

1. はじめに

スポーツ選手の視線情報を取得して解析することは、状況判断能力等の選手の能力改善への利用が期待されている。とりわけ、サッカーではスキャンニングなどの周囲を確認する動作や、パスを受ける際の注視点位置などに関する研究が行われている [2, 3]。また先行研究では能力改善として、視野の狭さに着目した育成事例が報告されている。しかしながら、これまでの研究ではプロサッカー選手を対象とした一人称視点の動画像撮影、及びその解析は充分になされていない。そこで、本研究ではサッカーにおけるトレーニング時の視線行動についてプロサッカー選手とアマチュアサッカー選手の比較を定量的に分析することを目的とする。合わせて、プロサッカー選手の視線推定モデルを作成し、プロサッカー選手の視線と比較した際にアマチュアサッカー選手に気づきを促すことを目的とする。

2. 関連研究

2.1. サッカーにおける視線情報

サッカー選手を対象とした視線の研究はスキャンニング回数の比較 [2] と注視点比較 [3] がある。Phantakら [2] はサッカー選手権ユーロ 2016 において、セントラルミッドフィルダーを対象にスキャンニング回数が多い選手はパスの成功率が高いことを報告した。ここで、スキャンニングとは周囲確認のために頭部を振る動きである。Aksumら [3] ではエリートサッカー選手の試合において、スキャンニング中のフレーム内にチームメイトより対戦相手が多いことを報告した。

2.2. 視線推定

視線推定は、一人称視点の動画像から注視点座標を推定する手法である。代表的な視線推定手法に DINet [1] がある。この手法では、Dilated ResNet と Inception 構造の畳み込み層を用いることで特徴マップの解像度を落とさずに視線推定が行える。しかし、DINet は 2 次元視線推定であるため、時系列が考慮できないという問題点がある。

3. 提案手法

本研究では、プロサッカー選手とアマチュアとして大学生サッカー選手の視線の差異を確認し、プロサッカー選手の優れている点を定量的に評価する。次に、モデル化したプロサッカー選手の視線推定モデルを作成することで、プロサッカー選手と大学生サッカー選手の視線を定性的に評価する。

3.1. 測定概要

視線計測の際に、同条件下で 2 種類のトレーニングを行う。測定対象者はセントラルミッドフィルダーのプロサッカー選手 (Pro) と大学生サッカー選手 2 名 (Uni1), (Uni2) である。

測定項目は、ボールポジショントレーニング (BPT) と Small-sided games (SSG) である。ここで、SSG は攻撃フェーズ (Att) と守備フェーズ (Def) に分けて分析を行う。それぞれのトレーニングの条件を表 1 に示す。また、Uni2 は BPT のみを測定する。

測定変数は、スキャンニング回数とスキャンニング時の平均時間、スキャンニング時の注視点位置とする。また、スキャンニング時の注視点位置はフレーム内のボールがなくなったタイミングからボールが映るフレームまでとし、100msec 以上の時にフレームを目視で確認し、人手でクラスラベリングを行う。

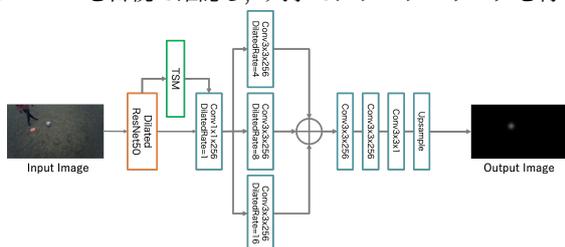


図 1: 提案手法のネットワーク構造

表 1: 各トレーニングの条件

	BPT	SSG
人数	7(攻)vs4(守)	11vs11
範囲	14m × 14m	52.5m × 40m
時間	10min	6min
条件	2 タッチ以下	フリータッチ

クラスは、Teammate(味方), Opponent(敵), Space(スペース), T/O(味方と敵), T/S(味方とスペース), O/S(敵とスペース), T/O/S(味方と敵とスペース) である。

3.2. 測定結果

プロサッカー選手と大学生サッカー選手の各トレーニングにおける測定変数の結果を表 2 に示す。

表 2: 各トレーニングにおける測定結果

		回数 (回)	注視点時間 (sec)	注視点位置 (%)						
				Teammate	Opponent	Space	T/O	T/S	O/S	T/O/S
BPT	Pro	117	0.187	35.1	15.8	30.7	12.3	4.4	1.8	0.0
	Uni1	52	0.160	44.2	15.4	30.8	0.0	7.7	1.9	0.0
	Uni2	31	0.162	29.0	16.2	32.3	9.7	6.5	6.5	0.0
SSG	Att_Pro	93	0.353	21.7	24.6	14.5	23.2	5.8	2.9	7.2
	Def_Pro	93	0.353	25.0	37.5	25.0	1.2	4.2	0.0	4.2
	Att_Uni1	49	0.216	21.6	13.5	37.8	16.2	2.7	8.1	0.0
	Def_Uni1	49	0.216	25.0	33.3	16.7	0.0	8.3	16.7	0.0

表 2 より、各トレーニングで Pro は Uni1, Uni2 と比較し、スキャンニング回数が多く、スキャンニング時の平均時間が長いことがわかる。先行研究 [2] から、プロサッカー選手の方が優れた選手であると言える。また、スキャンニング時の注視点は Pro が BPT では T/O, SSG では Opponent の割合が高く、視線の特徴として挙げられる。

視線分析結果から、各トレーニングにおいてプロサッカー選手と大学生サッカー選手で注視点位置が異なり、プロサッカー選手の視線を模倣することが、パフォーマンスの向上を促す要素であることが確認できた。

3.3. 視線推定モデルの構築

次に、視線推定モデルを構築する。従来の DINet は画像ベースの視線推定モデルであるため、時系列の考慮ができない問題がある。そこで本研究では、時系列を考慮した視線推定モデルである、TS-DINet を提案する。提案手法のネットワーク構造を図 1 に示す。提案手法では、時系列考慮をするために Dilated ResNet50 の後に TSM [4] を用いる。TSM は、特徴量の一部を前後のフレームとシフトし、次の畳み込み層へ伝達することで、時系列の特徴を獲得している。

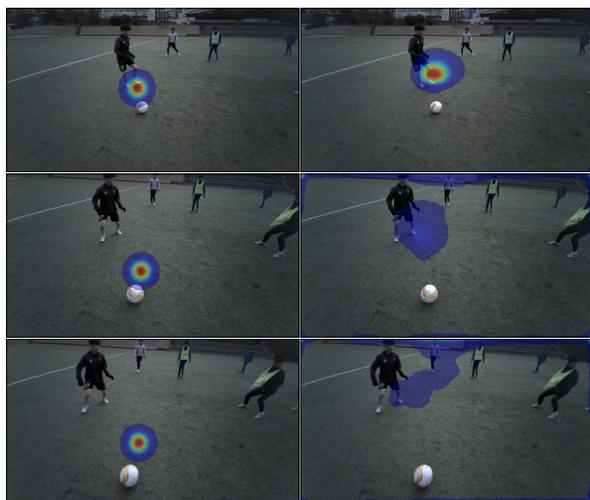
視線推定の出力結果は画像全体における人間の視野の確率分布と見なすことができる。予測された顕著性マップと真値を確率分布に変換する場合、これらの値を正規化する必要がある。一般的に用いられる L1 ノルムを用いて、線形正規化を行う。損失関数を以下の式 (1) に示す。ここで、 $x = (x_1, x_i, x_N)$ は正規化前の値である。

$$L(p, g) = \sum_i |p_i - g_i| \tag{1}$$

$$\text{where, } p_i = \frac{x_i^p}{\sum_{i=1}^N x_i^p}, \quad g_i = \frac{x_i^g}{\sum_{i=1}^N x_i^g}$$

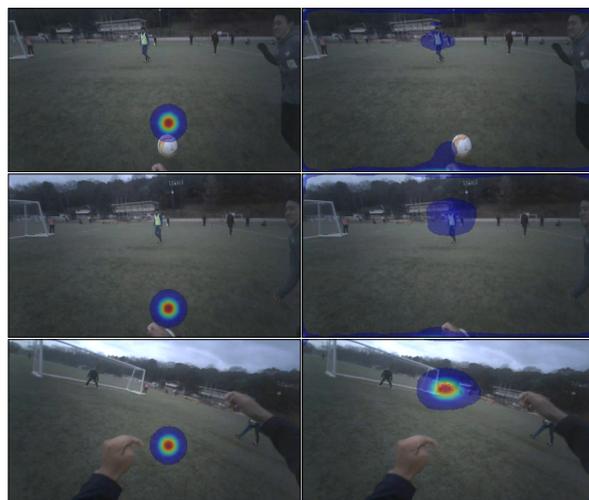


図 2: 定性的評価



大学生サッカー選手 プロサッカー選手

(a) BPT



大学生サッカー選手 プロサッカー選手

(b) SSG

図 3: 各トレーニングにおける出力結果比較

4. 評価実験

本章では、時系列考慮の有無でプロサッカー選手の視線の模倣できたかを検証する。比較対象として、DINet を用いる。

4.1. 実験概要

TSM で考慮するフレーム数 f は短い時系列考慮 ($f = 2, 4$)、スキニングの平均時間に合わせた時系列考慮 ($f = 6$)、長期的な時系列考慮 ($f = 12, 24, 36, 48$) で行う。使用するデータセットはプロサッカー選手の各トレーニングの動画フレームとする。BPT は学習に 8,884 枚/評価に 1,121 枚、SSG は学習に 6,944 枚/評価に 1,029 枚用いる。評価指標は Normalized Scanpath Saliency (NSS), TheLinearCorrelationCoefficient (CC), Kullback-Leibler (KL) を用いる。

4.2. 定量的評価

各トレーニングにおける定量的評価を表 3 に示す。表 3 より、従来手法である DINet と比較し、提案手法の方が精度が高いことが確認できる。

表 3: 定量的評価

		DINet	提案手法 (TS-DINet)							
			$f=2$	$f=4$	$f=6$	$f=12$	$f=24$	$f=36$	$f=48$	
BPT	NSS ↑	7.19	7.17	7.17	7.30	7.19	7.18	7.11	7.16	
	CC ↑	0.44	0.51	0.51	0.52	0.51	0.51	0.51	0.51	
	KL ↓	1.99	1.97	1.98	1.98	1.97	1.97	1.97	2.00	
SSG	NSS ↑	5.66	6.02	5.99	6.01	5.77	5.73	5.95	6.10	
	CC ↑	0.42	0.43	0.43	0.43	0.48	0.51	0.51	0.53	
	KL ↓	2.30	2.42	2.43	2.44	2.42	2.29	2.16	2.17	

BPT では、 $f = 6$ で最も精度が高いことが確認できる。BPT は限られた正方形の空間内でパス回しを中心にトレーニングであるため、味方の位置を確認するスキニングが最も重要である。よって、スキニングの平均時間に合わせた $f = 6$ の結果が最も優れているのではないかと考える。

SSG では、 $f = 48$ で最も精度が高いことが確認できる。これは、SSG が実際のゲームを模したトレーニングとなっており、スキニング時間も長いことから長期的な時系列考慮が有効であったためと考える。 $f = 12, 24, 36, 48$ では、いずれの評価指標もフレーム数が上がるにつれて評価が高くなる傾向が見られる。このことから、SSG において長期的な視線を考慮することは、実際のゲームで視線を模倣するモデルを作成する際にも重要であることが考えられる。

4.3. 定性的結果

各トレーニングにおける定性的評価を図 2 に示す。また、表 3 より BPT は $f = 6$ 、SSG は $f = 48$ を用いて評価を行った。

BPT では、DINet と比較して提案手法が真値に近い出力であることが確認できる。また、SSG も BPT と同様に DINet より提案手法が真値に近い出力であることが確認できる。DINet は手前の人物に注視する推定結果に対して、提案手法では真値と同様に、奥のボール保持者に注視するように推定できている。これは、提案手法は、時系列考慮に

よりボールを保持している情報を捉えることができ、奥の人物を注視できたためと考えられる。これらの定量的、定性的結果から、提案手法で BPT ではスキニングの平均時間に合わせた時系列考慮、SSG では長期的な時系列考慮によりプロサッカー選手の視線を模倣する結果が確認できた。

5. 大学生サッカー選手との視線比較

プロサッカー選手の視線を学習した TS-DINet を用いて、大学生サッカー選手の視線を評価する。大学生サッカー選手とプロサッカー選手の各トレーニングにおける視線比較を図 3 に示す。また、視線推定の信頼度が理由は、学習に使用したデータ数の問題であると考えられる。

BPT では、大学生サッカー選手は、ボールを追うような視線の動きをしているのに対して、プロサッカー選手の視線マップはフレーム上部に集中している。プロサッカー選手は、パスを出した後にスキニングや、周りの味方や人を確認する視線の動きをする。そのため、プロサッカー選手の視線では遠方の味方及び、パスを出した味方を注視する推定結果が出力されたと考えられる。

SSG では、大学生サッカー選手はゴール前に BPT と同様にボールに注視が寄っている。BPT と同様に、プロサッカー選手はボールより目の前の味方を注視する推定結果が出力された。

これらの結果から、プロサッカー選手の視線推定モデルは各トレーニングにおいて有効な結果を示しており、大学生に気づきを促せるレベルであると考えられる。

6. おわりに

本研究では、サッカーのトレーニングにおける注視点比較と視線推定を行った。データ分析の結果、プロサッカー選手の方が大学生サッカー選手より優れていることが定量的に判明した。視線推定は DINet と比較し、提案手法を用いることでプロサッカー選手の視線を模倣することができた。大学生サッカー選手との視線比較では、今後は、大学生サッカー選手に気づきを促すことでプレーの変化を確認する。

参考文献

- [1] S. Yang, *et al.*, "A Dilated Inception Network for Visual Saliency Prediction", TMM, 2019.
- [2] A. Phantak, *et al.*, "Keep Your Head Up : Correlation between Visual Exploration Frequency, Passing Percentage and Turnover Rate in Elite Football Midfielders", MDPI Sports, 2019.
- [3] K. Aksum, *et al.*, "What Do Football Players Look at? An Eye-Tracking Analysis of the Visual Fixations of Players in 11 v 11 Elite Football Match Play", IJSSC, 2021.
- [4] J. Lin, *et al.*, "TSM: Temporal Shift Module for Efficient Video Understanding", ICCV, 2019.

研究業績

- [1] 伊佐 稜 等, "物体追跡による経路データセットの自動生成", 電気・電子・情報関係学会 東海支部連合大会, 2021. (他 2 件)