

大規模言語モデルを活用したシュートシーン解説の自動生成

片山 葉一朗* 近藤 生也* 伏見 龍樹* 落合 陽一*

概要. サッカー中継の解説は「スペースを空ける動き」や「ビルドアップ」といったサッカー未経験者には理解しづらい表現が使用され、サッカーを詳しく理解し楽しみたい初学者には向かない。低い人的コストで詳しい解説を得る方法として解説の自動生成が試みられており、その中で選手の自動認識や可視化のサポートは実用化している。しかし、戦術を含めた解説の自動生成は近年の汎用的な LLM 技術をもってしても実現していない。その理由として、汎用的な LLM ではサッカーの分析に必要な情報抽出の仕組みが足りていないためだと考えられる。本研究ではサッカー初学者向けに丁寧な解説を自動生成することをタスクに位置づけ、既存の汎用的な LLM にサッカーに適したヒューリスティックな情報抽出を組み込むことで性能の向上が見られるかを検証し、また効果的な情報はどのようなものかを調査する。そしてシュートシーンのハイライト動画に対する解説生成と可視化までを行い、ユーザー体験の向上を目指す。

1 はじめに

サッカー中継の解説は「スペースを空ける動き」や「ビルドアップ」といったサッカー未経験者には理解しづらい表現が使用され、サッカーを普段からプレーしたり観戦したりしていないサッカー初学者には分かりづらい。その要因として専門用語が多く、素早い展開に合わせた解説を中継放送では必要とするため、用語が簡略化されやすい。一方で、解説者の仕事は専門性が高いため解説者を増やすことも難しい。自動化のアプローチとして選手の自動認識や可視化のサポートは実用化されているが、戦術を含めた解説の自動生成は実現していない。昨今の大规模言語モデル (LLM) や大规模視覚言語モデル (VLM) の急激な発展により、画像や動画から情報を取得し、自然言語で説明文を生成するシステムの開発が容易になっている。しかし、汎用的な LLM, VLM に直接データを取り込んでも時空間を俯瞰した戦術を解説できず、状況の表層的な描写しか出力できない。また、映像の内容とは合致しない内容を出力するなどサッカーの分析に必要な情報抽出の仕組みが足りていない。

このような課題を踏まえ、本研究では、LLM を用いてサッカーのシュートシーンにおけるハイライト映像から解説文を自動生成し、その内容を可視化するシステムを開発する (図 1)。LLM への入力には、既存のコード [1] で試合映像から取得した選手のトラッキングデータやサッカーにおけるセオリーなどのヒューリスティックな情報を利用した。解説の可視化には LLM が出力した解説文から着目している選手はどれか、走った向きはどうかなどの可視化に

必要な情報を抽出し、図形や矢印を用いて機械的に可視化する。なお、試合の網羅的な解説は難しいため、本研究ではシュートシーンの解説に絞っている。



図 1. 出力のイメージ図。図下の文章は LLM が出力した解説であり、矢印は選手の進行方向、半透明の楕円は生じた空間を示す。

2 関連研究

IT を用いたサッカーの分析方法の多くは選手のトラッキングデータを用いて、独自の指標を定義し、選手やチームを定量的に評価している [2]。また、選手のつながりをグラフで表すことでグラフニューラルネットワークを構築し、得点につながりやすい動きを提案した技術 [3] が開発され、プロサッカークラブも活用している。トラッキングデータを用いず、試合映像を直接入力として用いて VLM に戦術的分析と解説を生成する手法 [4] も提案されているが、2 秒以上の動画に対してはハルシネーションを起こしやすいことが指摘されている。

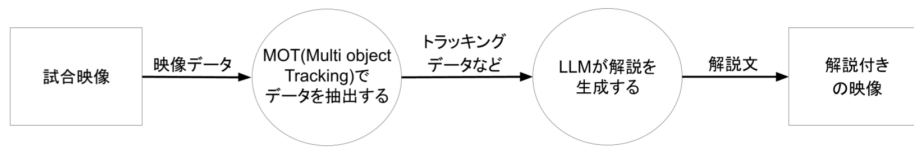


図 2. 手法の全体像。

3 手法

図2は手法の全体像である。まず、試合映像から選手などを追跡し、平面上に正射影する MOT(Multi Object Tracking) 技術 [1] を用いて、選手とボールの2次元座標を1フレームごとに取り出し、csvファイルに出力した。さらに、所属チームとゴールキーパー (GK) の特定に必要な情報を csv ファイルに加えた。次に、作成した csv ファイルとその中身の説明、コートの大きさ、そしてゴールの座標を ChatGPT5[5] に入力として与え、ボールと選手の座標を考慮し、「どのようなシーンか?」を解説するように指示した。結果を 4.1 節に示す。4.1 節に基づき、浮き球を蹴ったフレームと落ちてきたフレーム、ヘディングの競り合いがあったこと、「ヘディングの競り合いはボールの落下地点の予測が難しい」や「ボールの速度が変わる＝選手がボールに触れた」といったヒューリスティックな情報を LLM への入力に加え、攻撃がうまくいった要因も推論するようにプロンプトを変更した。その結果を 4.2 節に示す。

4 結果と考察

4.1 結果 1

出力として、選手とボールの座標のプロット図3が得られ、シュートシーンであること、攻撃側がどちらかのチームか、シュートがゴールに入らず GK に止められた可能性が高いことの3つが理解できていた。一方で、ボールの高さがデータに含まれていないため、ボールが浮いているのか分からないこと、ボールのスピードを考慮できていないことでシュート位置の推測が誤っていることの2つの課題があった。この課題を解決するために3章で述べたようにプロンプトを変更した。

4.2 結果 2

出力として、シュート時のシューターやボール、GK、ゴールの位置関係をプロットした図4と次の文章を得た。「(中略) 要するに (1) 浮き球で守備の予測と迎撃の精度を下げたこと、(2) ヘディングに対してセカンド回収の設計があったこと、(3) 競り合いのこぼれの不確実性を前提に、前向きをサポートがあったこと、この3つが噛み合い、前進する構造を保てたことが攻撃成功の要因です。」セカンド回収やこぼれ、サポートといった入力したシーンを戦術的に解説するためのキーワードが出力された。

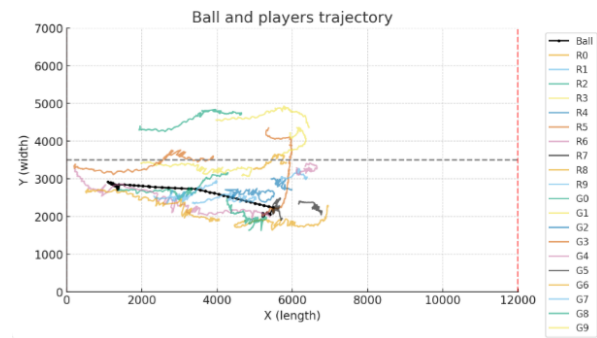


図 3. 選手とボールの座標のプロット図。ゴールの座標を [0,3500], [12000,3500] とし、コートの大きさを縦 7000, 横 12000 とした。

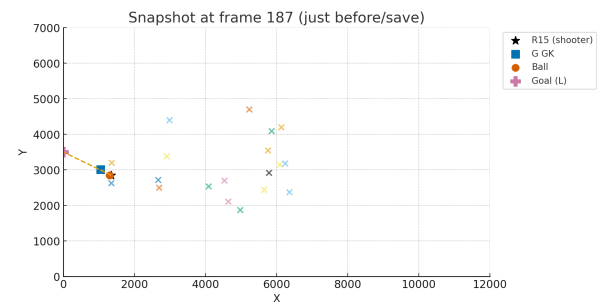


図 4. シュート時の選手、ボールの座標のプロット図。

5 今後の課題

本研究の目的達成のためにはプロンプト設計の改良、入力データの拡充、そして映像可視化手法の高度化が必要である。まず、プロンプト設計の改良として、専門用語をサッカー初学者にも理解しやすい表現へ変換する指示を加えるとともに、サッカーにおけるセオリーやヒューリスティックな知識を明示することで、生成される解説文の精度向上を図る。次に、入力データの拡充として、選手やボールの速度情報に加え、映像からボールの高さを推定可能な画像処理モデルを活用し、キックとヘディングを識別可能なシステムが必要である。さらに、映像可視化手法の高度化として、図1のような可視化を実現するために、生成された解説文から着目選手や走行方向などの可視化に必要な情報を抽出し、それに基づいて自動的に可視化する仕組みの実装が必要である。

参考文献

- [1] SkalskiP. Soccerai, 2024.
- [2] Javier Fernández. Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. 2019.
- [3] Zhe Wang, et al. Tacticalai: an ai assistant for football tactics, 2023.
- [4] Tiancheng Jiang, Henry Wang, Md Sirajus Salekin, Parmida Atighehchian, and Shinan Zhang. Domain adaptation of vlm for soccer video understanding, 2025.
- [5] OpenAI. Chatgpt. <https://chatgpt.com/>.

未来ビジョン

本研究の目的はサッカー中継の解説がサッカー初学者にとって理解しづらいという問題を解決することだが、その根底には、将来的にサッカーを日本の文化にしたいという筆者の目標がある。このように思ったのは高校時代に聞いたドイツのサッカー文化の話がきっかけである。ドイツでは多くの家庭で一家全員で試合のある週末にはスタジアムに行く、あるいはテレビの前で地元のチームを応援するという。彼らにとってそのことは非日常ではなく、生活の一部であり、チームが強かろうと

弱かろうと関係ない。そして、サッカーが別に好きでない人でもサッカーを家族との団らん場の場と捉えている。日本のサッカー文化がドイツのように発展していくには、より多くの人がサッカーをプレーしたり観戦したりすることが不可欠であり、そのためにサッカーの観戦体験を向上や観戦するうえでの障壁を取り除く、または緩和する研究を行っている。