

リアルタイム物体検出と視線追跡による運転時の視線評価システムの開発 Development of a gaze evaluation system in driving using real-time object detection and gaze tracking

タクツラ ジャガト[†]
Thakulla Jagat

平石 広典[†]
Hironori Hiraishi

1. はじめに

本研究では、運転中の視線評価システムを提案する。本システムは、Unity と YOLOv8x を用いてリアルタイムに物体を検出し、視線追跡センサーの監視下で被験者がどこを見たかを自動的に評価するものである。システムは見るべき領域を緑色のバウンディングボックスで、見てはいけない領域を赤色のバウンディングボックスでそれぞれ表示する。本システムにより、被験者の視線がどの領域に向けられたかに基づいて自動的にスコアを算出することが可能である。本研究の成果は、初心者ドライバーの視線行動を改善するためのトレーニング手法の開発や、安全運転の向上に貢献するものである。

2. 視線評価システム

図 1 は、本研究で設計された視線評価システムを示している。本システムは、事前に録画された車内映像を再生するコンピュータ上で動作し、ユーザーがその映像を見ることで視線の評価を可能にする。視線計測センサー (Tobii Eye Tracker 5, Tobii Technology) はモニター下部に設置され、ユーザーの視線を計測し、映像中のどの領域を見たかを評価する。見るべき領域 (例: 自動車, 信号機, 標識など) は運転中に注目すべき対象を示し、見てはいけない領域 (例: 広告看板, 駐車中の車) は運転中の注意を妨げる対象を示す [1]。本システムは、運転者が適切に見るべき領域に注意を向け、見てはいけない領域を回避できているかを評価する。評価結果として得られるスコアは、モニターの左上に表示される。

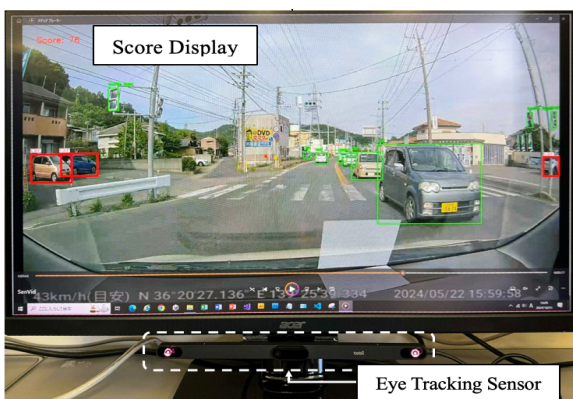


図 1 視線評価システム

YOLOv8x[2]というリアルタイム物体検出が可能な深層学習モデルを採用することで、本システムは映像に基づいて見るべき領域および見てはいけない領域を自動的に設定

する。システムの実装には、Unity ゲーム開発エンジンを用いた。運転映像中に YOLO によって検出された物体と、Unity から取得した視線データをもとにスコアを算出する Python プログラムを作成した。図 2 は、評価システムの開始から終了までの一連の流れを示す。

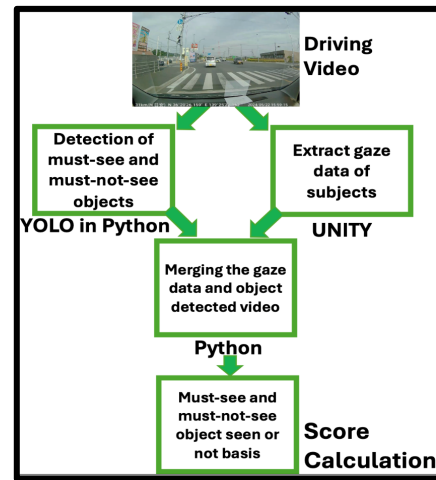


図 2 システムの流れ

運転中の映像は、車両前方のフロントウィンドウ中央に取り付けられたドライブレコーダーを用いて撮影された。ドライブレコーダーのカメラは、車両の前方視界を適切に捉えるように調整されており、その様子を図 3 に示す。取得された映像は本実験に使用され、「左側処理」と「右側処理」の 2 つのプロセスに分けて解析された。これらはそれぞれ、YOLO in Python 処理と、Unity による表示処理に対応している。



図 3 ドライブレコーダーで動画撮影

図 2 の左側のプロセスでは、Python 上で YOLOv8x を用い

[†] Ashikaga University, Faculty of Engineering

て実験映像に自動アノテーションを施し、YOLO が検出できなかった物体については手動アノテーションおよびモデルの再学習し認識できるようにした。その結果、見るべき領域および見てはいけない領域を含んだ映像が得られた。右側のプロセスでは、被験者に Unity 画面上で再生される実験映像を見てもらい、その際の視線は Tobii アイトラッキングセンサーにより取得された。その後、左側プロセスで得られたアノテーション結果と、右側プロセスで取得された被験者の視線データとを、Python プログラムを用いて統合した。この視線評価システムは、被験者が見るべき領域または見てはいけない領域を見ていたかどうかを判断し、式 (1) に基づいてスコアを算出する。

$$\text{スコア} = 100 \times \left(\frac{n}{N} + \frac{M-m}{M} \right) \div 2 \quad (1)$$

このスコアは、見るべき物体を見た割合と見てはいけない物体を見なかった割合の平均を表している。ここで、 N は実験映像に含まれる見るべき物体の総数、 n は被験者が実際に見た数を表す。同様に、 M は見てはいけない物体の総数、 m は被験者が実際に見た数を表す。これらの値は、各物体が一度だけカウントされるように Python のセット機能を用いて自動的に計算された。トラック ID をベースとするこのロジックにより、注視の評価が精密かつ頑健となり、被験者が映像全体を通して特定のターゲットをどの程度見ていたか、あるいは見抜けたかを公正に評価することができる。

3. 評価実験

本システムは、視線の動きに基づいて運転者が特定の対象に注意を向けたかどうかを判定し、それが見るべき領域（緑色のバウンディングボックス）か、見てはいけない領域（赤色のバウンディングボックス）かに応じて、自動的にスコアを算出する。

自動スコアの比較は、熟練者 2 名と初心者 3 名の合計 5 名の被験者に対して実施された。熟練者には、運転歴 13 年以上の 30 代の大学院生と、運転歴 30 年以上の大学教授が含まれた。初心者は全員 20 代の大学生であり、運転歴は 1 年未満であった。テストコースとして大学周辺の約 1km を撮影した車載映像（1 分間）を使用して実験を実施した。

物体検出には、2 つの視覚クラスを含むカスタムデータセットで学習させた YOLOv8 モデルを使用した。データのアノテーションには CVAT (Computer Vision Annotation Tool) を使い、クラス 0 は見るべき領域（車両、オートバイ、信号機、歩行者、自転車）、クラス 1 は見てはいけない領域（広告看板、駐車車両）をそれぞれ表す。クラス 0 は計 429 枚の画像、クラス 1 は 347 枚の画像にアノテーションが行われた。フレーム間での物体の一貫性を保つため、Deep SORT トラッキングアルゴリズム [3] を使用した。

さらに、周辺視野を反映し人間の注意のモデルをより現実的に表現するため、長方形の視線領域を導入した。この領域は、13 ピクセル幅と 36 ピクセル高さを有する固定サイズの矩形として設定され、効果的な注視範囲が広がり、より包括的な視線と物体のインタラクション評価が可能となった。表 1 は、5 人のドライバーに対する自動スコア

リング結果を示している。B1, B2, B3 は初心者ドライバーを、E1 および E2 は熟練者ドライバーをそれぞれ表す。

なお、追跡中に一つの課題が確認された。それは Deep SORT による誤検出、いわゆる「ゴースト」バウンディングボックスの生成である。これは、最近の検出がない状態でもモーション予測によってボックスが継続的に生成されることにより発生し、誤った視線の対応やスコアへの不正確な影響を引き起こすことがあった。この問題への対応として、追跡処理の前に信頼度しきい値によるフィルタリングを実施した。具体的には、YOLO の検出結果のうち、信頼度スコアが 0.5 を超えるもののみを Deep SORT トラッカーに渡すようにした。このフィルタリングにより、誤検出されたトラックが削減され、自動スコアリングシステムの精度が向上した。

表 1 スコア比較

Subjects	N	n	M	m	Score
B1	191	13	196	9	51.11
B2	191	18	196	16	50.63
B3	191	22	196	33	47.34
E1	191	25	196	11	53.74
E2	191	30	196	5	56.58

おわりに

本研究では、YOLOv8x を用いたリアルタイム物体検出と Deep SORT アルゴリズムによる追跡を組み合わせた、運転中の視線追跡に基づく評価システムを設計した。本システムは、被験者の視線が運転において見るべき対象に向けられているかどうかに基づいて自動的にスコアを算出することが可能である。このスコアに基づき、熟練ドライバーの視線パターンを参考にして初心者ドライバーの訓練を行うことを目的としている。今後は一分以上の映像や異なる運転ルートをを用いた実験を行う予定である。これにより、初心者と熟練者の視線行動における違いがより明確に現れることが期待される。さらに、今後の課題としては、周辺視野のモデリング精度の向上および被験者の多様性の拡充に取り組み、自動スコアリングの精度と汎用性を高めていく予定である。

参考文献

- [1] David Vetturi et al., "Use of eye tracking device to evaluate the driver's behavior and the infrastructures quality in relation to road safety", *Transportation Research Procedia* 45 (2020) 587–595, 2020.
- [2] Sharma, A., & Kapoor, R. (2023). A Comprehensive Review of YOLO Architectures: From YOLOv1 to YOLOv8. *arXiv preprint arXiv:2304.00501*.
- [3] N. Wojke, A. Bewley, and D. Paulus, "Simple Online and Realtime Tracking with a Deep Association Metric," in *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, 2017, pp. 3645–3649.