

列車前方監視用の AI の判断ミスの要因推定手法

長峯 望* 合田 航*

Method for Estimating Causes of AI Decision-making Errors in Train Forward Surveillance

Nozomi NAGAMINE Wataru GODA

In the railway industry, there is a growing trend of using AI and camera images to automate various tasks that were previously performed visually. However, when considering the future application of AI in safety-critical operations such as autonomous driving, it is essential to investigate the causes of AI decision-making errors. Therefore, we have developed a tracing method for train-forward surveillance systems that estimates the causes of AI decision-making errors, such as missed detections, at three stages: input images, AI structure, and training data. This paper reports the developed method for estimating the causes of AI decision-making errors and its verification results.

キーワード：列車前方監視，人工知能，物体検知，判断ミス

1. はじめに

近年の機械学習，特に深層学習の進歩により，AI（Artificial Intelligence:人工知能）は画像認識，自然言語処理，予測分析等，複数の領域で高度な用途にも活用可能な性能を発揮するようになった。これに伴い，医療，製造，金融等，教育等様々な分野において AI の利活用は進んでおり，少子高齢化に伴う労働人口の減少等の社会的課題の解決に向けた強力なツールとして期待が高まっている。鉄道業界においても，運転士や保守作業員の確保が以前より困難になっており，業務の効率化・省力化に向けて，鉄道事業者やメーカー等が AI の研究開発並びに導入に取り組んでいる。

一方で，AI（特に深層学習）は，判断プロセスが不透明であり，AI がどのようにして特定の出力や決定に至ったかを理解することが難しいという課題も存在する。これは，AI が判断ミスを起こした際に，その説明や根拠の提示が難しいことを意味する。

この課題は，総務省の「AI 利活用ガイドライン」¹⁾においても言及されており，「AI を利活用する際の社会的文脈を踏まえ，どのような説明が求められるかを分析・把握し，必要な対応を講じること」と示されている。特に，鉄道の安全に関するシステム等においては，判断ミスが発生した際に，原因究明と対策を正しく説明し，透明性を確保できる技術を開発することは重要であると考えられる。

透明性の高い，いわゆるホワイトボックスな AI を開発するという方向性も模索されているが，画像処理に関する深層学習等の AI においては，現時点までで有効かつ実用的な手法は提案されていない状況である。さらに，透明性を有さない AI が他分野で社会実装が始まっている実情も踏まえると，AI の透明性を一定程度確保する枠組みが，現実的な対策として必要となることが想定される。

このような背景から，本稿では，AI を用いて列車前方の支障物を検知する前方監視システムを対象とし，AI による支障物の見逃し等の判断ミスの要因を推定する手法について述べる。前方監視システムは，乗務員の支援のほか，列車の自動運転への適用も期待されており，安全に関わる重要な用途であるため，対象として選定した。

この手法を用いることにより鉄道事業者は，AI を用いた前方監視システムの支障物の検知性能を導入

* 情報通信技術研究部 画像解析研究室

前の検査や、導入した前方監視システムが支障物を見逃した際の原因究明および事後対策の検討を行うことが可能となる。

2. 列車前方監視における AI の活用

本稿では、図 1 に示す通り、列車前頭にセンサ類を設置し、センサから取得される情報を基に AI が判断を行うことで前方を監視するシステムを想定する。将来的に列車前方監視システムが実用される場合には、カメラ、LiDAR、ミリ波レーダー等のセンサが、単一またはそれらの組み合わせにより構成されることが考えられる。このようなセンサ類の中でも、可視光カメラは、安価で設置しやすく、さらに鉄道用のドライブレコーダ等の用途で普及しており、撮影される画像に対して AI が認識を行う技術も発展している。そこで、現時点での普及状況及び、今後の活用見込みを勘案し、センサとして可視光カメラを対象とし、列車の前方監視を行う AI (以下、前方監視 AI) の判断ミスの要因推定手法の開発を行うこととした。また、想定する前方監視 AI は、支障物の種類と位置を推定する物体検知 AI とした。

3. AI の判断ミスの要因推定手法

3.1 手法の全体像

可視光カメラで列車前方を監視する場合の危険事象として、支障物を見逃しがあげられる。そこで、前方監視 AI が支障物を見逃した場合の要因の推定を行う手法を開発した。開発した手法は、入力画像の検査、AI 構造の検査、学習データの検査に関する 3 段階のアプローチにより要因の推定を行う。提案する要因推定手法の全体フローを図 2 に示す。

要因推定手法の全体フローは、3 段階の検査のいずれかで要因を推定できれば、その要因をもとに改善に向けた対策の検討が行える。そこで、要因が推定された際に対策が容易である順に検査を行うこととした。これは判断ミスが起きた際には、迅速な対策が必要であることを想定している。フローの最初に位置する入力画像の検査では、AI そのものではなく、主にカメラの撮影方法等に起因する要因の推定を行う。撮影方法等に起因する要因の対策は、カメラ設定変更等が主体となるため、AI そのものが要因である場

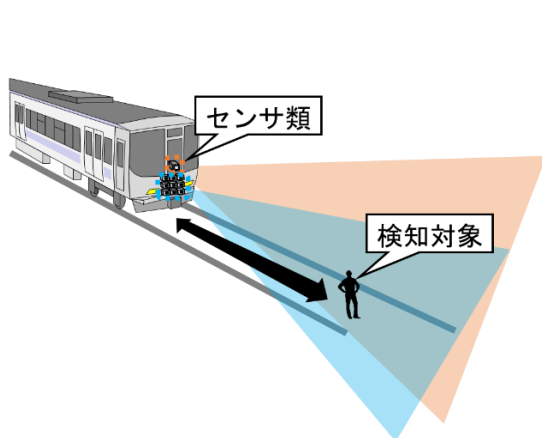


図 1 想定する列車前方監視システム

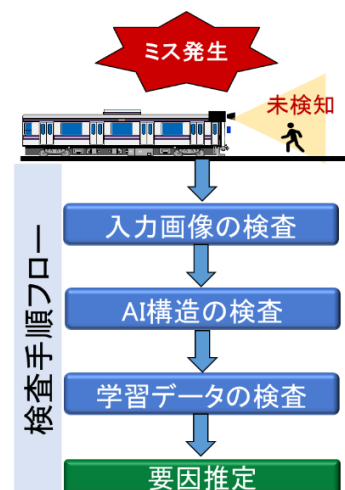


図 2 AI による判断ミスの要因推定の全体フロー

合に比べ、対策が容易である可能性が高いと考え、初めのステップとして検査を行うこととした。

次に、AI 構造の検査では、画像の認識方法に起因する判断ミスに関する検査を行う。AI は構造の違いによって画像に対する認識方法が異なるため、この認識方法に起因する判断ミスが発生する可能性がある。認識方法が要因の場合、AI の構造の変更・追加やチューニングを行うことで対策が可能であると考えられるため、2 番目のステップとして設定した。

最後に、学習データの検査では、学習データの分布や特徴から、判断ミスの要因となるデータの偏り（バイアス）を調査し推定する。これらの要因を解消するには、偏っている学習データを間引くこと（アンダーサンプリング）や足りていない学習データの追加等が考えられる。特に、学習データの追加は、学習用データ収集と作成が伴い、作業量が多く時間がかかることが想定される。したがって、学習データの検査を最後の 3 段階目のステップとして位置付けた。

次節以降で、それぞれの検査の詳細について述べる。

3.2 入力画像の検査

入力画像に起因した判断ミスの検査手法を図 3 に示す。図 3 左側に示す通り、見逃し等の判断ミスが発生した画像に対して、各種カメラ設定の変更を再現する画像処理を行い、再度、前方監視 AI による検知を行うことで判断が改善されるかを確認する。その結果、改善された項目があれば、それを判断ミスの推定要因とする。具体的な項目として、回転、位置、コントラスト、ホワイトバランス、彩度、明るさ、ピンボケ、ブレ、焦点距離（光学ズーム倍率）、ノイズの 10 項目を設定した。設定した 10 項目は、これまで我々の AI による物体検知に関する研究で観測できた判断ミス要因から、カメラ設定の変更等で対策が可能なものを選定している。また、各項目については設計変数であり、必要に応じて適宜新たな検査項目の追加が行える。

具体的な設定項目（判断ミス要因）と、それを再現する画像処理方法、その対策を照らし合わせたものを表 1 に示す。判断ミス要因からカメラ設定の変更等実行可能な対策までの対応関係が示されている。この対応表に沿って検査を行う。この検査により判断ミスの要因が推定された場合には、カメラ設定が要因であると言えるため、表 1 の右側の列に示された対策を施すことが推奨される。

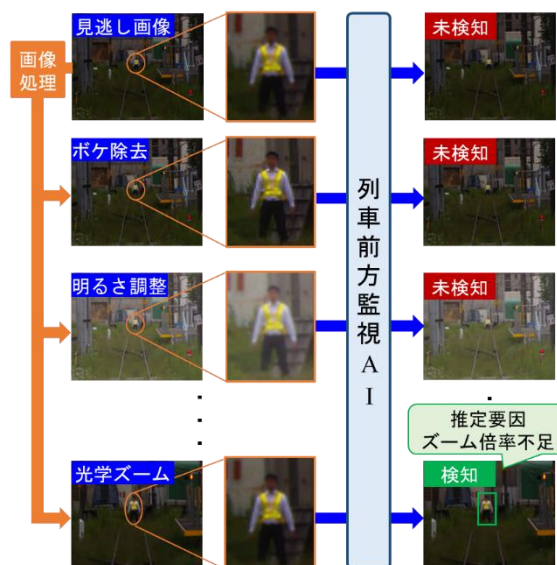


図 3 入力画像の検査による判断ミス要因推定

表 1 設定項目（判断ミス要因）への対応表

設定項目(判断ミス要因)	再現する画像処理方法例	実行可能な対策
ピンボケ	ピンボケ除去フィルタ	ピント調整
ブレ	ブレ除去フィルタ	シャッタースピード変更 列車速度変更
明るさ	HSV 色空間における V パラメータの調整 ガンマ補正	シャッタースピード変更 被写界深度変更 沿線に光源を設置
コントラスト	ヒストグラム平坦化	コントラスト設定変更
ホワイトバランス	色温度調整	ホワイトバランス設定変更
彩度	HSV 色空間における S パラメータの調整	カメラの彩度設定変更
回転	アフィン変換	カメラ設置位置の変更
位置	アフィン変換	カメラ設置位置の変更
光学ズーム	クリッピング+超解像	カメラレンズの変更
ノイズ	メディアンフィルタ バイラテラルフィルタ	ISO 感度変更 シャッタースピード変更

3.3 AI 構造の検査

深層学習をベースとした物体検知アルゴリズムは、ニューラルネットワークのネットワーク構造（以下、AI 構造）の違いによって、画像の認識方法が異なることが知られている²⁾。AI 構造に関する検査では、このような認識方法の違いが要因となる判断ミスを推定することを目的としている。

AI 構造の違いとしては、CNN（Convolutional Neural Network）や Transformer 等のベースとなるアーキテクチャの違いやニューラルネットワークの中間層の深さ等の違いがあげられる。アーキテクチャの違いは認識方法の違いを示す。例えば、CNN は、画像に対して空間的な周波数を考えたときに空間的に細かい特徴量を重視するハイパスフィルタのような特性を持ち、Transformer では、空間的に大雑把な特徴量を重視するローパスフィルタのような特性を持つことが報告されている²⁾。また、ニューラルネットワークの中間層の深さはより複雑な特徴量の表現性を示す。

また、入力解像度も特徴量の算出や認識に影響を与える。入力解像度は、大きいほど細かい特徴量まで考慮することができ、逆に小さい場合には細かい特徴量を捉えられないという傾向がある。

これらの傾向が判断ミスの要因になっていることを確認するため、提案する検査手法では、同一の学習データによって作成された入力解像度と AI 構造が異なる前方監視 AI を作成し、判断ミスが起きた画像に対して検知を行わせる。これにより、検知できた AI が存在する場合には、判断ミスが起きた条件（AI 構造と入力解像度の組み合わせ）と検知ができた条件の差異が判断ミスの要因であると推定できる。AI 構造の検査による判断ミス要因推定のイメージを図 4 に示す。なお今回の検証では、AI 構造と入力解像度の組み合わせは、表 2 のように設定した。ここで用いる AI 構造の選定は、検証を実施する際に、状況に応じて適宜変更可能である。この検査により判断ミスの要因が推定できた場合は、検査手法によって改善

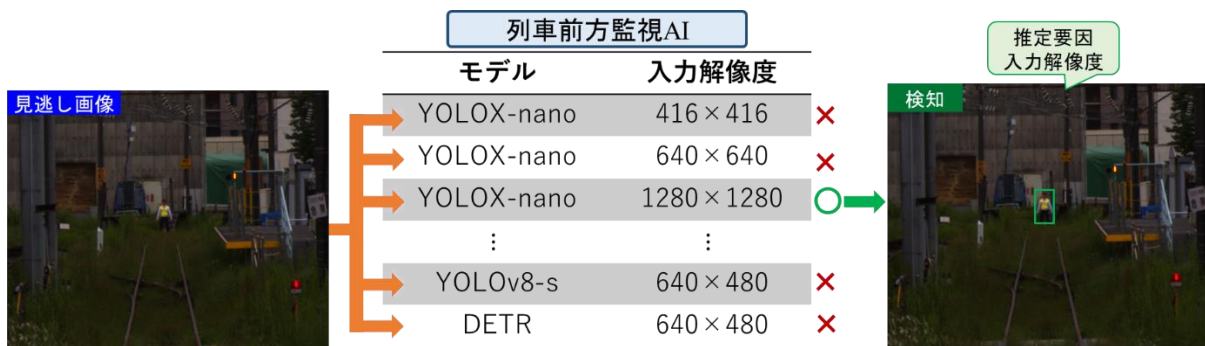


図 4 AI 構造の検査による判断ミス要因推定

表2 モデル（ネットワーク構造）と入力解像度

モデル(ネットワーク構造)	入力解像度
YOLOX-nano ³⁾ (CNN)	(416,416), (640,640), (1280,1280)
YOLOX-s ³⁾ (CNN)	(416,416), (640,640), (1280,1280)
YOLOX-l ³⁾ (CNN)	(416,416), (640,640), (1280,1280)
YOLOv8n ⁴⁾ (CNN)	(640,480), (1280,960)
YOLOv8s ⁴⁾ (CNN)	(640,480), (1280,960)
YOLOv8l ⁴⁾ (CNN)	(640,480), (1280,960)
DETR ⁵⁾ (Transformer)	(640,480), (1280,960)
Co-DETR ⁶⁾ (Transformer)	(640,480), (1280,960)

された AI 構造（検知できた AI 構造）を用いて前方監視システムを再構築することが対策として推奨される。

3.4 学習データの検査

学習データに起因する AI の判断ミスとして、データの分布が要因となる場合が複数報告されている。例えば、人材採用システムに用いる AI において、女性に対して差別的な評価をしてしまうことにより、運用を取りやめた事例⁷⁾がある。これは過去 10 年間に提出された履歴書のパターンを学習データとしており、そのほとんどが男性であるというデータの偏りによって発生したことが知られている。他にも、今回の対象と同じ画像に関する AI についても、画像認識ソフトウェアにおいて撮影した写真を分類し、保存する機能でアフリカ系アメリカ人の人物がゴリラとして誤って分類されてしまう事例⁸⁾がある。これについても、学習データに、他人種に比べてアフリカ系の人物のデータが少なかった等のデータの偏りが原因であったとされている。上記のように、AI の判断ミスが学習データ分布に起因する可能性があることが分かる。

このような学習データ分布が要因となる判断ミスは、次のように推定する。まず、学習データの画像の特徴量の分布を求め、判断ミスの画像の特徴量が学習データ分布のどの位置となるか可視化する。学習データの分布が少ない位置に判断ミスの画像が該当すれば、このような学習データの分布が要因であると推定できる。

人間を検知対象とした場合を例として、可視化する特徴量を図5に示す。物体検知における学習データは、画像中の検知対象の検知枠（Bounding Box）情報が事前に付与されたものである。この情報も活用し、

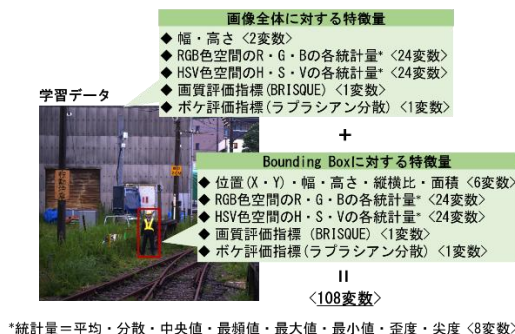


図5 学習データから可視化される特徴量

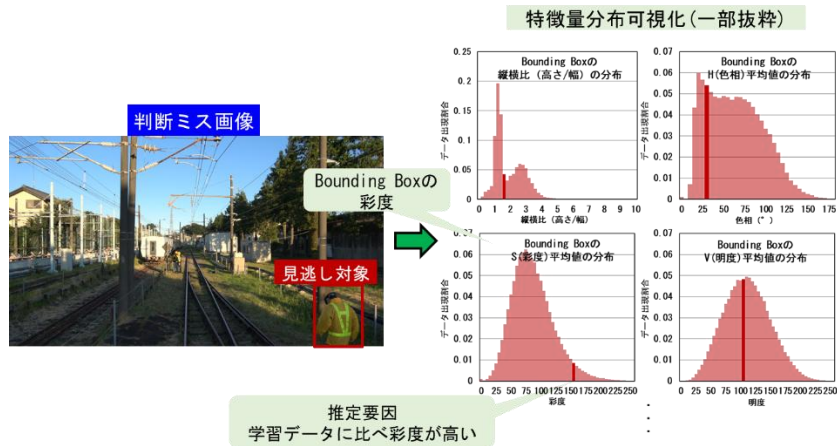


図 6 特徴量の分布を用いた学習データの検査による判断ミス要因推定

画像全体に対する大域的な特徴量と Bounding Box 領域に対する局所の特徴量を組み合わせた特徴量 108 個を可視化する変数 (図 5 中に記載) としている。分布を可視化するには判断ミスが起きた画像に対しても、見逃した Bounding Box の情報を付与し、同様に特徴量を求め、学習データの特徴量の分布のどの位置に該当するかを確認することでデータの偏りによる要因かを推定する。

特徴量の分布を用いた学習データの検査による判断ミス要因推定のイメージを図 6 に示す。図 6 の左側は見逃した画像を示し、右側は各特徴量 (Bounding Box の縦横比、色相、彩度、明度) を横軸とし、縦軸にデータの出現割合とした分布のグラフを示す。

グラフの濃い色の縦棒が当該画像の分布の中での「位置」を示している。図 6 の Bounding Box の彩度 (吹き出しが示すグラフ) のグラフでは、黄色の作業服を着た人物は、分布の頻度が少ない「位置」に該当することが確認でき、Bounding Box の彩度が判断ミスの要因として推定できる。この場合は、彩度の高い人物のデータを学習データに追加する等の対策も考えられる。

特徴量の分布を用いた学習データの検査では、比較的単純な特徴量に関する分布を確認することで判断ミスの要因の推定を行った。しかしながら、AI のニューラルネットワーク内では非線形変換によって、より複雑な特徴量に基づいて高度な判断を行っている。したがって、前述のような単純な特徴量のみではなく、複雑な特徴量に関する検査についてもできることが望ましい。一方で、このような複雑な特徴量は一般には人間がそのまま理解することは難しく、仮にある複雑な特徴量に判断ミスが起因することが推定できたとしても、それを解釈し、対策することは困難であることが想定される。これは AI が透明性を有さないとされる原因でもある。

そこで、複雑な特徴量も考慮するため、AI によって出力される複数の複雑な特徴量の類似度を算出することによって、学習データ上にどの程度似た画像が含まれているかを確認することを考える。これにより、複雑な特徴量を解釈することなく、判断ミスが起きた画像に類似した画像を類似度順に抽出することができ、判断ミス画像が学習データから見た際に珍しい状況かどうかを確認可能とする。

類似度を算出するための AI として、画像検索にも活用されている言語画像 AI モデル CLIP⁹⁾を使用する。CLIP とは、2021 年 2 月に OpenAI によって公開された言語と画像を紐づけて異種の情報を統合的に学習 (マルチモーダル学習) した AI モデルである。CLIP は、図 7 のように画像とテキストのペアデータにより学習が行われている特徴があり、画像とテキストをそれぞれ特徴ベクトルに変換する 2 つの特徴量抽出器であるエンコーダを持つ。学習時には、ミニバッチと呼ばれる n 個のペアデータに対して、エンコーダからそれぞれ特徴ベクトル群を求め、ベクトル群同士の演算により、各成分にコサイン類似度を持

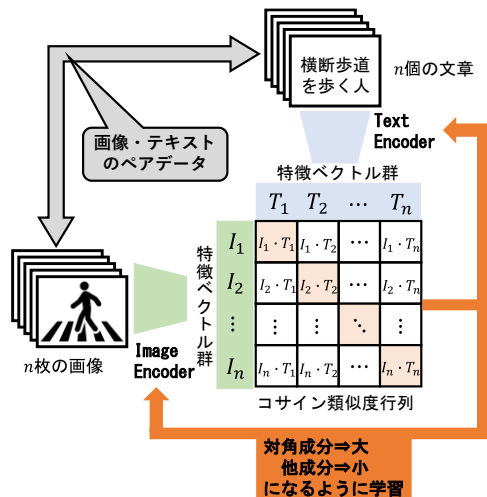


図7 言語画像 AI モデル CLIP の概要



図8 言語画像 AI モデル CLIP による類似度検索

つ行列を計算する。このコサイン類似度行列の対角成分が大きく、その他成分が小さくなるように画像・テキストそれぞれのエンコーダを学習する。これにより、画像とテキストの特徴ベクトル表現が合致するようなエンコーダの学習が行われる。CLIP により学習された画像のエンコーダは、テキスト表現も同時に学習しているため、画像に対してより人間の理解に近い特徴量を獲得していると解釈することもできる。

次に、見逃し等の判断ミスが起きた画像に対して、CLIP を活用して類似度検索した結果のイメージを図8に示す。図8より判断ミスが起きたときと類似した状況の画像を抽出できている様子が確認できる。撮影箇所の前後や画角、雲の様子等が多少変化しても類似した画像として抽出できている。これは、CLIP がテキスト表現と同時に学習することで比較的人間に近い特徴量を獲得しているため、人間が目視で見たときの類似性の判断と大きな矛盾が無いと考えられる。このように、見逃した状況と類似する画像がどの程度学習データ中に含まれているかを解析することで、学習データの中で希少な状況であった場合には、判断ミスの要因として推定する。この検査によりミスの原因が推定された場合は、AI の学習に用いた学習データの中に検知対象となる画像の条件が無い（もしくは少ない）ことが原因であるため、検知対象となる画像を含めた学習データで AI を学習しなおすことが対策として推奨される。

4. 手法の評価

4.1 評価用画像の作成

提案した判断ミスの要因推定手法を評価するにあたり、実際に前方監視 AI が判断ミスにより見逃した画像が必要となる。この際、必要な画像は以下の条件を両方満たす必要がある。

- ・支障物が存在する危険状況に関する画像である。
- ・支障物の存在を前方監視 AI が見逃した画像である。

しかし実際には、どちらの条件も単体でも稀なケースであり、加えて複合条件となっているため、評価用のデータ収集の難易度が非常に高い。

特に、支障物が存在する危険な画像については、実際に状況を再現して撮影する場合、試験線や条件の整った特定の線区等に場所が限られてしまう。一方、判断ミスの要因は天候や日光、線区条件等多様な可能性が考えられ、評価する画像は、様々な条件を含むように実際の営業線の映像等、通常起こりうる



(a)生成した危険画像(晴天条件)



(b)生成した危険画像(降雪条件)

図9 営業線の前方面像から作成した危険画像の例

条件が網羅されていることが望ましい。

そこで、近年急速な発展を見せる生成 AI によって実際の営業線の前方面像に支障物を疑似的に生成し、危険な状況を模擬した画像を作成した。画像生成 AI には、GLIGEN¹⁰⁾を使用した。GLIGEN は、言語から画像を生成するタスク (text-to-image) において、Bounding Box による画像を生成する位置の制御を可能としている。生成する位置は、レールの幅 (狭軌の場合 1067mm) を基準にして、適切な遠近感を保つように人物等を設定する。これにより、前方面像さえ取得すれば、危険な状況を模擬できる。図 9 の例のように、営業線のデータさえあれば、晴天条件や降雪条件等の様々な条件で危険画像を生成可能である。この画像群に対して前方監視 AI が支障物等を見逃した画像を抽出し、判断ミスの要因推定手法の性能評価用の画像とする。

4.2 評価用画像による手法の評価

前節で説明した評価用の危険画像の生成手法を用いて、実際に走行している列車の前頭から撮影した前方監視映像約 215 時間 (データサイズ 7.6TB) に対して危険画像を生成した。動画は、2007 年から 2023 年までに撮影された映像であり、日本全国の様々な線区、季節、天候、時刻を含み、撮影機材も様々であることから、解像度も HD (1,440×1,080) から 5.3K (5,312×2,988) の範囲を含む。映像中で、危険な状況を生成するフレームは 10 秒に 1 回とし、生成する対象は人物として、評価用の危険画像を約 75,000 枚作成した。

次に、これまで鉄道総研で開発した前方監視システム¹¹⁾に使用している前方監視 AI を用いて、この評価用の危険画像に対して検知を行ったところ、約 75,000 枚のうち、2,668 枚において見逃しの発生が確認できた。

ここで、見逃しの発生した 2,668 枚の画像に対して判断ミスの要因推定手法を適用し、手法の性能を評価した。今回生成した評価データにおいては、2,668 枚のうち原因の推定ができない画像は存在しなかった。評価の結果、要因を推定できた見逃しの例を図 10～図 13 に示す。ここで、情報保護の観点から、本稿ではモザイク加工を施しているが評価では未加工の画像を使用している。

入力画像の検査の例を図 10 に示す。この例では、明るさ補正の変更によって検知可能なことが分かり、これが要因と推定できた。明るさは、光源の設置やシャッター スピードの変更、ゲインの変更が対策案となる。

AI 構造の検査の例を図 11 に示す。この例では、ネットワーク構造の変更によって検知可能なことが分かり、これが要因と推定できた。前述の通り、CNN は細かい情報を中心に認識し、Transformer はシルエットのような大域的な情報を中心に認識していると考えられている。今回の対象は、顔等の細かい人間らしい部分の露出が少ないため、CNN が苦手とする認識対象であった可能性がある。対策としては、ネットワーク構造の追加や変更が考えられる。

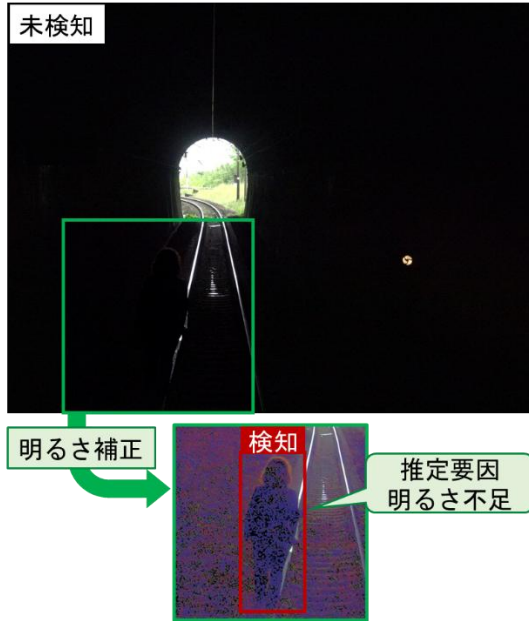


図 10 入力画像の検査手法による推定例



図 11 AI 構造の検査手法による推定例

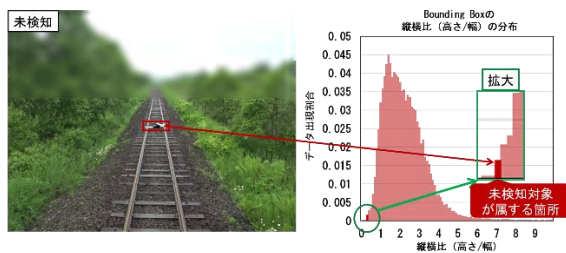


図 12 学習データの検査手法による推定例 (特徴量の分布)



図 13 学習データの検査手法による推定例 (類似度検索結果)

学習データの検査の例を図 12 に示す。この例では、倒れている人物が線路上に存在しているが、Bounding Box の縦横比 (高さ/幅) の分布を確認すると、このように幅が大きい Bounding Box は、学習データに含まれる割合が少ないことが分かった。使用している学習データには、物体検知の学習でよく用いられる一般に公開されたデータセットの人物が多く含まれている。それらは基本的に立位で撮影されており、縦横比が 1 を超える場合 (縦長の Bounding Box) が支配的である。したがって、倒れている人物 (横長の Bounding Box) のデータが少ないことが判断ミスの要因として推定できる。対策としては、倒れている人物に関する学習データの追加等が考えられる。

学習データの検査手法のうち、類似度検索について評価を行った結果を図 13 に示す。見逃した画像のうち図 13 のような列車前頭のガラス面に水滴が付着したケースの類似度検索を行ったところ、水滴が付着しているような同様の学習データは数枚程度しか存在しないことが確認できた。したがって、このような画像が少ないことが判断ミスの要因として推定できる。対策としては、学習データに対して、同じ状況の画像の追加等が考えられる。

5. まとめ

本稿では、列車前方の支障物を検知するための前方監視システムを対象とし、前方監視 AI による見逃

し等の判断ミスの要因の推定手法について述べた。

前方監視 AI の判断ミスを推定するための手順を提案し、入力画像の検査、AI の構造の検査、学習データの検査の 3 段階の検査で要因の推定が可能である。

手法の性能評価では、発生するのが稀な危険な状況を再現するため、実際の列車前方画像に対して生成 AI で加工した危険な状況の画像を約 75,000 枚作成した。その中で前方監視 AI による見逃しが発生した画像 2,668 枚に対して手法を適用し、評価を行った。評価データにおいて、判断ミスの要因推定ができないケースは無く、要因の推定が可能であることが確認できた。

これにより将来的に AI を用いた前方監視システムを鉄道事業者が導入する際に、前方監視システムの支障物の検知性能の検査を導入前に実施できる。また、導入した後においても、前方監視システムが支障物を見逃した際の原因究明および事後対策の検討に活用することが期待される。

今後、前方監視だけでなく設備のメンテナンスを対象とする AI への適用検討など、AI の説明性・透明性に関する研究に取り組み、鉄道分野における AI 活用の促進に貢献していきたいと考えている。

文 献

- 1) 総務省 AI ネットワーク社会推進会議「AI 利活用ガイドライン」：https://www.soumu.go.jp/main_content/000637097.pdf (参照日：2024/08/21)
- 2) Park, Namuk, and Songkuk Kim : "How do vision transformers work?", arXiv preprint arXiv:2202.06709, Vol.2202, No.06709, pp.1-10 (2022).
- 3) Ge, Zheng, et al. : "Yolox: Exceeding yolo series in 2021", arXiv preprint arXiv:2107.08430, Vol.2107, No.08430, pp.1-12 (2021).
- 4) Ultralytics YOLO : <https://github.com/ultralytics/ultralytics> (参照日：2024/02/27)
- 5) Carion, Nicolas, et al. : "End-to-end object detection with transformers", European conference on computer vision, Cham: Springer International Publishing, Vol.12345, No.67890, pp.1-15 (2020).
- 6) Zong, Zhuofan, Guanglu Song, and Yu Liu : "Detsr with collaborative hybrid assignments training", Proceedings of the IEEE/CVF international conference on computer vision, Vol.2023, No.123, pp.100-110 (2023).
- 7) ロイター『焦点：アマゾンが AI 採用打ち切り、「女性差別」の欠陥露呈で』：<https://jp.reuters.com/article/amazon-jobs-ai-analysis-idJPKCN1ML0DN> (参照日：2024/08/21)
- 8) Medium「How Machines Inherit Their Creators' Biases」：<https://medium.com/coinmonks/ai-doesnt-have-to-be-conscious-to-be-harmful-385d143bd311> (参照日：2024/08/21)
- 9) Radford, Alec, et al. : "Learning transferable visual models from natural language supervision", International conference on machine learning, PMLR, Vol.2021, No.1234, pp.123-134 (2021).
- 10) Li, Yuheng, et al. "Gligen: "Open-set grounded text-to-image generation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2023).
- 11) 影山 椋, 長峯 望, 吉野 純樹 : 列車前方監視のためのカメラとセンサを用いた支障物検知手法, 鉄道総研報告, Vol.38, No.3, pp.1-8, 2024