

テレマティクスデータを用いた高リスク車両挙動の抽出 Developing the Method to Extract “High-Risk Vehicle Behavior” from Telematics Data

飯田 克弘¹, 阪本 浩章², 山本 敬太³

Katsuhiro IIDA¹, Hiroaki SAKAMOTO², and Keita YAMAMOTO³

効果的な事故対策の立案には、事故実態のみならず事故に繋がる危険な状況の把握が重要である。これまでビデオ解析による車両挙動把握が行われ、近年では、CCTVや遠赤外線カメラ等を活用した事故の自動検出が試みられているが、事故に繋がる可能性のある車両挙動を任意の時間・空間範囲で簡便に把握する手法は確立されていない。本研究では、テレマティクスデータに対し、オートエンコーダとLSTMの2つの機械学習による異常検知を行うことで、高リスク車両挙動の抽出を試みた。松山自動車道・川内ICを対象に試行した結果、料金所ブース手前での急な進路変更、本線から出口への分流部での急な進路変更を含む走行軌跡のふらつきや路肩停車等が抽出できた。また、オートエンコーダの方が、高リスク車両挙動の抽出に適している可能性を見出した。

To devise effective traffic accident countermeasures, it is crucial to understand not only the actual traffic accident situations but also the dangerous conditions that could lead to traffic accidents. While some studies have analyzed vehicle trajectories using video analysis, and recent attempts have been made to automatically detect traffic accidents using technologies like CCTV or far-infrared cameras, a simple method for readily grasping the vehicle behavior that could cause traffic accidents within arbitrary spatiotemporal ranges has yet to be established. Therefore, in this study, we attempted to extract high-risk vehicle behavior from telematics data by applying two machine learning-based anomaly detection methods: Autoencoder and LSTM. The analysis, conducted at the Kawauchi interchange (IC) on Matsuyama Expressway, successfully extracted vehicle behaviors such as sudden lane changes in front of toll gates and wandering vehicle trajectories, including sudden changes in direction or stopping on the road shoulder at the diverging section from the mainline to an exit ramp. Furthermore, our findings suggest that the Autoencoder may be more suitable for extracting high-risk vehicle trajectories than LSTM.

Keywords: 高リスク車両挙動, テレマティクスデータ, 機械学習, オートエンコーダ, LSTM
High-Risk Vehicle Behavior, Telematics Data, Machine Learning, Autoencoder, LSTM

1. はじめに

交通事故(以下、事故)件数は、2004年の約95万件をピークに減少を続けていたが、2020年以降、約31万件前後でほぼ横ばいに推移し、毎年約2,600人が死亡、約35万人が負傷している¹⁾。事故件数の削減を推し進めるためには、道路交通を構成する、運転者、自動車、道路の各観点での事故対策の立案が必要となるが、本研究はこれらのうち、道路の事故対策(以下、事故対策)に着目する。

藪ら²⁾は事故対策の立案プロセスにおいて、着目する事故形態の発生状況だけでなく、事故に至らないまでも危

険であった状況(以下、危険事象)にも着目することを提唱している。危険事象の把握を試みた研究は、いくつか存在するが、たとえば飯田ら³⁾は、事故が多発する本線料金所において、ビデオ解析を用いて、事故に繋がる可能性のある車両挙動の把握を行っている。

しかし、事故や危険事象の発生場所・時間は予見困難であり、また、ビデオ設置場所の制約や夜間撮影の制約もあるため、ビデオ解析だけでは、任意の時間、任意の空間範囲における車両挙動の把握は困難であり、事故対策立案に必要なデータの全てを抽出できるとは限らない。また、撮影が可能である場合でも、データ取得や解析に

-
- 1 正会員, 博士(工学), 大阪大学大学院工学研究科
Member, Dr. Eng., Graduate School of Engineering, Osaka University
〒565-0871 大阪府吹田市山田丘2-1 e-mail: iida@civil.eng.osaka-u.ac.jp Phone: 06-6879-7611
 - 2 正会員, 修士(工学), 大阪大学大学院工学研究科
Member, ME, Graduate School of Engineering, Osaka University
 - 3 学生会員, 学士(工学), 大阪大学大学院工学研究科
Student Member, BE, Graduate School of Engineering, Osaka University

多くの労力, 時間, 費用を要するという課題がある. ビデオ解析以外の観点として, たとえばNEXCO 東日本では, CCTV の活用⁴⁾, あるいは可視光カメラと遠赤外線カメラの組合せ(多機能ポール)⁵⁾により, 高速道路上の事故や落下物の自動検知に取り組む等, 交通状況を把握する技術開発を進めている. しかし, CCTV は道路の全延長を監視できておらず, また, 多機能ポールも現時点では導入予定箇所が限定されていることから, 任意の時間・空間範囲で事故や危険事象の把握することは, やはり困難と言える.

そこで本研究では, 事故発生有無に関わらず, 任意の時間, 任意の空間範囲で発生した, 事故に繋がる可能性のある車両挙動(以下, 高リスク車両挙動)を選択的に抽出できる可能性のあるテレマティクスデータに着目した. また, 多量のデータを扱うことから, 高リスク車両挙動を目視判別で抽出することは困難であるため, 機械学習を活用することとした.

2. テレマティクスデータを用いた既往研究

テレマティクスデータから車両挙動の把握を試みた既往研究として, たとえば中山ら⁶⁾は, 主に事業用トラックに搭載されたネットワーク型のデジタルタコグラフから得た商用車プローブを用いている. このうち, 遠心加速度や速度に着目することで, 事故多発区間において, 事故に至る可能性の高い挙動の発生過程を推察している. しかし, これらの2指標に限定しての分析となっているほか, 任意の時間・空間範囲の車両挙動を簡便に抽出する手法の提案はなされていない.

川谷ら⁷⁾は, 会津若松市内を走る公用車および公共交通車両に搭載された機器により収集された走行情報履歴データ(位置情報および左右・前後・上下の加速度)から算出される前後方向の急ブレーキについて, その発生条件を機械学習の一つである SVM+FS 手法により判別した. その結果, 特定の車両が, 特定時間帯に特定場所を走行中に, 急ブレーキを発生させやすいことを明らかにした. しかし, 急ブレーキ以外の危険な挙動には着目していないこと, および, 対象車両が6台に限定され,

また, 特定の自治体内のデータに留まっており, 任意の場所で, 多量のデータ収集が出来ない点に課題がある.

水野ら⁸⁾は, ETC2.0 プローブデータを用いて, 阪神高速道路の入口への誤進入および最終的には高速道路に進入しなかった入口手前での迷走が, どの方向から来た車両に多く発生したかを明らかにし, 案内強化や案内充実の必要性に言及している. しかし, ETC2.0 プローブデータは, 約200メートル間隔でのデータサンプリングとなるため, 車両挙動の急変などの把握が困難という課題がある.

以上の課題を考慮し, 本研究では, 多量かつ時間解像度の高いテレマティクスデータから, 任意の時間・空間範囲で, 複数の高リスク車両挙動の実態を同時に, また簡便に把握する手法を検討する.

3. 研究方針

3.1 使用データ

本研究では, テレマティクスデータの一つとして, 損害保険会社が自動車保険の特約サービスに用いているドライブレコーダから取得した走行履歴データ(以下, 走行履歴データ)に着目した. このデータは, 日本全国における特約付帯契約の各車両の各トリップを1つのIDとして, ID毎に緯度や経度, 速度等の走行履歴が一定間隔で時系列順に記録されていることから, 任意の時間・空間範囲における車両挙動を高密度で把握可能であり, 2022年6月から2年分のデータを分析に用いた. 本研究で扱うドライブレコーダによる走行履歴データについては, 実際に取得された走行履歴から損害保険会社が1秒間隔に集計したデータの貸与を受け, 分析に用いることとした. なお, 使用したデータは各トリップのエンジンの始動および停止地点から一定範囲内の情報が除去され, 匿名性が担保されたものとなっている.

3.2 分析対象範囲

本研究では高速道路を選定した. 多くのテレマティクスデータが内包する課題として, 対象車両以外の交



図1 A・Cランプ分岐部

図2 B・Dランプ合流部

図3 下り本線Dランプ分流部

(図1~3のいずれも, 国土地理院地図⁹⁾を改良して作成)

通主体が記録されていない。高速道路であれば、歩行者や自転車等の飛び出し等の要因による急ハンドル、急ブレーキが排除でき、ドライバー自身に起因する高リスク車両挙動を抽出しやすいと考えた。

中でも、インターチェンジ(以下、IC)では合流、分岐、分流の機会が多く、実際に事故や逆走が発生していることから、高リスク車両挙動が単路部に比べて発生しやすいと考えた。そこで、一般的なIC形状であるトランペット型ICであり、逆走の発生件数も多く報告されている松山自動車道・川内ICを分析対象として選定した。なお、川内ICにおいては、2015年から2023年の9年間において、逆走に伴い確保された事案数が4件(故意によるものや、認知症疑いによるものは除く)報告されている。また、同9年間において、川内IC周辺の高速道路では交通事故が51件(NEXCO調べ)報告されており、その発生箇所は、本線、Dランプ、Aランプの順で多く、単独事故や追突事故等、多様な事故が発生している。

川内ICを走行する車両のうち、一般道路から料金所を通過し、上り本線、下り本線に向かうオンランプ(それぞれA、Cランプ)の分岐部(図1)、上り本線および下り本線から料金所に向かうオフランプ(それぞれB、Dランプ)の合流部(図2)および、下り本線からDランプへの分流部(図3)の3箇所を分析対象とした。分析対象範囲は、図1~図3それぞれ赤色枠線で示した範囲とした。

3.3 分析方針

本研究で行う多量の走行履歴データから高リスク車両挙動を抽出する過程を、異常検知と同様と考え、オートエンコーダとLSTM(4章で詳述)を用いることとした。オートエンコーダを選定した理由は、高リスク車両挙動の発生場所や具体的な挙動の全てを予め定義することが困難であるため、教師無し学習が適切と考えたためである。LSTMは、車両挙動は時系列データの集合であり、また、数ステップ前の車両挙動や走行位置が、現時点のそれに影響を与える可能性があるため、時系列順を考慮する必要があると考えて選定したものである。

分析の流れとして、まず、3.2節で述べた分析対象範囲の走行履歴データをオートエンコーダおよびLSTMにより処理し、異常な車両挙動を抽出する。抽出されたデータは、正常ではない車両挙動となるが、たとえば合流のタイミングを計るための一時的な低速走行等、必ずしも高リスク車両挙動とは限らないため、この段階では「高リスク車両挙動の候補」と呼称する。

次に、抽出された高リスク車両挙動の候補の走行軌跡図について、畳み込みニューラルネットワーク(以下、CNN:Convolutional Neural Network)の1種である、VGG-

16により特徴量を抽出する。ここで、CNNは、主に画像認識や画像分類、物体検出などのタスクに特化したディープラーニングモデルの一つで、画像の局所的な特徴を抽出することができる¹⁰⁾。また、VGG-16は、ImageNetと呼ばれる大規模画像データセットで学習された16層からなるCNNモデルで、広範囲のイメージに対する豊富な特徴表現を学習済みのモデルである。

VGG-16で抽出した特徴量は1,000次元に及び、その解釈が困難であるため、特徴量の可視化を目的として、特異値分解(SVD)により特徴量を2次元まで削減する。その後、ward法を用いた階層型クラスタリングを行う。

分析結果の詳細は5章で示すが、いずれの分析結果においても、クラスタ間の境界が明確でない、または複数象限に跨るクラスタが存在するため、本研究ではクラスタ毎の解釈の代替として、2次元の各ベクトルが表す特徴に基づき、高リスク車両挙動を特定することとした。なお、各ベクトルがどのような特徴量を表すかについては、分析者が考察する必要があるため、各走行軌跡図と、それらが持つ2次元の各ベクトルの特徴量を目視観察することにより解釈を行った。

4. 機械学習による分析の概要

4.1 オートエンコーダによる分析の概要

(1) オートエンコーダの概要

オートエンコーダは、図4に示すように、入力データからエンコードにより特徴抽出し、さらにこれを入力層と同じユニット数を持つ出力層にデコードして、入力データの復元処理を行う機械学習である¹¹⁾。この学習のために与えるデータ(以下、学習データ)は、本研究の場合は、正常な車両挙動となる。

異常検知を行う対象データ(以下、テストデータ)にオートエンコーダを適用した結果、出力データが入力データと異なる部分を多く含む場合に、異常であると判定する。この異なる部分を定量的に表現するために、異常度と呼ばれる指標が用いられる。すべてのテストデータの異常度を昇順に並べた際に、正常データと見なせるパーセンタイル値が異常検知の閾値となる。本研究では、高リスク車両挙動は稀有な事象であると考えられることから、全走行軌跡に含まれる正常でない走行軌跡の割合

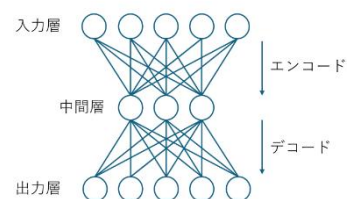


図4 オートエンコーダの概略図

は1%前後であると仮定し, 便宜的に閾値を99%と設定した. なお, 5.1(2)で詳述するが, オートエンコーダにおいて閾値99%で抽出した異常な走行軌跡が93走行分に対し, LSTMで抽出した異常な走行軌跡は134走行分であり, オートエンコーダの方が抽出数は少なかったが, オートエンコーダとLSTMそれぞれに含まれる高リスク車両挙動に差異がなかったことから, オートエンコーダの閾値99%の設定は, 妥当であると判断した.

(2) オートエンコーダによる分析手順

オートエンコーダでは, 入力データとして, 1 車両ごとの走行軌跡図を用いる. 走行軌跡図を作成する際, 解像度 300dpi で描画を行うこととし, プロットは直径 10 ピクセルの正円で表現することとした. なお, 画像サイズは, A・C ランプ分岐部では 240×176, B・D ランプ合流部では 395×406, D ランプ流出部では 759×362 で描画した. また, プロットした点の間隔によって当該車両の速度を表現できるよう, プロットした各点は時系列順に線分で結ばないこととした (たとえば, 5 章の図 6 参照).

次に, 学習データを作成するが, 本研究では, 分析対象データのうち, ID1~500 の 500 走行分から目視で判別した正常な車両挙動を学習データとした. 目視判別においては, 以下の 4 つ条件を満たすものを正常な車両挙動として扱うこととした.

- ①プロットした各点の間隔の変化が少なく, 一定であること (料金所の直近手前は除く)
- ②プロットした各点同士が重なるほどの低速走行や停止がないこと.
- ③滑らかな軌跡であり, 急な進路変更が見られないこと.

表 1 オートエンコーダにおける設定内容 (モデル構造)

部分	層番号	層の種類 (Layer)	フィルタ数	カーネルサイズ	活性化関数	備考
エンコーダ	1	Conv2D	64	3×3	ReLU	padding=same
	2	MaxPooling2D	—	2×2	—	特徴圧縮
	3	Conv2D	32	3×3	ReLU	padding=same
	4	MaxPooling2D	—	2×2	—	特徴圧縮
デコーダ	1	Conv2D	32	3×3	ReLU	padding=same
	2	UpSampling2D	—	2×2	—	特徴復元
	3	Conv2D	64	3×3	ReLU	padding=same
	4	UpSampling2D	—	2×2	—	特徴復元
出力層	—	Conv2D	3 (RGB)	3×3	Sigmoid	画像再構築

表 2 オートエンコーダにおける設定内容 (学習条件)

項目	設定内容
入力画像	128×128ピクセルに変換, RGBの3チャンネル
正規化	0~1にスケール
学習の繰り返し数	20エポック
バッチサイズ	16
データシャッフル	TRUE
検証データ割合	20% (validation split=0.2)
最適化アルゴリズム	Adam
損失関数	平均二乗誤差 (MSE)
閾値	再構築誤差の99パータイル値

④GPS 誤差も含め, 高速道路区域外に軌跡を持たないこと.

なお, 学習データ数を増やしすぎると, 人間の手作業の工数が増え, 機械学習を用いる利点が損なわれるため, 学習データ数は 500 走行分とした. ここで, 学習データを 700 走行分に増やして試行を行った場合であっても, 異常な走行軌跡の抽出結果に変化が認められなかったことから, 学習データ数は 500 走行分で良いと判断した.

テストデータは, ID501以降のデータおよび, ID1~500のうち, 学習データに用いずに除外したデータで構成した. オートエンコーダを実行する際は, 表1および表2の通りの設定とした.

4.2 LSTMによる分析の概要

(1)LSTMの概要

LSTM (Long Short-Term Memory) は, 再帰型ニューラルネットワーク (RNN : Recurrent Neural Network) の 1 種である. RNN とは, 数ステップ前に遡ってデータを反映することのできるニューラルネットワークであり, 時系列データ等の文脈を考慮した学習が可能である¹¹⁾. しかし, 遡るステップ数が増えた場合, 勾配消失問題により学習が上手くいかない場合がある. この問題を解決した手法が, LSTM である. 図 5 のように, 記憶セルに保存している 1 ステップ前の状態を参照する等, 各ゲートの開閉により出力を調整する¹¹⁾.

LSTMでは, データの時系列パターン (各時点での緯度経度や速度等の変遷) が正常, 異常のいずれであるかをラベル付けした学習データを基に, テストデータの時系列パターンの異常検知を行い, 各テストデータが異常である確率を0~1の値で返す. この異常である確率に閾値を設け, 閾値以上のデータを高リスク車両挙動の候補と判定する. 本研究では, LSTMにおける異常判定の閾値は, 次の通りの検討により, 0.999とした. まず, オートエンコーダで用いた閾値0.99をLSTMにおける仮の閾値として設定し, A・Cランプ分岐部を対象に, 異常な走行軌跡の抽出を行った. 次に, 抽出された走行軌跡データを走行軌跡図に変換したところ, 正常な走行軌跡と異常な走行軌跡の両方が含まれていたため, 目視判別により, 異常な走行軌跡のみを取り出した. 目視判別における基準は, オートエンコーダにおける学習データ作成時の4条件

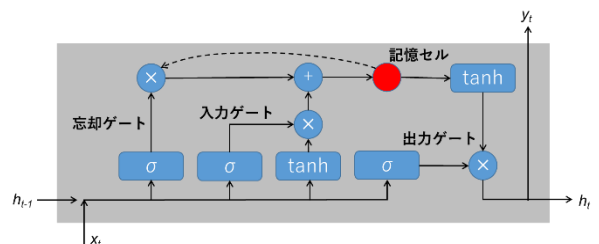


図 5 LSTM の概略図 (文献 11 を参照して作成)

表3 LSTMにおける設定内容

項目	設定内容
標準化方法	StandardScaler
シーケンス長	10
LSTMユニット数	64
LSTM活性化関数	ReLU
Dense出力層ユニット数	1
Dense活性化関数	sigmoid
評価指標	accuracy
学習の繰り返し数	10エポック
バッチサイズ	32
検証データ割合	20%(validation_split=0.2)
最適化アルゴリズム	Adam
損失関数	Binary_crossentropy
異常判定閾値	0.999

と同様である。ここで、目視判別後の異常な走行軌跡それぞれについて、LSTMが算出した「異常である確率」を確認したところ、そのすべてが、0.999から1の範囲内に収まっていることを確認した。したがって、LSTMにおける閾値を0.999と設定すれば、異常な走行軌跡を精度良く抽出できると考えた。

(2)LSTMによる分析手順

入力データは、CSV 上で時系列順に整理された走行履歴データのうち、図1~図3の赤枠で示した範囲内に含まれるデータである。また、分析対象とした走行履歴データの特徴量は、緯度、経度、速度、および前後加速度の4項目である。学習データは、走行履歴データのID1~500の500走行分とした。オートエンコーダの場合と同様の4条件(4.1(2)参照)に基づき目視による判別を行い、正常な車両挙動には0を、異常な走行挙動には1のラベルを付与し、500走行分すべてを学習データとして使用した。

ID501以降のデータをテストデータとし、計算した。この際、表3の通りの設定にて計算を行った。

次に、抽出された高リスク車両挙動の候補に対し、3.3で述べた特徴量抽出を行うために、時系列データを走行軌跡図に変換する。この際、視覚的な速度変化の判り易さを考慮し、プロットした各点を時系列順に線分で結んだうえで、たとえば図9のように速度に応じた配色を施すこととした。

5. 分析結果

5.1 A・Cランプ分岐部の分析結果

(1)オートエンコーダによる分析結果

AランプまたはCランプを通過した全9,734走行のうち、学習データとして444走行分を用いた。学習データに用いた走行軌跡図の例を図6に示す。

9,290走行分をテストデータとして計算した結果、93走行が高リスク車両挙動の候補として抽出された。この93走行に対し、VGG-16で特徴量を抽出し、特異値分解

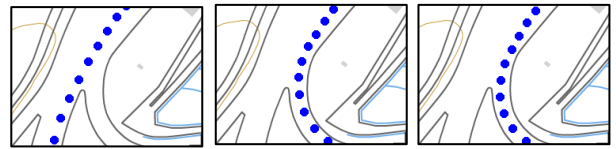


図6 学習データ(正常な走行軌跡)の例(国土地理院地図を加工して作成)

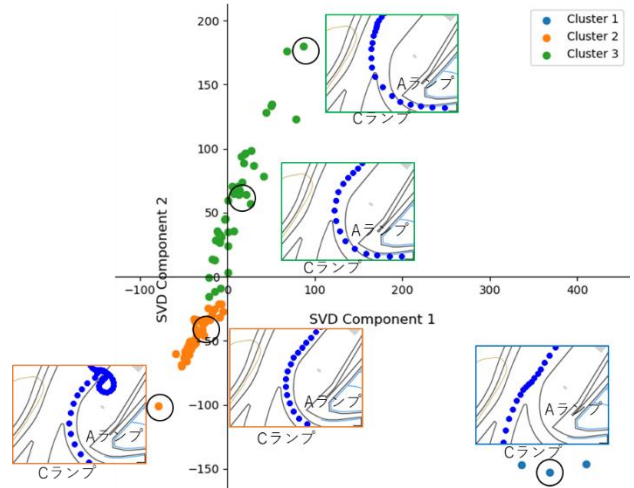


図7 特徴量の散布とクラスタリング結果

を行い、3クラスタに分類した結果を図7に示す。

第1ベクトル(SVD Component 1)の値が小さい(値の符号だけでなく、符号と絶対値の両面から大きさを総合的に判断。以降同様)ほど、Aランプへ分岐し、走行軌跡に半径の小さな円弧を含み、ランプ上では定速走行となる傾向が見られ、逆に値が大きくなると、Cランプへ分岐し、走行軌跡は直線的になり、ランプ上での加減速回数も増加する傾向がみられる。したがって、第1ベクトルは、分岐方向、走行軌跡の形状、およびランプ上での速度変化を示していると考えられる。

次に、第2ベクトル(SVD Component 2)の値が小さければ、道路内に走行軌跡のプロットが収まっており、値が大きいくほど逸脱量が大きくなる傾向が見られる。したがって、第2ベクトルは路外逸脱(GPSのズレ)の程度を表わすと考えられる。

分岐方向やGPSのズレは高リスク車両挙動とは関連が小さいことから、第1ベクトルで解釈した項目のうち、走行軌跡の形状および速度変化について、高リスク車両挙動の候補を目視で確認した。その結果、図7の青枠線で示すような、ランプ分岐部付近で急減速し、進路変更した車両や、橙色枠線のうち左図に示すような低速で円弧を描く挙動を示す車両が確認できた。このうち前者は、周辺車両の状況が不明なため断定はできないが、たとえば分岐部で行先に迷いが生じ、標識やナビを確認するために行われた可能性もあり、その場合には追突事故等のリスクを高める恐れがあることから、高リスク車両挙動の一つである可能性がある。一方後者は、料金所通過後に設置されている内プラザ駐車場に

立ち寄ったものと推定され, 特異な軌跡と言えるものの, 高リスク車両挙動には該当しないと考えられる。

(2)LSTMによる分析結果

500 走行分を学習データとし, 9,234 走行分をテストデータとして計算を行った結果, 134 走行が高リスク車両挙動の候補として抽出された。この 134 走行の走行軌跡図を描画し, (1)と同様に 3 クラスタに分類した結果を図 8 に示す。

第 1 ベクトル(SVD Component 1)の値が小さいほど, A ランプへ分岐し, 軌跡の形状が円弧を有し, 色の変化すなわち速度が変化する傾向が見られる。逆に, 値が大きくなると, C ランプへ分岐または C ランプに軌跡の一部が存在し, 軌跡が直線的あるいは折れ角を持ち, 滑らかな軌跡ではなくなり, 単色(ここでは赤一色の停止状態)となる傾向が見られる。したがって, 第 1 ベクトルは, オートエンコーダと同じく, 分岐方向, 走行軌跡の形状, 速度変化を示していると考えられる。なお, クラスタ 3 (緑色) に示した図は, 全域で速度 0km/h となっているため, 記録上のエラーと考えられる。

次に, 第 2 ベクトル (SVD Component 2) では, 値が小さければ走行軌跡のラインが道路内に収まるが, 値が大きいかほど逸脱量が大きくなる傾向にある。したがって, 第 2 ベクトルは路外逸脱 (GPS のズレ) の程度を表わすと考えられ, オートエンコーダと同様である。オートエンコーダと同様に, 高リスク車両挙動とは関連が小さい項目を除き, 走行軌跡の形状および速度変化

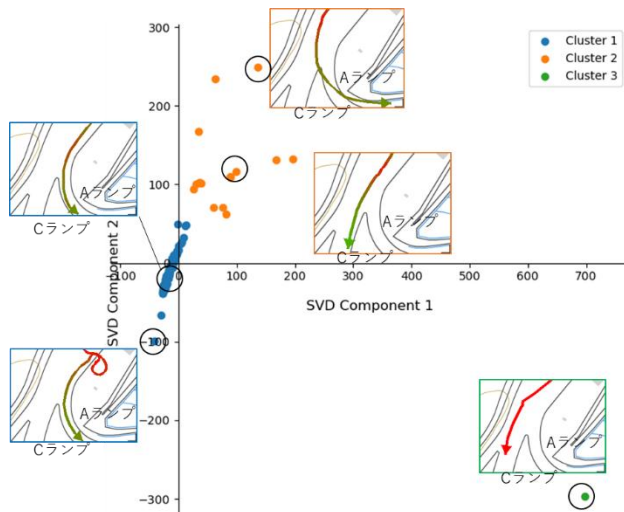


図 8 特徴量の散布とクラスティング結果

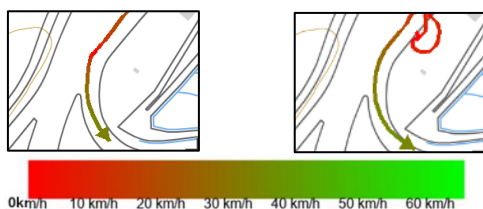


図 9 LSTM で抽出された高リスク車両挙動の候補の例 (国土地理院地図を加工して作成)

について, 高リスク車両挙動の候補を目視で確認した。その結果, (1)と同様に, A・C ランプ分岐部付近で急減速し, 図 9 左図のように進路変更した車両や, 図 9 右図の右上部分において, 低速で円弧を描く挙動を示す車両が含まれていることが確認できた。

以上のことから, A・C ランプ分岐部では, オートエンコーダ, LSTM とともに, 高リスク車両挙動の可能性のある走行軌跡を抽出することができ, 抽出された種類に大きな違いは見られないことが分かった。

ただし, 留意点として, LSTMの方がオートエンコーダよりも抽出された高リスク車両挙動の候補数が約 44%多い点が挙げられる。これは, 4.1 節(1)で述べた通り, オートエンコーダでは, 閾値の設定を便宜的に 99%としており, 高リスク車両挙動の候補はテストデータの 1%分に固定されるためである。したがって, 高リスク車両挙動の種類が多数存在する場合には, オートエンコーダでは, 閾値の設定に留意する必要がある。なお, 今回の結果においては, いずれの手法でも同様の結果を得ていることから, 分析精度に影響はなかったと推察される。

5.2 川内 IC B・D ランプ合流部の分析結果

(1)オートエンコーダによる分析結果

B ランプまたは D ランプを通過した全 7,675 走行に対し, 学習データとして ID1~500 の 500 走行のうち, 492 走行分を用いた。学習データに用いた走行軌跡図の例を図 10 に示す。

7,183 走行分をテストデータとして計算した結果, 70 走行が高リスク車両挙動の候補として抽出された。この 70 走行に対し, VGG-16 の適用および特異値分解を行い, 3 クラスタに分類した結果を図 11 に示す。

第 1 ベクトル (SVD Component 1) の値が小さければ D ランプを走行しており, 値が大きければ B ランプを走行している傾向にある。よって, 第 1 ベクトルは, ランプ走行位置を示していると考えられる。

次に, 第 2 ベクトル (SVD Component 2) の値が小さいほど, 進路の急変のない滑らかな走行軌跡となり, 値が大きくなると, 走行軌跡に折れ角が発生し, 進路の急変が存在する傾向が見られる。よって, 第 2 ベクトルは, 進路の急変の有無を表わすと考えられる。



図 10 学習データ (正常な走行軌跡) の例 (国土地理院地図を加工して作成)

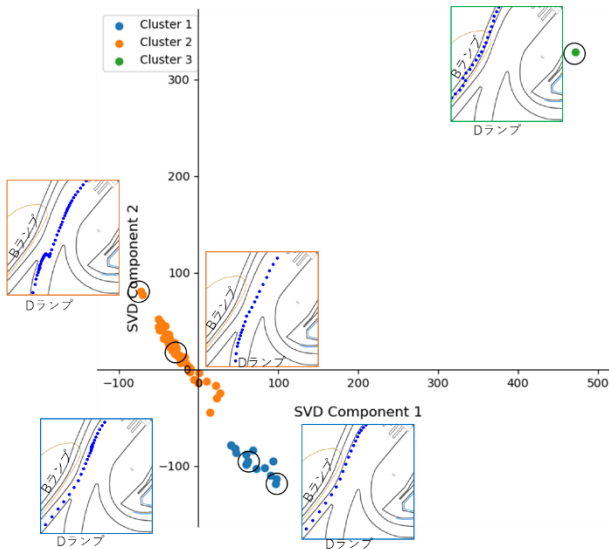


図 11 特徴量の散布とクラスタリング結果

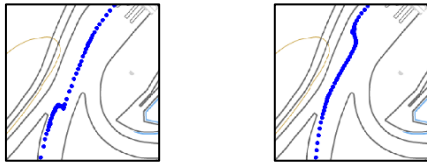


図 12 オートエンコーダで抽出した高リスク車両挙動の候補例 (国土地理院地図を加工して作成)

解釈した項目のうち、走行ランプは高リスク車両挙動との関連が小さいことから、進路の急変について、高リスク車両挙動の候補を目視で確認した。その結果、B・D ランプ合流部や料金所ブース手前で進路が急変した走行軌跡 (それぞれ図 12 左図, 右図) が確認できた。

こちらも断定はできないが、図 12 左図では合流時、B ランプ走行車両に気付くのが遅れ、ドライバーが急ハンドルを行った可能性や、前方に見える料金所のレーン選択の判断を変更した結果として、進路を急変させた可能性が考えられる。また、図 12 右図では、たとえば非 ETC 車が誤って ETC レーンに向かい、料金所直前で一般レーンに向かった場合や、隣接レーンが空いているために移動した可能性が考えられる。図 12 ではいずれの場合も、他車への接触事故あるいは他車による自転車への追突事故が生じるリスクが高まる恐れがあると考えられることから、高リスク車両挙動に該当する可能性がある。

(2)LSTMによる分析結果

500 走行分を学習データとし、7,175 走行分をテストデータとして計算を行った結果、36 走行分が高リスク車両挙動の候補として抽出された。この36 走行分の走行軌跡図を描画したのち、VGG-16 の適用および特異値分解を行い、3 クラスに分類した結果を図 13 に示す。

第 1 ベクトル (SVD Component 1) の値が小さければ B ランプを走行し、特に値が小さい領域では路外逸脱 (GPS のズレ) が顕著になる。一方、値が大きければ D

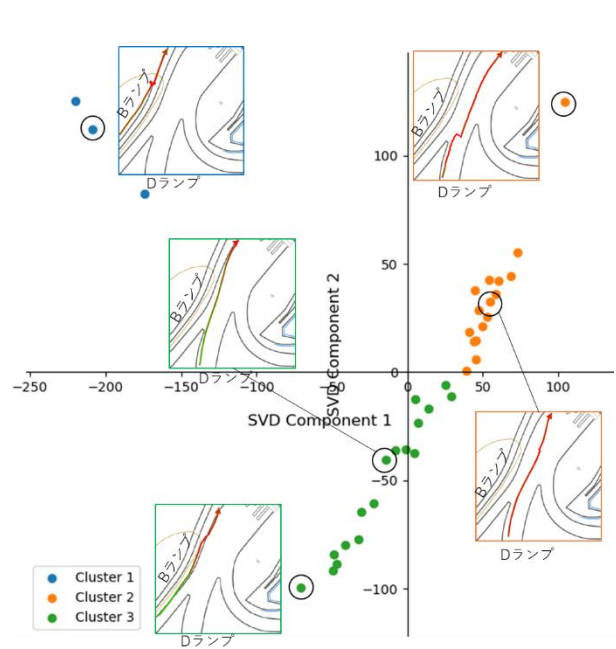


図 13 特徴量の散布とクラスタリング結果

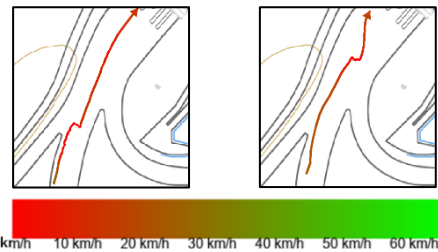


図 14 LSTM で抽出された高リスク車両挙動の候補の例 (国土地理院地図を加工して作成)

ランプを走行し、走行軌跡も道路内に収まる傾向が見られる。したがって、第 1 ベクトルは、路外逸脱も含めた走行位置を示していると考えられる。

第 2 ベクトル (SVD Component 2) の値が小さい場合は、走行軌跡図の速度を表す色が、赤色から緑色まで変化が大きい傾向にあり、値が大きくなれば、速度の色が単色になる傾向が見られる。したがって、第 2 ベクトルは、速度の安定度合いを表していると考えられる。

解釈した項目のうち、高リスク車両挙動と関連の小さい路外逸脱も含めた走行位置を除き、速度の安定度合いについて、高リスク車両挙動の候補を目視で確認した。その結果、図 14 のように、B・D ランプ合流部手前で減速したのち、料金所ブース手前まで低速で走行する挙動が確認できた。この場合も断定はできないものの、たとえば、B・D ランプ合流部付近から、前方に見える料金所において、どのレーンに進入するかを意識が向き、停止に近いほどの低速走行に至った可能性が考えられ、このような極端な低速走行は、後続車からの追突リスクを高めるため、高リスク車両挙動に該当する可能性がある。

以上のことから、B・D ランプ合流部でも、オートエンコーダ、LSTM とともに、高リスク車両挙動の可能性が

ある走行軌跡を抽出することができたが、抽出された種類については、オートエンコーダではB・Dランプ合流部付近や料金所手前での進路の急変、LSTMではB・Dランプ合流部付近からの極端な低速走行という異なる結果となった。

両者で結果が異なったのは、特に第2ベクトルの解釈がオートエンコーダとLSTMで異なる結果となったためであるが、その原因として、学習データの構成の違いが考えられる。具体的には、オートエンコーダでは正常な車両挙動のみを学習データとしているのに対し、LSTMでは正常と異常の両方の車両挙動を学習データとしている。本節のLSTMの学習データに含まれる「異常」と判定したデータが8走行分と少なく、不均衡データに基づく学習が異常検知精度を低下させ、進路の急変を高リスク車両挙動の候補と上手く判断できなかった可能性がある。実際に、抽出された高リスク車両挙動の候補の数は、LSTM(36走行分)はオートエンコーダ(70走行分)の約半分であることから、LSTMの異常検知が上手くできなかった可能性が窺える。

5.3 川内IC 下り本線Dランプ分流部

(1)オートエンコーダによる分析結果

下り本線のうち、Dランプとの分流部より上流側に走行履歴を持つ5,881走行のうち、学習データは496走行分を用いた。学習データに用いた走行軌跡図の例を図15に示す。

テストデータ 5,385 走行に対してオートエンコーダによる分析を行った結果、54走行が高リスク車両挙動の候補として抽出された。この54走行に対し、VGG-16の適用および特異値分解を行い、5クラスタに分類した結果を、図16に示す。

第1ベクトル(SVD Component 1)の値が小さいほど、進路の急変を伴う挙動となり、値が大きければ、円弧を描くような滑らかな挙動となる。したがって、第1ベクトルは、走行軌跡の円滑性を示すと考えられる。なお、クラスタ4および5(それぞれ、赤色、紫色)は、進路の急変を含むものの、一般道路走行車両の軌跡であることに留意が必要である。

次に、第2ベクトル(SVD Component 2)の値が小さければ、道路内に走行軌跡のプロットが収まっており、値が大きければ路外への逸脱量(GPSのズレ)が大きくなる傾向が見られる。よって、第2ベクトルは、路外逸脱の程度を表わすと考えられる。

解釈した項目のうち、一般道路走行車両や、高リスク車両挙動と関連の小さい路外逸脱を除き、走行軌跡の円滑性について、高リスク車両挙動の候補を目視で確認した。その結果、Dランプの減速車線から下り本線へ向きを変えた直後に減速を伴いながら分岐端直前で再びDラ

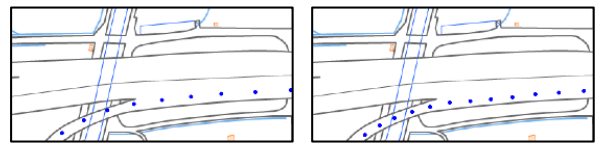


図15 学習データ(正常な走行軌跡図)の例(国土地理院地図を加工して作成)

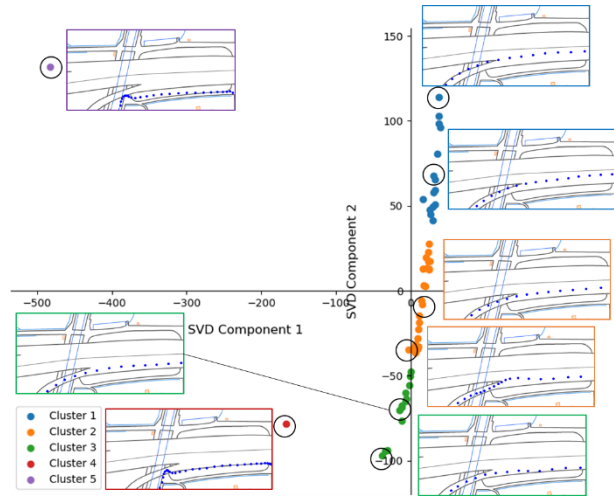


図16 特徴量の散布とクラスタリング結果

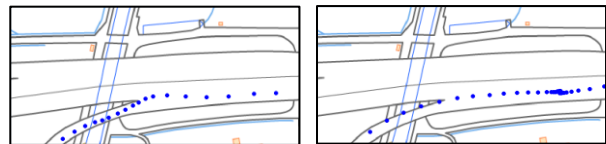


図17 オートエンコーダで抽出した高リスク車両挙動候補の例(国土地理院地図を加工して作成)

ンプに急な進路変更で進入するもの(図17左図)および、下り本線からDランプの減速車線へやや急な進路変更で進入し、路肩に一時停止するもの(図17右図)が確認できた。

図17左図では、走行軌跡に揺らぎが見られることから、流出すべきICかどうか、迷いが生じていた可能性がある。分岐端(ノーズ)直前での進路の急変は、Dランプを走行する後続車両との接触を回避する余裕長がほとんどないことを意味し、事故リスクを高める危険な挙動であると言える。また、分岐端の構造物への衝突リスクも生じる。さらに、仮に路面湿潤時であれば、スリップに伴う事故に至る可能性もある等、様々な形態の事故リスクを孕んだ高リスク車両挙動と言える。図17右図では、ランプ路肩停車は、Dランプに進入してきた後続車両が追突するリスクが考えられることから、これも高リスク車両挙動の一つと言える。

(2)LSTMによる分析結果

500走行分を学習データとし、5,381走行分をテストデータとしてLSTMによる分析を行った結果、4走行のみ、高リスク車両挙動の候補として抽出された(図18)。右図は高速道路と並行する一般道路の走行軌跡であるが、それ以外の3つについては、正常な車両挙動と大

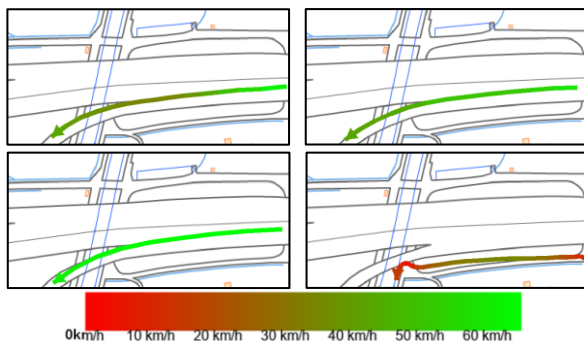


図 18 LSTM で抽出された高リスク車両挙動の候補
(国土地理院地図を加工して作成)

きな違いが観察されない結果となった。

以上のことから、オートエンコーダでは高リスク車両挙動を2種類抽出できたのに対し、LSTMでは抽出できなかった。これは、5.2節(2)と同様に、学習データの作成手法の差異に起因すると考えられる。今回のLSTMでは、学習データの中に含まれる異常な車両挙動は4走行分であったために、不均衡データに基づく学習の影響により、異常検知精度に影響した可能性がある。

6. おわりに

本研究では、近年下げ止まっている事故件数の更なる削減に向けた事故対策立案のためには、事故や事故に繋がる危険な状況の実態把握が重要であるが、任意の時間、任意の空間範囲において簡便かつ迅速にその発生状況を把握する手法がないことに着目し、その手法を検討することを本研究の目的とした。

任意の時間・空間範囲の車両挙動を抽出可能なデータとして、テレマティクスデータに着目し、中でも全国規模かつ高密度でデータ収集される、損害保険会社が自動車保険の特約サービスに用いているドライブレコーダから取得した走行履歴データに着目した。ここから事故に繋がる可能性のある危険な車両挙動(高リスク車両挙動)を抽出する過程を異常検知と見なし、オートエンコーダとLSTMの2つの機械学習を用いることとした。高リスク車両挙動を抽出しやすい環境として松山自動車道・川内ICの3地点を対象に、高リスク車両挙動を抽出した。

まず、料金所から本線へ向かうA・Cランプ分岐部では、分岐部手前で急減速して進路変更を行うという高リスク車両挙動の可能性のある挙動が、オートエンコーダ、LSTMともに確認できた。ただし、オートエンコーダでは閾値の設定に留意する必要があることもわかった。

次に、本線から料金所に向かうB・Dランプ合流部においては、オートエンコーダでは合流部付近や料金所手前での進路の急変、LSTMでは合流部付近からの極端な低速走行という高リスク車両挙動の候補が抽出された。両者の結果が異なった理由は、学習データに含まれる高リスク車両挙動が8件と少なく、不均衡データに基づく

学習となり、LSTMの異常検知精度に影響を及ぼしたためと推測される。

最後に、下り本線からDランプへの分流部においては、オートエンコーダでは、迷いによるものと考えられる進路の急変、および進路の急変を伴うDランプでの路肩停車という2種類の高リスク車両挙動を抽出できた。その一方で、LSTMでは高リスク車両挙動の候補を抽出できなかった。この理由も、不均衡データに基づく学習(高リスク車両挙動は4件のみ)による影響がLSTMの異常検知精度に現れたためと考えられる。

LSTMの異常検知精度を向上させる対策として、「異常」データを人工的に作成する等のオーバーサンプリングや、「正常」データを減らすアンダーサンプリングが考えられる。しかし、これらの手法は、第1章で述べた簡便に高リスク車両挙動を抽出することと相反する。

以上より、不均衡データに基づく学習により高リスク車両挙動を抽出できない可能性がLSTMにはあることを考慮すれば、閾値に関する留意点はあるものの、オートエンコーダの方が高リスク車両挙動の候補を簡便かつ精度よく抽出できる可能性があると言える。

今後の課題として、本研究で得られた知見を川内ICと類似したICに適用した場合も、本研究と同様に、複数種類の高リスク車両挙動を抽出できるかを確認する必要がある。また、オートエンコーダを適用する際に閾値の最適化や、異常検知以外の手法に基づく高リスク車両挙動の抽出等、さらなる手法改良の検討を進めることで、事故対策立案に有用な知見を、より簡便に収集できる環境の構築、ひいては、高速道路での安全性向上に繋げたい。

謝辞

本研究で用いた走行履歴データは、共同研究者・あいおいニッセイ同和損害保険株式会社から匿名加工したうえで提供を受けたものです(共同研究番号:2024共1466)。また、本研究の一部は株式会社高速道路総合技術研究所との共同研究において実施したものです(共同研究番号:2025共1224)。ここに、厚く御礼申し上げます。

参考文献

- 1) 警察庁:統計表, 年報, 【交通事故の発生状況について】, 令和5年(2023), 公表日:令和6年3月7日, URL: <https://www.npa.go.jp/publications/statistics/koutsuu/toukeihyo.html>, 令和6年3月7日公表。(閲覧年月日:2025年2月18日)
- 2) 藪雅行・武本東・尾崎悠太・神谷翔:交通事故の要因分析・対策立案に関する技術資料, 国土技術政策総合研究所資料第787号, 2014年4月。
- 3) 飯田克弘・多田昌裕・安部雅也・山田憲浩:本線料金所手前における車両走行軌跡の把握と交通コンフリ

- クト評価, 第33回交通工学研究発表会論文集, No.20, p.107-112, 2013.9.
- 4) NEXCO 東日本: 高速道路上の事故や落下物などの事象を交通監視カメラ映像から自動で検知する技術開発・実証を開始します, 2020年12月28日公表, URL: https://www.e-nexco.co.jp/pressroom/head_office/2020/1028/00008642.html. (閲覧年月日: 2025年3月5日)
- 5) NEXCO 東日本: moVision プロジェクト, 重点プロジェクト⑧: リアルタイム全線監視(多機能ポール), URL: <https://www.e-nexco.co.jp/assets/pdf/activity/safety/future/06.pdf>. (閲覧年月日: 2025年3月5日)
- 6) 中山達貴・中村俊之・宇野伸宏・Jan-Dirk Schmöcker: 商用車プローブデータを用いた潜在的事故危険性の把握手法の構築, 土木計画学研究・論文集, Vol.37, No.5, pp.I_1093-I_1104, 2017.
- 7) 川谷卓哉・廣川佐千男・峯恒憲・伊東栄典: 機械学習による走行プローブデータからの異常走行箇所発見, 情報処理学会研究報告. ITS, 高度交通システム, Vol.7, pp.1-6, 2019.
- 8) 水野翔太・河本一郎・山口樹・井上徹・西剛広・西岡悟史: ETC2.0 プローブデータを用いた都市高速入口への誤進入車両の挙動分析手法の提案, 交通工学研究発表会, Vol.42, No.16, pp.95-99, 2022.
- 9) 国土地理院: 地理院地図(電子国土 Web), URL: <https://maps.gsi.go.jp/#17/33.795394/132.904154/&base=pale&ls=pale&disp=1&vs=c1g1j0h0k0l0u0t0z0r0s0m0f2&d=m>
- 10) 斎藤康毅: ゼロから作る Deep Learning—Python で学ぶディープラーニングの理論と実装, オライリー・ジャパン, 2016.
- 11) 多田智史: あたらしい人工知能の教科書, 株式会社翔泳社, p.207-226, 2017