

人工知能を用いた効率的な事故防止対策 モデルの実用化に関する研究

成瀬 拓海¹・山脇 正嗣²・寺奥 淳³・倉科 慧大⁴・森本 章倫⁵

¹学生会員 早稲田大学大学院 創造理工学研究科建設工学専攻 (〒169-8555 東京都新宿区大久保 3-4-1)

E-mail: 7ruse.ta93@akane.waseda.jp

²正会員 株式会社建設技術研究所 国土文化研究所インテリジェンスサービスプラットフォーム

(〒103-0014 東京都中央区日本橋蛸殻町 2-14-5) E-mail: ms-yamawaki@ctie.co.jp

³正会員 株式会社建設技術研究所 中部支社 道路・交通部

(〒460-0003 愛知県名古屋市中区錦 1-5-13) E-mail: teraoku@ctie.co.jp

⁴学生会員 早稲田大学大学院 創造理工学研究科建設工学専攻 (〒169-8555 東京都新宿区大久保 3-4-1)

E-mail: bsbl.wsd6@fuji.waseda.jp

⁵正会員 早稲田大学理工学術院 (〒169-8555 東京都新宿区大久保 3-4-1)

E-mail: akinori@waseda.jp

我が国における交通事故の発生状況は年々改善の傾向を示しているが、交通事故による死亡者数については下げ止まりの状態である。そこで、先端技術の積極的な活用や地理的情報などに基づき、交通事故分析の高度化を図り、交通事故抑止に資する交通指導取締りの推進が期待される。そうした時代の中、既往研究では AI の一種である Deep Q-Network(以下、DQN)などを活用し、交通事故の発生を予測するモデルと効果的な取締り活動の実施方法を提案するモデルを構築した。そして本研究では、このモデルを実社会に適応する際の課題や検討すべき項目を整理したうえで、データの使用性の確認や適切なデフォルト値の設定を行い、汎用的な実用化モデルの開発を行った。今後は多様な地域での展開を目指し実証実験に向けて準備を進めていく。

Key Words: *Traffic Safety, ETC2.0, Traffic Accident, Traffic Enforcement, Practical Application*

1. はじめに

(1) 研究の背景・目的

我が国における交通事故の発生状況は、年々減少の傾向にある。交通事故による死者数は、平成で最多となった 1992 年の 11,452 件から減少し続け、2021 年には 2,636 件となっている^{注1)}。しかし、第 11 次交通安全基本計画^{注2)}では、2025 年までに交通事故による 24 時間死者数 2,000 人以下を掲げており、この目標の達成に向けて、ビッグデータや AI などの先端技術の活用により交通事故分析の高度化を図り、交通事故抑止に資する交通指導取締りを推進することを重点施策としている。

こうした背景を踏まえて、既往研究^{注3)}では、近年急速に活用が進む AI を用いて、ETC2.0 プローブデータを含むビッグデータを解析し、交通事故と取締り活動の関係を定量的に示した上で、取締り活動の最適な実施場所の

予測を行った。既往研究で構築した予測モデルが実社会へ適応できれば、現場の経験と併せて、より効率的な交通事故抑止が可能となる。

そこで本研究では、既往研究で構築した予測モデルを実用化することを目的とする。具体的には、実用化をする際の課題点を整理し、その解決策の提案を行う。さらに、多様な地域での展開が可能な実用化モデルを構築することで、効果的な取締り計画の作成の一助となることを目指す。

(2) 既存研究の整理

本研究に関する既存研究を「ビッグデータや AI を活用した研究」、「予測モデルの実用化に関する研究」の 2 つに大別して整理した。

a) ビッグデータや AI を活用した研究

ビッグデータや AI を活用した研究は近年急速に増加

しており、加藤ら²⁾は、ETC2.0 プローブデータを用いて一般道におけるボトルネック箇所とその影響範囲を評価する手法を提案した。従来手法と比較すると、より精緻かつ適正にボトルネック箇所を評価するために ETC2.0 プローブ情報を活用することが有効であることを示した。また、筆者ら³⁾は AI の一種である Q 学習を活用することで、最適な取締り活動の実施場所を予測するモデルを構築した。そして筆者ら³⁾の継続研究として島田ら⁴⁾は、重回帰ベイズモデルを用いて、取締り活動を考慮した交通事故の発生を予測し、交通事故と取締り活動の関係を評価した。さらに、DQN を用いて効果的な取締り活動方法を推計するモデルを構築した。Q 学習に Deep Learning を導入した DQN を用いることで、Q 値を深層ニューラルネットワークにより近似的に求め、複雑な事象の場合でも効率的な学習が可能となった。今後は、実用化に向けてモデルの有用性を検証するために、実証実験を行う必要がある。

b) 予測モデルの実用化に関する研究

学術研究で構築した予測モデルなどを、実務に活用した事例は少ない。そこで、構築したモデルを実用化の際の課題点を整理する。まず赤羽ら⁵⁾は、交通シミュレーションモデルの実用化に向けた課題点について言及した。使用データの現実的な精度やアクセス性が問題となることが多く、実際の交通状況を十分に再現できない場合は実用性があるとは言えないため、使用データの処理方法を検討する必要があると述べた。さらに露木⁶⁾は、数値予報モデルでは、その予測結果がデータの品質や、初期値の精度に大きく依存することを述べ、データ同化の精密化の必要性をまとめた。過去に観測されたデータを数値モデルの初期値の推定に取り込むことで、予測精度の高い初期値の推定が可能になることを示した。

(3) 本研究の位置づけ

既存研究より、交通分野でもビッグデータや AI を活用した研究が多く、その有効性が認められている。また交通安全分野においても、交通事故の防止対策に関する研究やモデルの開発は近年増加している。しかし、それらを実際に活用し、実社会に適応した事例は少ない。そのような中で、既往研究⁷⁾で構築した交通事故予測モデルと取締り活動方法推計モデル(以下、既往モデル)の実社会での運用は、効率的な交通事故抑止施策の実現に寄与すると考えられる。

また、ビッグデータを活用した研究や、構築したモデルを実社会に活用するためには、以下 4 つの課題を解決する必要があることが分かった。

(a) データの収集

収集精度やアクセス性の問題から、必要なデータが入手できない場合に備えて代替データの用意が必要である。

(b) データの処理

入手したデータの解像度が低い場合、モデルへの入力時に再現性を高める適切な処理を行う必要がある。

(c) 安全性の確保

個人情報の取扱いの整理や、システム障害を防ぐセキュリティ対策が必要である。

(d) コスト管理

モデル運用による費用対効果を鑑みて、適切な運用時の計画・管理が必要である。

そこで本研究では、既往モデルを実用化する上で(a)データの収集、(b)データの処理に関する課題を解決し、既往モデルのアルゴリズムを搭載した取締り活動方法最適化システムの構築を図る。学術研究として構築した取締り活動方法推計モデルを、実社会で運用するモデルへと発展させる点に新規性があると考えられる。

2. 本研究の概要

(1) 対象地と使用データの概要

本研究では、図-1 に示す東京都新宿区の JR 新宿駅周辺の 1km 四方を分析対象エリアとする。新宿区では交通事故が一定数発生しており、取締り活動も定期的を実施されている。また、同区内に繁華街と住宅地が共存しており、交通量の多い幹線道路から狭小な生活道路が存在するため、交通事故の発生状況を端的に表現できると考える。分析に際しては、本モデルを実務レベルで活用可能にするため、対象エリアを 50m メッシュに分割し (No.0-No.399)、メッシュ単位かつ日単位で分析する。また、本研究で使用するデータは以下の通りである。

(a) 2015～2018 年の交通事故データ

(b) 2015～2018 年の取締り活動データ

(c) 2015～2018 年の ETC2.0 プローブデータ

(d) 2015～2018 年の気象データ



図-1 対象地と分析対象エリア

(2) 既往研究の概要

既往研究では⁷⁾、重回帰ベイズモデルを用いて、取締り活動を考慮した交通事故予測モデルを構築し、交通事故と取締り活動の関係を定量的に把握することを可能に

した。また、DQN を用いて、効果的な取締り活動方法を推計するモデルを構築し、交通状況に応じた適切な取締り活動場所及び時間を日単位で提案することを可能にした。重回帰ベイズモデルと DQN モデルで使用した具体的な入出力データは表-1 に、研究全体の概要は図-2 に示す。表-1 と図-2 に示すとおり、既往モデルでは取締り活動を提案する当日以前のデータを用いて、モデルを実装している。したがって、交通量や最高速度などを含む ETC2.0 プローブデータのリアルタイムでの取得が可能なことや、取得されるデータに欠損がないことを前提としている。しかし、現状では ETC2.0 プローブデータのリアルタイムでの取得は困難である。そこで、既往モデルの実用化を図るために、適切な使用データのデフォルト値を設定することで、リアルタイムでデータを取得して分析した結果と同程度の精度を確保し、実社会での運用に備える必要がある。加えて、リアルタイムでのデータの取得が可能となった際に、交通状況の再現性が高まる処理方法について予め定めておく必要がある。そこで、3 章では ETC2.0 プローブデータの処理方法、4 章では使用データのデフォルト値の設定方法について言及する。

表-1 既往モデルの入出力データ

	INPUT	OUTPUT
重回帰ベイズモデル	<ul style="list-style-type: none"> 交通事故データ 取締り活動データ 交通量データ 最高速度データ 交差点データ 	事故リスク値
DQN モデル	<ul style="list-style-type: none"> 事故リスク値 取締り活動時間 	取締り実施後事故リスク値

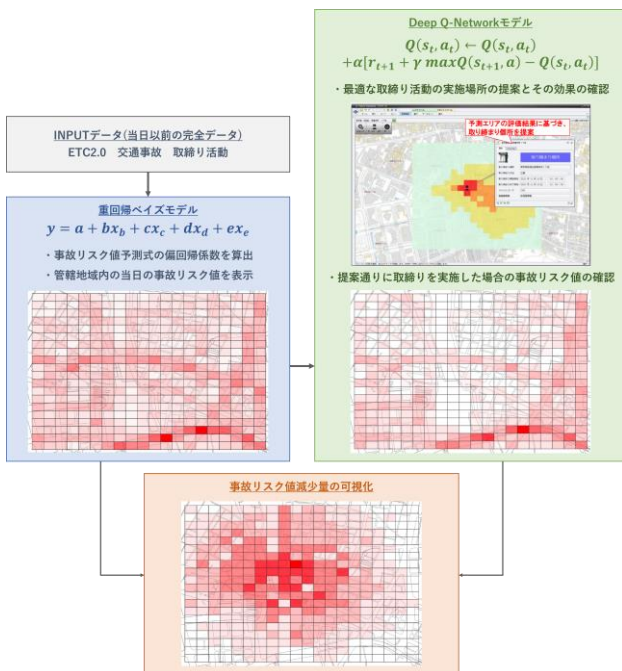


図-2 既往モデルの概要

(3) 本研究の流れ

本研究は以下の流れで進めていく。

(a) 既往モデルの概要把握 (2 章 2 節)

既往研究で使用したデータや構築した重回帰ベイズモデル・DQN モデルについて整理する。

(b) ETC2.0 プローブデータの処理方法の検討 (3 章)

使用するデータを処理する際の課題について解決策を検討する。

(c) データデフォルト値の設定 (4 章)

実用化モデルで使用した交通量データ及び最高速度データのデフォルト値を設定する。

(d) 既往モデルとの精度比較 (5 章)

(a), (b) で挙げた解決策を実施した場合の実用化モデルと、既往モデルの精度を比較する。

3. ETC2.0 プローブデータの処理方法の検討

本研究で使用する ETC2.0 プローブデータは、国土技術政策総合研究所が構築した道路プローブ情報収集システムから出力される走行履歴情報のデータである。データの項目には、表-2 に示すような時刻、緯度・経度、走行速度、道路種別などがあり、前回蓄積した地点から 200m 走行、または進行方位が 45 度以上変化した時点でデータが記録される。測位間隔が 200m であることを踏まえると、50m メッシュ単位での集計時に、車両が通行した全てのメッシュに通過記録が残るわけではないといった課題が生じる。したがって本章では、データ処理時の課題への対応策について言及する。補足として、以下の 2 点に留意する。①プローブデータはドライバーのプライバシー保護のため、起終点から 500m 程度のデータが消去、1 日単位で車両 ID、数時間単位でトリップ番号が変更されるように加工がされていること。②データ容量の都合上、最初に遭遇する ITS スポット以前の 80km 分のデータのみ蓄積される制約があること。

表-2 走行履歴情報のデータ項目 (一部抜粋)

走行履歴の記録情報	車両運行 ID
	GPS 時刻
	トリップ番号
	道路種別
マップマッチング処理後の付加情報	マッチングフラグ
	緯度・経度
	(道路) 管理者コード

はじめに既存の走行履歴情報データに基づきメッシュ毎に交通量を集計すると、図-3 に示すように、交通量の多い幹線道路上であっても、隣接するメッシュの交通量

に必ずしも相関があるとは言えず、正確に交通量の集計ができていないことが分かる。そこで、走行履歴として記録されるプローブデータ内の複数のドットデータを連携して、一つの経路選択データを生成し、より正確な交通量を集計する。

経路選択データの生成にあたっては、まず ETC2.0 プローブデータを対象に、走行履歴情報内の緯度、経度、GPS 受信時刻、車両運行 ID などを用いて、空間・時間条件を適用してトリップ単位での分割、欠損データの補間、そして経路利用実績データを生成する。具体的には、同一車両運行 ID 及び同一トリップ番号のデータを集め、GPS 時刻をもとに時系列順に並べる。その後、各 2 データの緯度・経度から 1 次スプライン補間を行い、2 データ間に等間隔で 100 地点を生成する。以上の補間法により、経路を推定した結果の一例を図-4 に、補間後の隣接するメッシュの交通量の関係を図-5 に示す。図-3 と図-5 を比較すると、補間により隣接するメッシュ間の相関が非常に高くなることが分かる。

次に、最高速度データの補間を行う。交通量データと同様に、各データ間は走行履歴が集計されていないため、通過実績が記録されず、走行速度データが不足している。そこで、走行履歴が蓄積された 2 点間では、走行速度が線形的に増減していると仮定した上で、位置情報の補間と同様に一次スプライン補間を行う。これにより、走行履歴が観測されていないメッシュにおける走行時の速度を推測することが可能である。こうして位置・速度情報を補間した後に、次章以降の分析を行う。

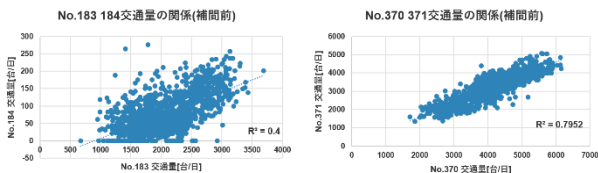


図-3 隣接するメッシュ間の交通量関係(修正前)

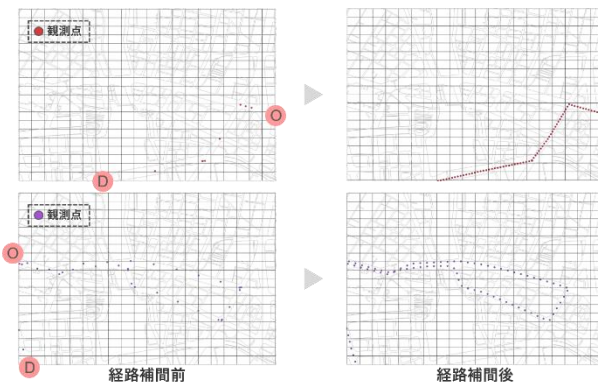


図-4 位置情報修正結果

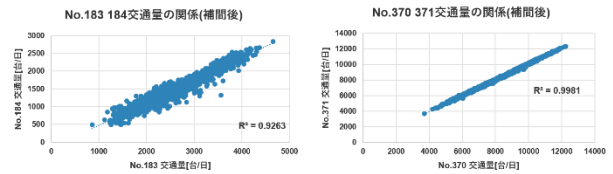


図-5 隣接するメッシュ間の交通量関係(修正後)

4. 使用データのデフォルト値の設定

既往モデルでは、リアルタイムで取得が可能であり、データに欠損がない前提のもとで、ETC2.0 プローブデータから予測日前日までのメッシュ毎の交通量と最高速度データを抽出して、事故リスク値を予測している。そこで、リアルタイムで完全なデータが取得できない場合を想定し、実データの代わりとして使用可能なデフォルト値を設定しておく必要がある。ここで、対象となる 400 メッシュの中には、ETC2.0 車載器搭載車の通行が 4 年間一度も観測されていないメッシュが 130 個存在した。したがって以降の分析では、130 個のメッシュは交通量が 0[台/日]、最高速度が 0[km/h]であると仮定し、残りの 270 個のメッシュのみを考慮して交通量と最高速度を推計する。

(1) 交通量データのデフォルト値の設定

まず、交通量のデフォルト値を推定する。手法には、非線形関数の近似を得意とする Neural Network(以下、NN)を用いる。NN の具体的な構造を表-3 に示す。説明変数には、日ごとの変動を考慮するための休祝日ダミーや、道路環境を考慮した降水量や日照時間、道路構造を考慮した道路占有率や幹線道路ダミーなど、多面的に分析ができるような変数を用いた。分析の有用性を調べるために、分析期間については、2015～2016年の 641 日間(訓練期間)で NN による予測式を構築し、2017年の 365 日間(テスト期間)で予測を行い、実測の交通量データと比較して予測精度を検証する。

NN により予測した結果を図-6 に示す。予測交通量と実測交通量間の自由度調整済み決定係数は 0.931 と非常に高い精度が確認された。また、最小二乗法による近似式を算出した結果は式(1)となった。したがって、NN によるアウトプットを近似式に基づき補正した結果を、交通量のデフォルト値として設定する。

$$V_{act} = 1.10 \times V_{pre} \quad (1)$$

ここで、

V_{act} :交通量の実測値

V_{pre} :NNによる交通量の予測値

表-3 NNの詳細構造

項目	詳細
入力層 (入力データ)	1層 ・メッシュ番号・年/月/日・曜日 ・降水/降雪量・日照時間 ・最大風速・道路占有率
中間層(ノード)	3層(128-64-32)
出力層	1層(メッシュ交通量[台/日])
エポック数	500
バッチサイズ	32
学習率	0.0001
活性化関数	ReLU
最適化関数	Adam

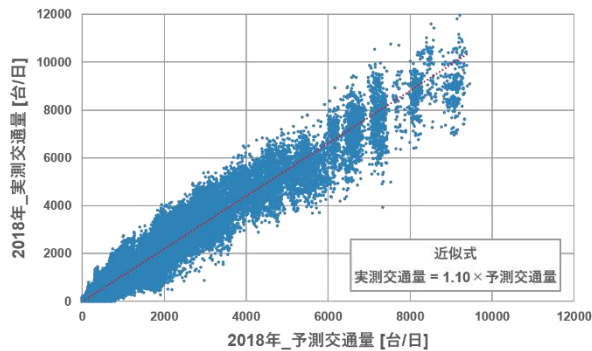


図-6 NNによる2018年の交通量予測の精度検証

(2) 最高速度データのデフォルト値の設定

次に、最高速度のデフォルト値を推定する。まず図-7に、訓練期間における各メッシュの最高速度をヒストグラムで表示する。その結果、最高速度が40~60[km/h]付近と80[km/h]の区間で度数が高いことが分かる。また、分析対象エリア内の道路の法定速度に鑑みると、40[km/h]付近に最頻値を持つ分布、60[km/h]付近に最頻値を持つ分布、80[km/h]に最頻値を持つ分布の3つに分類できる。そこで、メッシュ番号に応じて訓練期間中の最高速度を3つのクラスターに分類し、クラスターごとに最高速度のデフォルト値を設定する。ここでクラスター分析の手法は、データ数の大小を問わず、変数が1次元であること踏まえてk-means法を採用する。具体的には、メッシュ毎に訓練期間中の最高速度データから度数分布表(0~80[km/h]を80階級)を作成する。さらに、階級値とその度数を用いてk-means法を実施し、270個のメッシュを3つのクラスターに分類する。k-means法により分類されたクラスターごとに、最高速度のヒストグラムを作成し、その結果を図-8、図-9、図-10に示す。クラスター分析の結果から、クラスターIに属すメッシュは最頻値44[km/h]を最高速度のデフォルト値として設定する。同様に、クラスターIIでは最頻値60[km/h]、クラスターIIIでは最頻値80[km/h]をデフォルト値とする。

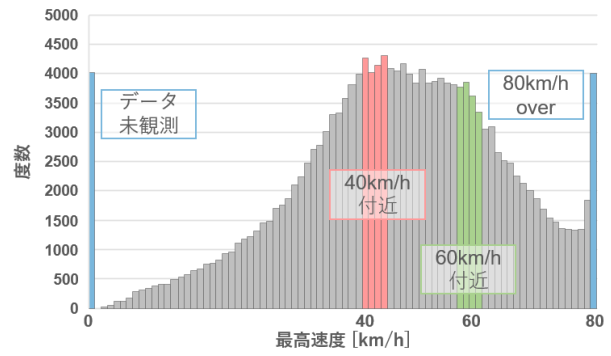


図-7 最高速度のヒストグラム

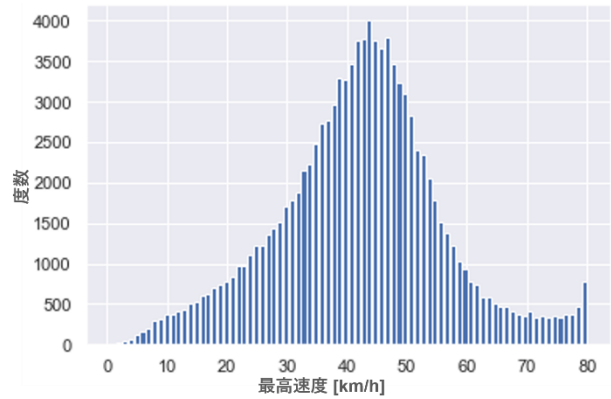


図-8 クラスター分析の結果(クラスターI)

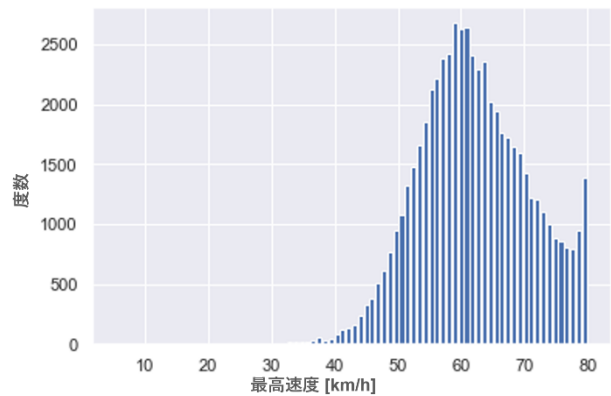


図-9 クラスター分析の結果(クラスターII)

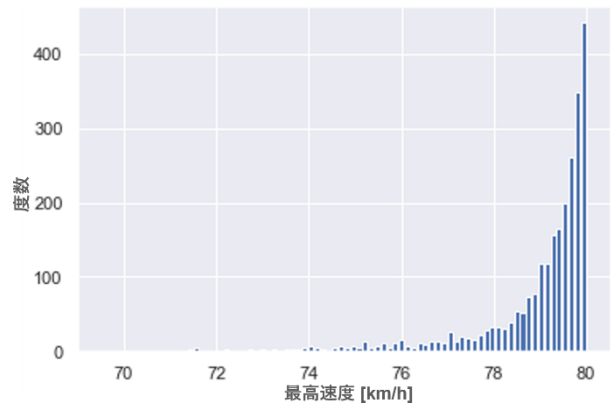


図-10 クラスター分析の結果(クラスターIII)

5. 既往モデルと実用化モデルの精度比較

本章では、3章で実施した位置・速度情報の補間により再現性が向上した交通量と最高速度データと、4章で設定した交通量と最高速度データのデフォルト値を用いて、重回帰ベイズモデルを構築し事故リスク値を予測する。具体的には、直近1年間はデータベースが更新されていない場合を想定し、2015～2016年は実測データ、2017年はデフォルト値を用いて重回帰ベイズモデルを構築し、2018年の事故リスク値を予測する。構築したモデルにおける、取締り活動の偏回帰係数の平均値を表-4に示す。交通安全週間を含めた場合、除外した場合のいずれも偏回帰係数の平均値が負の値となり、取締り活動による事故抑止効果を確認できた。さらに精度比較のために、リアルタイムでデータを取得する場合を想定し、2015～2017年に実測データを用いて予測した既往モデルと、デフォルト値を用いて予測した本モデルにより算出した事故リスク値を、実測の事故リスク値と比較した結果を表-5、表-6に示す。いずれも、既往モデルとの予測精度に大きな差が見られなかった。したがって、既往モデルと同程度の精度が確保されたため、交通量と最高速度のデフォルト値の設定が適切であると考えられる。

表-4 実用化モデルの偏回帰係数

	取締り活動の 偏回帰係数の平均値 μ_i
交通安全週間を含む	-9.92×10^{-4}
交通安全週間を除外	-7.27×10^{-4}

表-5 既往モデルとの精度比較(交通安全週間を含む)

決定係数(本モデル)		決定係数(既往モデル)	
最大値	0.318	最大値	0.315
最小値	0.213	最小値	0.182
平均値	0.260	平均値	0.243

表-6 既往モデルとの精度比較(交通安全週間を除外)

決定係数(本モデル)		決定係数(既往モデル)	
最大値	0.315	最大値	0.309
最小値	0.169	最小値	0.109
平均値	0.233	平均値	0.199

6. おわりに

(1) 得られた知見

本研究では、既往モデルを実用化する際の課題整理と、その対応策を提案した。

まず、データ処理時の課題に対して、交通状況の再現性を高めるために、走行履歴情報から走行ルートの特

定した。一次スプライン補間により、交通量データは隣接メッシュ間の相関が高まり、最高速度データは日ごとの大幅な変動が収まったことで、交通状況の再現性の向上に加え、メッシュ単位での分析を容易にした。

次に、データ収集時の課題に対して、実データの代わりとなるデフォルト値を設定した。交通量データは、NNを用いることで、自然環境や道路構造などの多角的視点から日単位での設定を可能にした。最高速度データは、k-means法によりメッシュごとの特徴を分類し、デフォルト値の設定を可能にした。

さらに、上記の解決策を既往モデルに適用して事故リスク値を予測した。取締り活動による事故抑止効果を定量的に示すとともに、既往モデルと同程度の予測精度を持つことを確認した。したがって、モデルを実用化する上で課題となるデータ収集・処理時の課題を解決したことから、本研究の目的である実用化モデルの構築に成功したといえる。

(2) 今後の展望

本研究では、データの分析や既往モデルの改良を行う過程で、複数の課題が見つかった。ETC2.0プローブデータの処理については、本研究では除外した適切な位置情報を持たないデータの復元方法や取り扱い方を明確に定める必要がある。加えて、最高速度データのデフォルト値の設定にあたっては、日単位での予測を可能とするために、データの収集可能性や有意性を考慮した変数の追加が必要である。そして、本研究で構築した実用化モデルでは、1.(3)で挙げた課題の(c)安全性の確保、(d)コスト管理についても解決策を示すことで、実用化モデルの運用が期待できる。

謝辞：本研究は、警視庁よりデータ提供を受け、国土交通省と連携して実施しました。記して謝意を示します。

NOTES

注1) 警察庁：道路の交通に関する統計，令和4年中の交通事故死者数について，pp.1-2, 2023.

注2) 内閣府：第11次交通安全基本計画，pp.6-13, 2021.

REFERENCES

- 1) 島田大輔，山脇正嗣，寺奥淳，成瀬拓海，森本章倫：Deep Q-Networkを用いた効果的な取締り活動方法の推計に関する研究，土木計画学研究講演集，Vol.65，No.8，2022. [Shimada, D., Yamawaki, M., Teraoku, J., Naruse, T. and Morimoto, A.: A Study on Estimating Effective Methods of Traffic Enforcement Activities Using Deep Q-Network, *Infrastructure Planning Review*, Vol. 65, No. 8, 2022.]
- 2) 加藤哲，橋本浩良，瀬戸下伸介，松田奈緒子：ボトルネックとその影響範囲を特定するためのETC2.0プ

- ローブ情報活用の有効性に関する研究, 土木学会論文誌, Vol. 73, No. 2, pp. 242-250, 2017. [Kato, S., Hashimoto, H., Setoshita, S., and Matsuda, N.: A Study on the Effectiveness of Utilizing ETC2.0 Probe Data to Identify Bottleneck and Its Influence Range, *Infrastructure Planning Review*, Vol. 73, No. 2, pp. 242-250, 2017.]
- 3) 成瀬拓海, 山脇正嗣, 寺奥淳, 森本章倫: Q 学習を活用した最適な交通取締り活動の実施場所の予測に関する研究, 交通工学論文集, Vol. 2, No. 8, pp. 232-239, 2022. [Naruse, T., Yamawaki, M., Teraoku, J. and Morimoto, A.: A Study on Prediction of Optimal Traffic Enforcement Activity Locations Using Q-Learning, *Journal of Traffic Engineers*, Vol. 2, No. 8, pp. 232-239, 2022.]
- 4) 赤羽弘和, 大口敬, 吉井稔雄, 堀口良太: 交通シミュレーションモデルの実用化に向けての課題, 土木研究学研究講演集, No. 20, pp. 521-523, 1997. [Akabane, H., Oguchi, T., Yoshii, R. and Horiguchi, R.: Issues for Practical application of Traffic Simulation Models, *Infrastructure Planning Review*, No. 20, pp. 521-523, 1997.]
- 5) 露木義: データ同化と初期値敏感性, 日本気象学会春季大会シンポジウム「予測可能性—カオスへの挑戦」の報告書, pp. 179-184, 1998. [Tsuyuki, T.: Data Assimilation and Initial Value Sensitivity, *Symposium at the Spring Meeting of the Meteorological Society of Japan*, pp. 179-184, 1998.]

(Received March 1, 2023)

(Accepted March 1, 2023)

Study on Practical Application of an Efficient Accident Prevention Model Using Artificial Intelligence

Takumi NARUSE, Masashi YAMAWAKI, Jun TERAOKU,
Keita KURASHINA and Akinori MORIMOTO

The number of traffic accidents in Japan has been improving year by year, but the number of fatalities due to traffic accidents has remained at a low level. Therefore, it is expected that traffic accident analysis will be advanced through the active use of advanced technology and geographical information, and that traffic guidance and control will be promoted to contribute to the deterrence of traffic accidents. In such an era, previous research has utilized Deep Q-Network (DQN), a type of AI, to construct a model that predicts the occurrence of traffic accidents and proposes effective enforcement methods. In this study, we identified issues and items to be considered when adapting this model to the real world. By confirming the usability of the data and setting appropriate default values, we developed a general-purpose model for practical use. In the future, we will prepare for demonstration experiments to deploy the model in various regions.