss\_P, recei のいずれでも other が最大割合を占め、次いで keepers と defender が続き、ball は個人差が大きい傾向を示した。また miss\_P では defender が相対的に増加し ball が減少する傾向が見られ、失敗局面で守備者への注意配分が変化する可能性が示唆された。一方で other が大きいことは、自然環境下の遮蔽や急速な頭部運動、検出不能フレーム、自分色の混入などを含む指標であり、観測系と競技状況の双方の影響を受ける点に留意する必要がある。

今後は、イベント窓を単位とした統計検定や階層モデルにより、成功・失敗差と個人差を同時に評価する枠組みを導入する。また、結果で見られた個人差を説明するには、選手配置や選手の組み合わせなど視覚探索に影響しうる要因を加味した上での、視線行動の分析が必要となる。ロンド練習における視線行動の定量指標を、練習設計やフィードバックへ接続する知見として精緻化する。

**謝辞** 本研究は、大阪成蹊大学 学内共同研究費の支援を受けたものである。また、本研究に協力いただきました、びわこ成蹊スポーツ大学の山田庸 教授ならびに大阪成蹊大学データサイエンス学部の自主ゼミのメンバーの皆さん、実験参加者の皆さんに謝意を表します。

## 参考文献

- [1] Hunter AH, Smith N, Santiago PRP, Wilson RS. Visual Scanning and Technique Improve Performance in a Standardized Soccer Passing Task. *Applied Sciences*, 15(20):11045, (2025).
- [2] Aksum, K. M., Brotangen, L., Bjørndal, C. T., Magnaguagno, L., Jordet, G. Scanning activity of elite football players in 11 vs. 11 match play: An eye-tracking analysis on the duration and visual information of scanning. *Plos one*, 16(8), e0244118, (2021).
- [3] de Oliveira, T. A. C., Davids, K., Denardi, R. A., Zalla, S., Corrêa, U. C. Interpersonal coordination tendencies and perception of visual information for decision-making in futsal. *Psychology of sport and exercise*, 66, 102403, (2023).
- [4] Oppici, L., Panchuk, D., Serpiello, F. R., Farrow, D. Long-term practice with domain-specific task constraints influences perceptual skills. *Frontiers in psychology*, 8, 1387, (2017).
- [5] Kredel, R., Vater, C., Klostermann, A., Hossner, E. J. Eye-tracking technology and the dynamics of natural gaze behavior in sports: A systematic review of 40 years of research. *Frontiers in psychology*, 8, 287392, (2017).
- [6] Kredel, R., Hernandez, J., Hossner, E. J., Zahno, S. Eye-tracking technology and the dynamics of natural gaze

behavior in sports: an update 2016–2022. *Frontiers in Psychology*, 14, 1130051, (2023).

- [7] Dwyer, B. and Nelson, J. and Hansen, T., others. Roboflow (Version 1.0) [Software], <https://roboflow.com>,(accessed 2025-12-20), (2022).
- [8] Jocher, G., Qiu, J. Ultralytics YOLOv11, <https://github.com/ultralytics/ultralytics>, (2024).
- [9] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L. Microsoft COCO: Common Objects in Context, European Conference on Computer Vision (ECCV), pp. 740–755 (2014).