機械系研究インターンシップ報告書

物体検出アルゴリズム YOLO v8 を用いて自動運転の発進判断に用いるフリースペース情報を直接学習し検出させる手法の提案と改良

機械情報工学科 2 年　椿 道智
指導教員　中野 公彦 教授

1　研究目的

本研究の目的は東京大学柏キャンパスの周辺を走る自動運転バスが右折する際の発進判断に用いる，対向車線に自動車や歩行者がいないという「フリースペース情報」を取得手法の開発である．柏キャンパス周辺の T 字交差点の信号機に設置された株式会社 IHI 様の路側器から取得された動画データのみを用いてフリースペース情報を取得する．図 1 の黄色と赤の網掛け部がフリースペースであり，この部分に自動車や歩行者がいないことを確認することができれば，自動運転バスが右折できる．

図 1: 路側器によるデータ収集の概要

なお，本研究における路側器のデータの収集と解析は，東京大学倫理委員会の承諾を得た方法に基づいて行った (No. 17-14).

2　関連手法

フリースペース情報を取得する手法としては，これまで，物体検出アルゴリズムを用いて自動車や自転車，歩行者といった物体の検出を行い，特定のエリアに何も物体が検出されなかった場合，それをフリースペースと定義することが考えられてきた.（特許第 7361482 号「画像処理装置、撮像装置、移動体及び画像処理方法」京セラ）

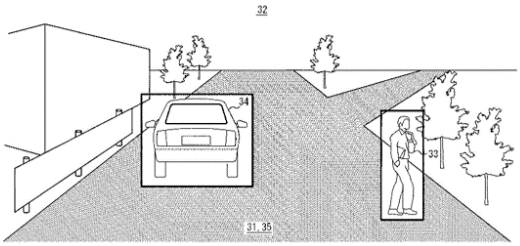

図 2: これまで考えられてきた手法（京セラ）

3　提案手法

本研究では，この手法とは反対に，YOLO v8 などの物体検出アルゴリズム[1]に道路のフリースペース (アスファルト) 自体を直接学習させれば,「物体」の検出を介さずとも直接フリースペース情報を取得できるのではないかと考えその手法を提案する.

また，路側器データのうち LiDAR 情報を用いず，動画データのみを用いてフリースペース情報を取得できることから，本手法は，自動運転技術の普及において，低コストで実現可能な手法であるといえる.

[1]物体検出アルゴリズムは文字通り「物体」があることを検出するアルゴリズムである．YOLO の場合は (1) 信頼度（物体がある場合を 1, 背景の場合 0 とした連続値）による物体の有無の検出と (2) 何が移っているかの情報（各物体のクラスに対する確率）の (1) と (2) の結果をマージすることによって物体検出を行っている.

4　実験 1　手法の実現可能性の調査

まずはじめに提案手法の実現可能性を調査する目的で実験を行った.

4.1　前処理について

株式会社 IHI 様よりご提供いただいた 10 分間の動画ファイル (.avi) のファイル形式を.mp4 に変換し，フリースペース情報を得るのに必要な対向車線の部分だけをクロップ（画面の切り取り）した．次に，クロップした動画ファイル 600 秒を 2 秒に 1 回静止画 (.png) として保存し，学習や評価に使う 300 枚の画像ファイルを生成した.

4.2　アノテーションと学習

前処理で取得した画像ファイル 300 枚に対してアノテーション作業（アンカーボックスとラベルの付与）を行い，学習データを作成した．アノテーションには専用ツール「labelImg」を用いた．ラベルは，Google Map の距離測定機能（図 4）による測定と，実験画像上の街路樹など特徴的な部分を対応させ，ものさし（図 3）を作成し，その基準に基づき,「0m」（＝フリースペースなし）「10m」「20m」「40m」「60m」「80m」「100m」「150m」「200m」の 9 種類のいずれかのラベルを与えた.

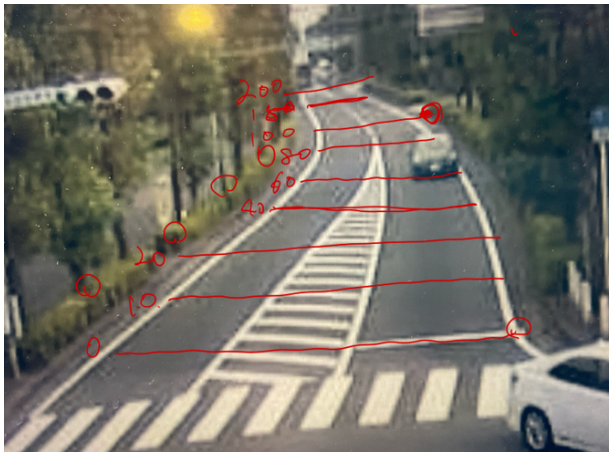

図 3: 距離のものさし（基準）

図 4: Google Map による距離測定

また，従来の物体検出では背景となるようなただの黒いアスファルトを学習させただけではすべての道路を検出してしまい，フリースペースの大きさを検出できないと考えた．よって，アスファルトを，物体として認識させるために，アンカーボックスのつけ方を工夫した（図 5）．自動車と停止線で挟まれる平面を物体として認識させようと考え，アンカーの下端を道路の停止線やや下の図 5 の★におおむね固定し，上端を自動車の全面が道路に落とす影の直下になるようにして，アンカーボックスは基本的に高さのみ変えた．この高さが，フリースペースの大きさに対応する.

図 5: アンカーボックスのつけ方の工夫

アノテーションを行った画像でデータセットを作成し，Ultralytics 社の YOLO v8 という Python ライブラリを用いて物体検知の学習を行った．学習は Google Colab の GPU 上で行い，学習したモデル（重みファイル）をローカルに落として，動画ファイルに対して，フリースペースが検出できるかなど検証をした.

4.3　結果・課題点

フリースペースを検出でき，提案手法の実現可能性が示された（6 上）．停止線上を自動車が通過する場合に検出できないなどの課題点が挙がった（6 下）．

図 6: 実験 1 結果（動画より抜粋）

5　実験 2 データセットの成型

5.1　目的・概要

実験 1 の結果より，提案手法の実現可能性が示された．実験 2 以降では，本手法のモデルの課題点の改善及び精度の評価，安全性の検証を行う．実験 2 では，まず実験 1 で用いたデータセットの構成を整えた．図に示すように 600 秒間の動画計 300 枚の画像のうち時系列順にはじめの 180 枚を学習用データセット，次の 60 枚を検証用データセットとし，残りの 60 枚をテスト用データセットとした．これは，out of time validation という手法に基づいており，学習に用いたデータセットとは異なる時系列のデータセットを用いてモデルの精度を評価することで，シナリオを学習してしまうことによる精度の過大評価を防ぎ，モデルの汎化性能を評価することができる．

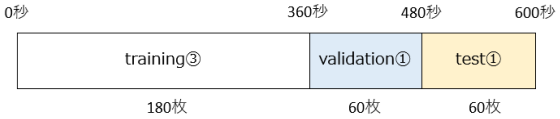

図 7: データセットの構成

5.2　モデルの精度

機械学習における精度の指標として，recall（再現率）と precision（適合率）がある．recall は，

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

であり，precision は，

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

である．ここで，TP は True Positive，FP は False Positive，FN は False Negative である．

ラベルがフリースペース有無（検知・非検知）の 2 値ならば，FN が起こると重大事故につながるため，recall は事故原因の指標といえる．一方で，FP が起こると，フリースペースが無いと判定されてしまうため，precision は交通の円滑性快適性の指標といえる．ただし，今回はラベルがフリースペースの距離の指標になっているため，False Negative は，フリースペースを検出できなかった場合だけでなく，「実際とは違うラベルを予測した場合」も含むことになる．よって，recall の低さが安全性の評価に直結するとは一概に言えないことに注意が必要である．なお，本研究の安全性指標については実験 4 で定義する．

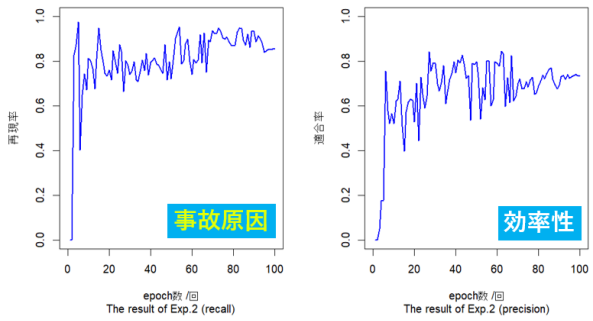

図 8: 実験 2 のモデルの精度

5.3　結果・課題点

実験結果の抜粋画像を図 9 に示す．実験 1 と比べて学習データ数や epoch 数が増えたため，学習モデルの精度

が向上し，実験 1 で課題となっていた停止線上を自動車が通過する際の検出漏れが改善された．

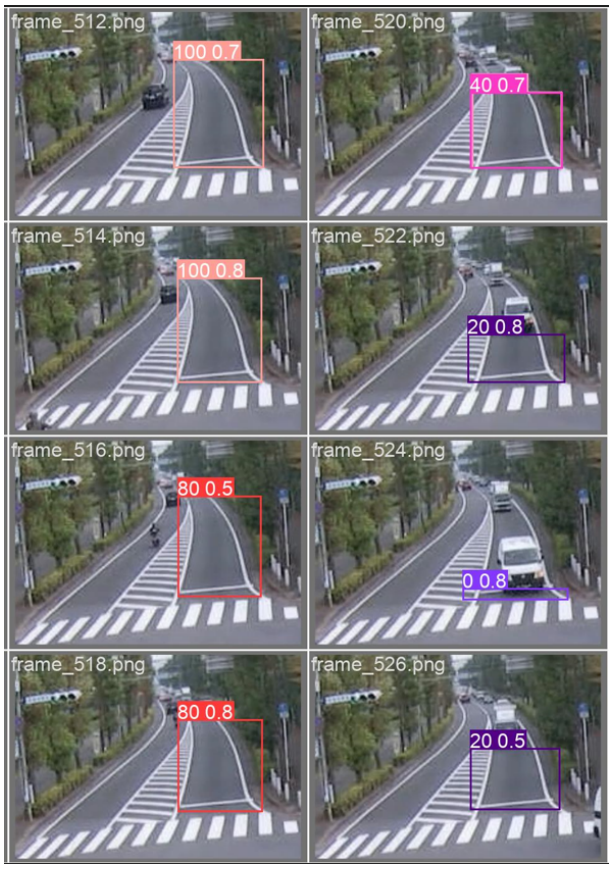

図 9: 実験結果（検証データ／抜粋）

しかし，信号無視をして女性が停止線と横断歩道の間を歩いている（交通違反をしている）ときに，フリースペースが無いのにフリースペース有と誤検知してしまうという課題点が残った．このようにフリースペース無なのに有と誤検知してしまうことは自動運転技術としては欠陥であり，モデルの改善が必要だと考えた．

6 実験 3 False Negative の改善

6.1 目的・概要

実験 2 での課題点を改善するために，モデルの改良を行った．実験 1 では，信号無視をして女性が停止線と横断歩道の間を歩いている（交通違反をしている）ときに，フリースペースを 0m と表示できなかった．これは重大事故につながりうるよくない誤検知である．よって，本実験では，このような誤検知を改善するために，学習データセットにこの女性が停止線と横断歩道の間を歩いている様子を含む画像を集中的に追加し，再学習を行った．（図 10）

図 10: 信号無視をする女性

6.2 改善手法

信号無視をする女性が停止線の前を通過する 2.6 秒間について 0.1 秒刻みで改めて静止画を生成し，アノテーションを行った．実験データ中にこの信号無視のシナリオが含まれるのはこの 1 回のみであるため，この部分については out of time validation の手法を用いることができない．ここでは，学習データとして 18 枚，検証データとして 4 枚，テストデータとして 4 枚を実験 2 のデータセットに単純に追加した

6.3 モデルの精度

モデルの精度は，実験 2 のモデルの精度と比較して評価した．結果は図 11 の通りである．なお，先述した通り，recall の低さが安全性の評価に直結するとは一概に言えないことに注意が必要である．

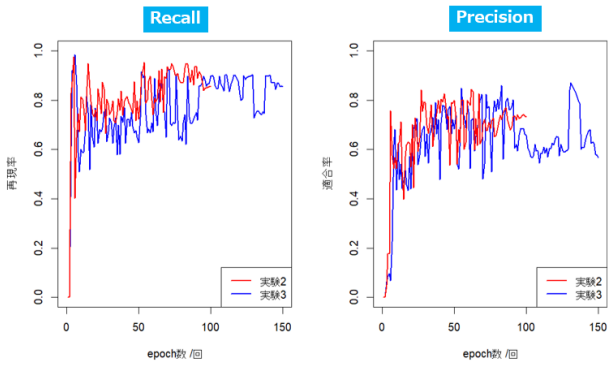

図 11: 実験 3 のモデルの精度 (実験 2 との比較)

6.4　結果・課題点

追加したテストデータ 4 枚についてすべてでフリースペースを 0m と検出できた.（図 12）動画でも渡り始めから渡り終わりまで常に 0m を検出できていることを確認した．よって，よくない誤検知が改善された.

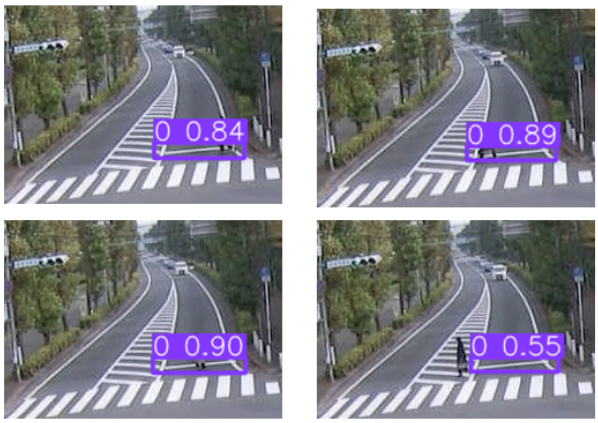

図 12: 実験 3 結果

ただ，実験 3 の結果を見ると，200m の次が 100m になることが多く，150m のラベルの検出がほとんどできないないことが分かった．これは，ラベルの付け方に問題があると考えられる．実際に，学習データに含まれるクラスごとの物体の割合を確認すると，図 13 の通りとなった.

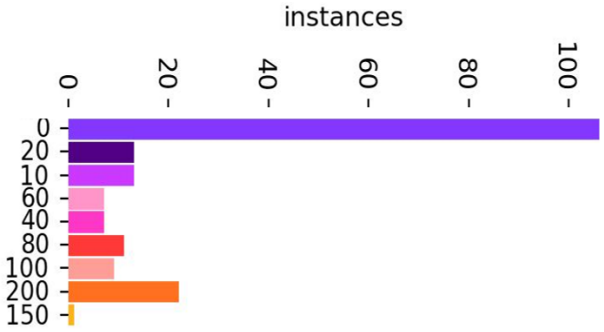

図 13: 実験 3 のモデルの精度 (実験 2 との比較)

0m のクラスの割合が他のクラスに比べて突出していることが分かる．均等に 2 秒おきにサンプリングした実験データに対して out of time validation を適用したので，インスタンス数にムラが出るのは当然であり，信号停止などを考えれば，0m のインスタンスが多くなるのも当然であると考えられる．また，0m はフリースペースが無いという意味で，事故防止という観点から考えれば重要な情報なので多くても差し障りない．一方で，150m のインスタンスが少ないことは，実際には 150m のフリースペースがあることが少ないわけではないので，このようなラベルの付け方は適切でないと考えられる．この点が，次の実験に向けた課題点となる.

7　実験 4　安全性の定量的評価及び 150m のラベルの追加学習

7.1　目的・概要

実験 3 での課題点を改善するために，実験 4 では，150m のラベルの追加学習を行った.（図 14）

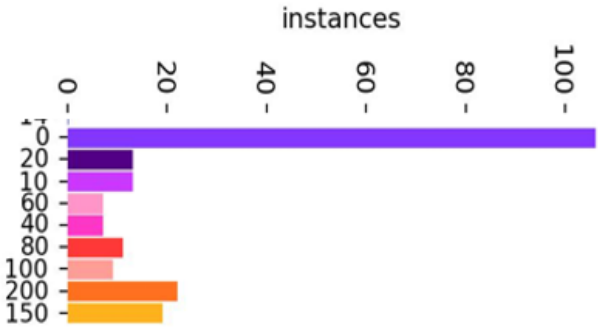

図 14: 150m の追加学習後のデータの割合

また，本実験が研究の最終実験であるため，ここで改めて自動運転における実用を考慮した安全性の定義を行い，定量的評価を行った.

7.2　安全性の定量的評価

自動運転技術においては，交通事故を防ぐことが最も重要な課題である．よって，本研究においても，安全性の評価を行うことが重要である．フリースペース情報の取得に関しては，フリースペースを実際よりも小さく見積もることは，交通の安全性や円滑性では問題だが，安全性の観点では問題とならない．一方で，フリースペースを実際よりも大きく検出してしまうことは，事故につながる重大な過誤である．よって，本研究における安全性の指標として以下のとおり「危険率」を定義した.

$$\text{危険率} = \frac{\text{実際よりも大きくフリースペースを予測した件数}}{\text{全テストデータ数}}$$

テストデータに対して，この危険率を求めた．実験可能なテストデータの件数は 64 件であり，そのすべてについて検証した．与えたラベルよりも大きな値のラベルを予測した件数は 2 件であり，危険率は 3.1%であった.

このうち 1 件は 150m のテストデータについて 200m と誤った例で，これは実験データの画質の粗さや，遠近法による遠くの観測の限界に起因する過誤である．本件は提案手法の限界といえる．もう 1 件は 40m のラベルを 60m と誤ってしまった事例でもっとも有害な事例といえる．(図 15)

図 15: 40m のラベルを 60m と誤ってしまった事例

自動運転における実用に耐えうるためには，危険率はほぼ 0%であることが望ましいと考えられ，提案手法はさらなる改善の余地があるといえる．例えば,「関連」で提案する手法や LiDAR 値などを用いてシステムの冗長化を行うことで，自動車に伝達するフリースペース値ベースで危険率を低減することができると考える．また，今回の実験データには路側を走る二輪車が含まれていないため，二輪車が含まれる実験データに対しても同様の安全性評価を行うことが望ましい．

8 関連 バウンディングボックスによるフリースペース情報の数値化

自動運転に用いることを想定すると，フリースペース情報を数値化することが望ましい．提案手法では，バウンディングボックスがちょうど車の前方の陰影から停止線までをうまく囲えていることが動画で確認された．よって，バウンディングボックスの高さを用いてフリースペース情報を数値化することを試みた．

バウンディングボックスの高さから遠近法を考慮してフリースペースの大きさを推定する逆問題を考えればよく，遠近法により，間隔は遠くなるほど小さく表れ，その間隔は反比例の関数 $(1/x)$ なので，その地点までの距離を表す関数はその積分の $\log x$ である．この逆関数は e^x である．これにノイズや比例定数などを考慮して，

$$f = ae^{bh} + \epsilon$$

上記の非線形モデル f を用いて，バウンディングボックスの高さ h からフリースペースの大きさを推定することを試みた．なお，a, b, ϵ は未知定数であり，実験データによって最小二乗法などから求められる．

本手法により求められたフリースペース情報と物体検出から得られたクラス情報とを組み合わせることでより正確にフリースペース情報を得ることができると考えられる．

9 まとめ

従来は自動運転では路側器から得られた動画を用いて物体検出をしたり，LiDAR 情報などで，物体が無いとされた場合をフリースペースとしていた．しかし，本研究ではその手法とは反対に，物体ではなくフリースペースそのものを画像学習させ，フリースペースを直接検出する手法を提案した．実験 1 の試行調査で，提案手法の実現可能性を示されたので，実験 2 ではモデルの精度を評価し，実験 3 では有害な誤検知を改善し，実験 4 では安全設計の観点から安全性の定量的評価を行った．また，関連としてバウンディングボックスによるフリースペース情報の数値化を試みた．今後は，実験 4 で示した有害な誤検知に対するシステム上のアプローチの必要性や自転車が路側帯を通行するシナリオの場合どうなるかなどの課題点を改善し，さらなる安全性の評価を行うことが望ましい．

参考文献

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. (2015). You Only Look Once: Unified, Real-Time Object Detection. https://doi.org/10.5334/dsj-2019-009.

[2] YOLO v8 技術資料 https://docs.ultralytics.com/ja

[3] 特許情報プラットフォーム「特許 7361482」京セラ https://www.j-platpat.inpit.go.jp/c1801/PU/JP-2020-162013/11/ja