

深層学習による 首都高速道路の事故発生予測モデル

森本 裕治¹・吉井 稔雄²・坪田 隆宏³・全 邦釘⁴

¹学生会員 愛媛大学 工学部環境建設工学科 (〒790-8577 愛媛県松山市文京町3番)

E-mail:morimoto.yuji16@cee.ehime-u.ac.jp

²正会員 愛媛大学大学院 生産環境工学専攻 教授 (〒790-8577 愛媛県松山市文京町3番)

E-mail:t.takahiro@cee.ehime-u.ac.jp

³正会員 愛媛大学大学院 生産環境工学専攻 講師 (〒790-8577 愛媛県松山市文京町3番)

E-mail:yoshii@cee.ehime-u.ac.jp

⁴正会員 東京大学大学院 総合研究機構 特任准教授 (〒113-8654 東京都文京区本郷7丁目3番1号)

E-mail:chun@i-con.t.u-tokyo.ac.jp

本稿では、時間的・空間的に広範囲の感知器情報を入力に用いて、予測時点直後の特定道路区間における事故の起こりやすさを予測するConvolutional Neural Networkモデルの構築を試みる。具体的には、首都高速道路の複数路線を対象として、路線上に存在する全感知器による過去1時間の観測情報を入力として、予測時点から2時間先までの間に事故が発生する確率を予測するCNNモデルを構築する。続いて、過去9年間のデータを用いてモデル学習を行い、構築される事故発生予測モデルの検証を行う。検証の結果、多数の織り込み区間が存在する路線に於いて事故リスク予測精度が低下するとともに、一定のサンプルが獲得できれば高精度で事故発生確率を予測することができるとの結果が得られた。

Key Words : 事故発生予測モデル, Convolutional Neural Network, ROC,AUC

1. はじめに

単位距離を走行した際の期待事故発生件数（以下では“事故リスク”）は、道路構造、交通環境、あるいは交通状況によって時々刻々変化すると考えられるが、首都高速道路をはじめとする道路管理者は、道路構造データ、交通環境データ、気象データ等をマッチングさせた交通事故データベースを構築しており、感知器などを用いて獲得される交通観測データと組み合わせることで事故リスクの定量評価が可能となった。

道路構造が事故リスクに与える影響について、Shankar, V¹⁾は単位道路区間におけるカーブ区間の数や最大勾配と事故件数の間に比例関係があることなどを示した。また、交通状況が事故リスクに与える要因について、井上ら²⁾は、交通流状態を渋滞/非渋滞の二つに分類して事故リスクの比較を行い、渋滞時に事故の危険性が高まることを示している。さらに、交通環境による事故リスクの変化について、Andry & Yagar³⁾は非降雨時と降雨時の事故リスク比較を行い、降雨時において事故リスクが増加すること、昼間との比較において夜間や薄

暮時において事故リスクが増大することなどを示している。一方、時々刻々と動的に変化する事故リスク値を用いた交通マネジメント手法として、村上ら⁴⁾はドライバーに対する現時点における情報提供が安全な経路選択を促す事故削減効果を有することを示した。

これらの研究が、過去または現在の時点における事故リスクを対象とした研究であるのに対して、本研究では、直後に事故リスクの上昇が予測される状況に際して、ランプ流入制御やペースカーの導入など、事故リスクの上昇を抑制するための適切な交通マネジメントを実施することを最終目的とするものである。以下、本稿では、特定の道路区間を対象に近未来の事故発生件数の期待値を定量的に把握する事故発生リスク予測モデルを構築し、同モデルの予測精度検証を行う。

2. モデルの評価指標

本研究で構築するモデルは、「事故有り」「事故無し」のいずれかにて予測結果を出力する方法と、事故が発生

する確率を数値で出力する方法を考え、以下の2指標を用いて予測結果の評価検証を行う。

(1) Area Under the Curve 値 (以下では“AUC値”)

事故発生の有無について、直接的に「事故有り」「事故無し」のいずれかにて予測を出力するモデルを構築した際の予測結果評価指標としてAUCを用いる。

表-1に示すように、事故の有無で2値のいずれかを出力するモデルでは、「事故有り」と予測した際に「実際に事故が起きた」True Positive(TP), 「実際には事故が起きなかった」False Positive(FP)のいずれかの状況が出現する。また、「事故無し」と予測した際も同様に「実際に事故が起きた」False Negative(FN), 「実際には事故が起きなかった」True Negative(TN)のいずれかの状況, 計4つの状況が出現する。このとき、モデルの予測精度は式(1)に示す的中率として算出評価することが可能である。もちろん予測的中のTPとTNの割合が高くなれば予測精度が高いということを示す。ここで、本研究のように稀少事象を対象として取り扱う場合には、例えばすべての出力を「事故無し」としておくことで高い中率が実現するとの状況が出現しうる。一方、本研究によるモデルは、計算によって[0,1]の出力値 p ($0 \leq p \leq 1$) を出力するモデルとなっており、予め設定された閾値と p との大小関係から「事故有り」「事故無し」の判定を行うモデルであることから、閾値の値を変化させた場合に、上記4つの状況の出現割合が異なる結果を得ることができる。そこで、閾値の値を[0,1]で変化させた場合のそれぞれで、式(2)、式(3)に示す False Positive Rate (以下では“FPR”) と True Positive Rate (以下では“TPR”) を算出する。前者のFPRは、事故が起きなかった場合に「事故無し」と予測していた割合、後者のTPRは、事故が起きた場合に「事故有り」と予測していた割合を示し、閾値の増加に伴って前者の値は単調に増加し、後者の値は単調に減少する。例えば、閾値を0.0として、全ての予測結果を「事故有り」とした場合には、(FPR, TPR)= (0.0, 1.0) となり、逆に閾値を1.0として全ての予測結果を「事故無し」とした場合には (FPR, TPR)= (1.0, 0.0) の結果が得られる。

続いて、算出結果を図-1に示すFPR-TPR平面にプロットし、閾値ごとに異なる各点を連続的につなぐROC曲

表-1 閾値による予測データの分類表

		予測	
		事故あり	事故なし
実際	事故あり	True Positive(TP)	False Negative(FN)
	事故なし	False Positive(FP)	True Negative(TN)

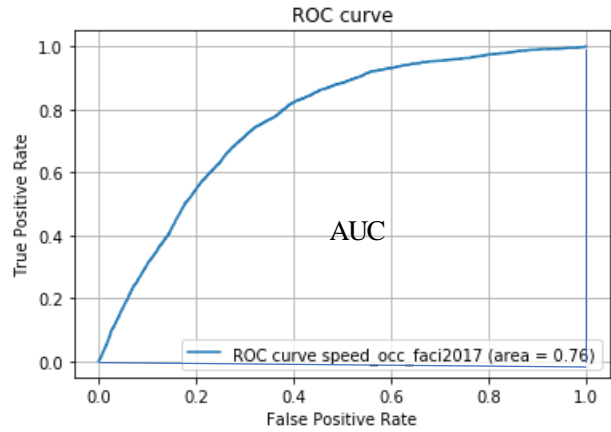


図-1 ROCcurve の例

線 (Receiver Operating Characteristic curve) を用いて予測結果を評価する。図中の座標 (0.0, 1.0) は、「事故あり」と予測したときには全て実際に事故が発生し、「事故無し」と予測した場合には全て事故が発生しない状況、すなわち100%予測的中している状況を示し、各点が図中左上に位置すればするほど良い結果を示すこととなる。そこで、ROC曲線も左上に近づけば近づくほどモデルの予測精度が高いと結論づけられる。このことから、ROC曲線右下部の面積、すなわちAUC値を以て予測精度の評価指標とする。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$False\ Positive\ Rate = \frac{FP}{FP + TN} \quad (2)$$

$$True\ Positive\ Rate = \frac{TP}{TP + FN} \quad (3)$$

(2) 事故発生割合

「事故有り」「事故無し」、2値での予測結果出力方法に代えて、事故が発生する確率を数値で出力するモデルを構築した際の予測結果評価手法には、事故発生確率的中度合いを用いる。具体的には、[0.0, 1.0]の範囲で出力される事故発生確率を0.1刻みの階級に分けて予測値出現回数を集計し、予測値が出現した回数の中に占める実際に事故が発生した回数の割合 (以下“事故発生割合”) を算定する。同割合が各階級の階級値に近い値を示せば高い精度で事故発生確率を予測していることから、事故発生割合と階級値との乖離を以て的中度合いの評価指標とする。

3. モデル概要

図-2に示すように、事故発生予測モデルには、Convolutional Neural Network（以下“CNN”）モデルを用い、以下に示す車両センサー情報を入力として、事故発生の有無、ならびに事故発生確率として予測時点から2時間先までの事故リスク値を出力するモデルを構築する。

入力データ(予測時点直前の1時間) Output(予測時点から2時間先まで)



図-2 モデル概略図

(1) 入力データ概要

予測対象路線を表-2に示す。オンランプ、オフランプ、さらには放射状各路線のジャンクションが互いに近接する路線として、首都高速道路環状線内回り、外回り、ならびに放射状路線を代表して、首都高速道路4号新宿線上り、下り、5号池袋線上り、下り、計6路線とする。入力値としたのは、各路線に設置された全車両センサーの平均速度、交通量、オキュパンシーの各5分集計値である。さらに、各路線5分単位での事故発生の有無を示した施設接触を除く事故データ、それぞれ2009年4月1日から2018年3月31日までの9年間分である。このうち、全期間の7月以外のデータを学習用データ、全機関の7月データを検証用データとする。

表-2 予測対象路線

	路線延長[km]	センサー区間数	事故件数[件/9年]
環状線内回り	14.8	34	4829
環状線外回り	14.8	32	4877
新宿線上り	13.5	24	3025
新宿線下り	13.5	20	2409
池袋線上り	21.5	36	5050
池袋線下り	21.5	35	3465

入力データのイメージを図-3に示す。5分ごとに得られるセンサーデータは、平均速度、交通量、オキュパンシーのそれぞれをセンサー数×知器時間帯の行列にして入力する。各行は、予測時点直前1時間分の5分データを時間の順に並べたものであり、各列は、各センサー区間における観測データを上流から下流に向けてセンサーが位置する順に並べたものである。

一方の、事故データについては、予測時点から2時間先までの各5分間での事故発生の有無を、事故有り (1)、事故無し (0) として時間の順に並べた列ベクトルとし

て入力する。

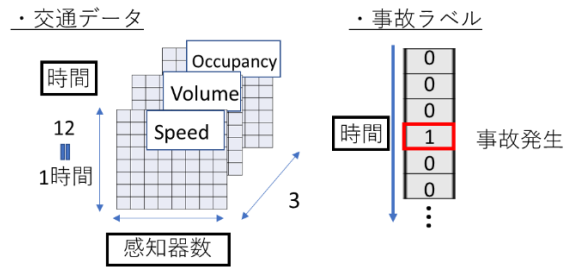


図-3 入力データ

(2) モデル構造

本研究で用いたCNNモデルの構造例を図-4に示す。畳み込み層を3層用意したのちデータを1次元化して全結合層につなげるネットワークとし、事故リスク値を[0.0, 1.0]の値として出力するモデルとした。

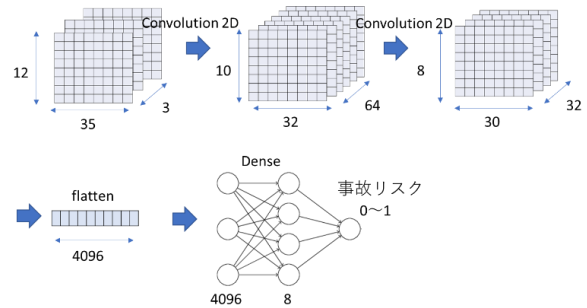


図-4 モデル構造の例

4. 結果

(1) AUC値による検証結果

対象とした6路線におけるAUC値による評価検証結果を図-5-1～図-5-6に示す。AUCのスコアは5号池袋上りりりで0.76と最も良いスコアを記録している。一方で環状線の予測結果外回り内回り共にスコアが0.7を下回る結果が得られた。環状線では、オンランプ、オフランプ、さらには放射状各路線のジャンクションが近接していることから、多数の織り込み区間の存在が事故リスク予測精度の低下を招いたと推察される。

(2) 事故発生割合による結果

事故発生割合による検証結果を図-6に示す。予測精度を示す各階級の階級値と事故発生割合との乖離は、サンプル数の多い予測値の小さい階級においては小さく、サンプル数の少ない予測値の大きい階級において大きくなっている。このことから、一定のサンプルが獲得されれば精度良く事故発生確率を予測することができるという結果が得られた。

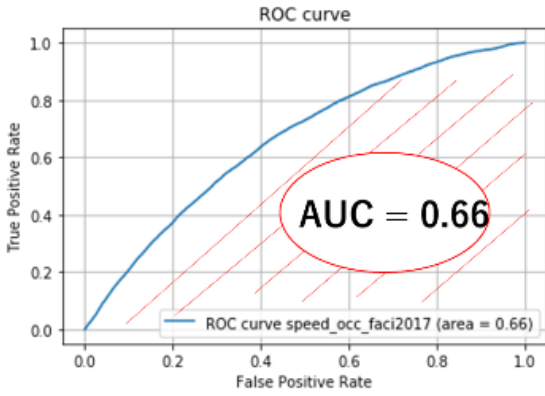


図-5-1 AUC値による検証結果（環状線内回り）

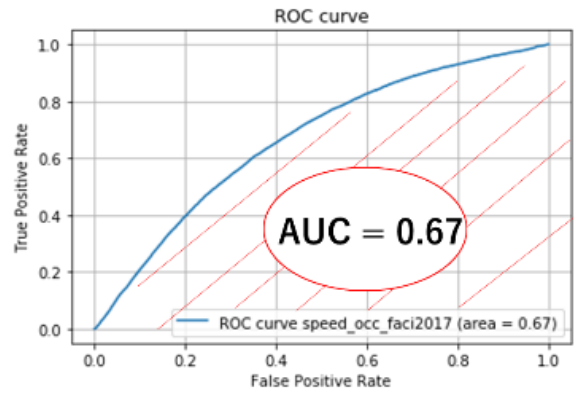


図-5-2 AUC値による検証結果（環状線外回り）

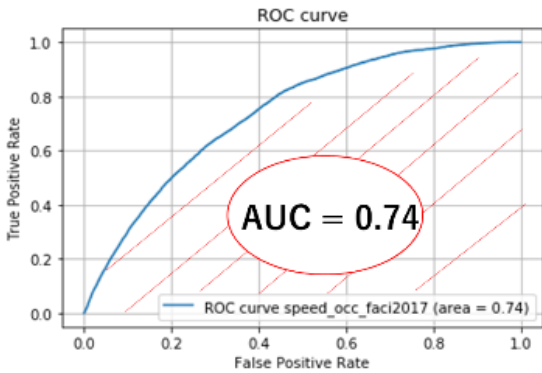


図-5-3 AUC値による検証結果（3号新宿線上り）

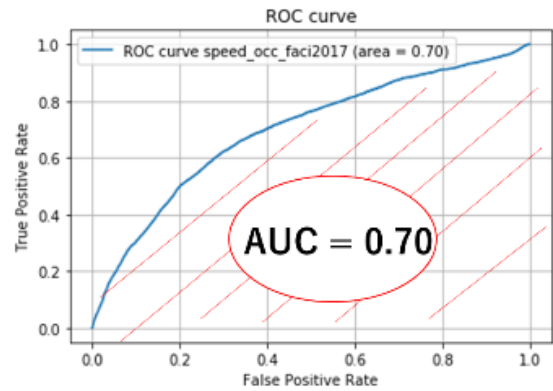


図-5-4 AUC値による検証結果（3号新宿線下り）

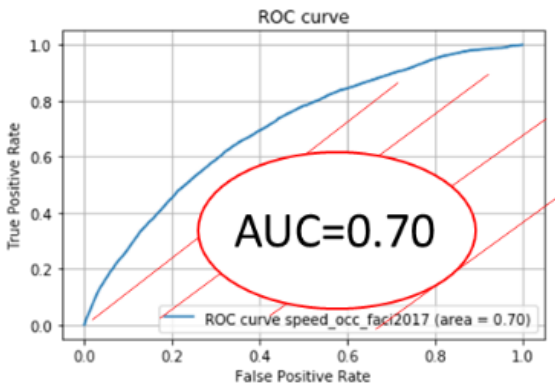


図-5-5 AUC値による検証結果（5号池袋線上り）

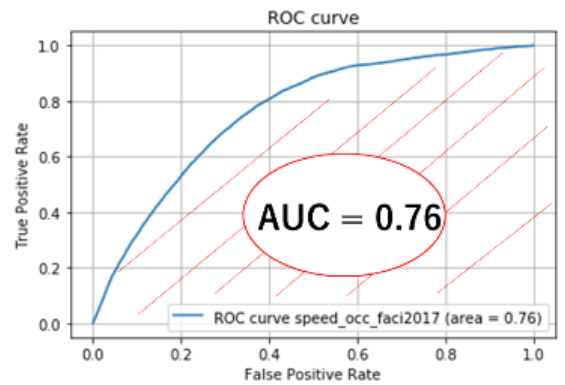


図-5-6 AUC値による検証結果（5号池袋線下り）

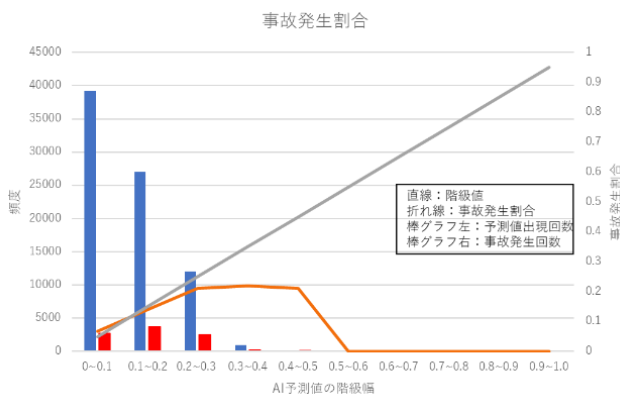


図-6-1 事故発生確率による検証結果（環状線内回り）

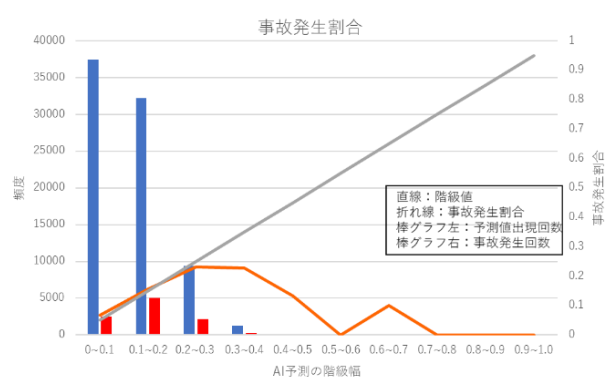


図-6-2 事故発生確率による検証結果（環状線外回り）

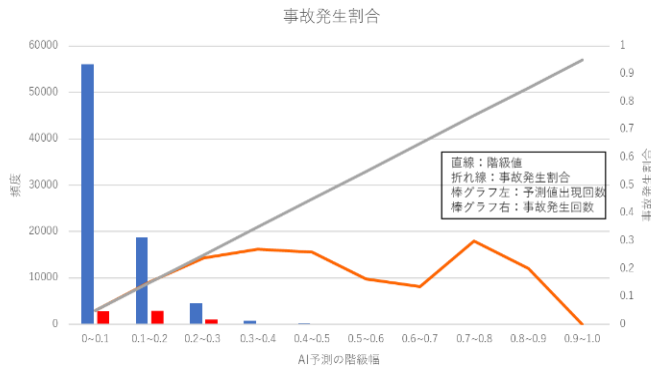


図-6-3 事故発生確率による検証結果 (4号新宿線上り)

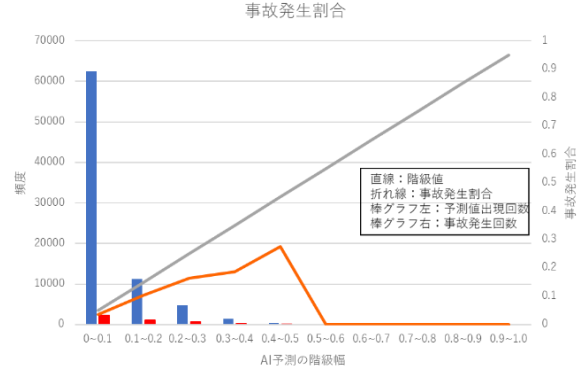


図-6-4 事故発生確率による検証結果 (4号新宿線下り)

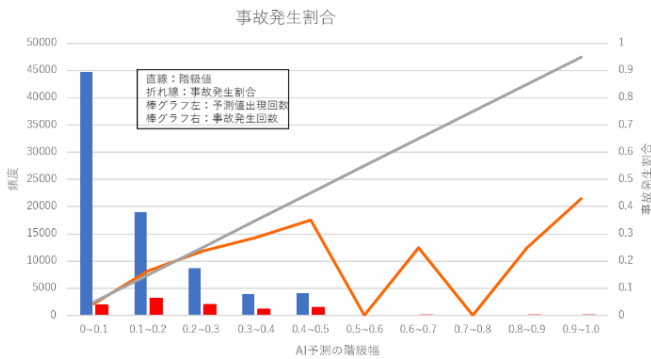


図-6-5 事故発生確率による検証結果 (5号池袋線上り)

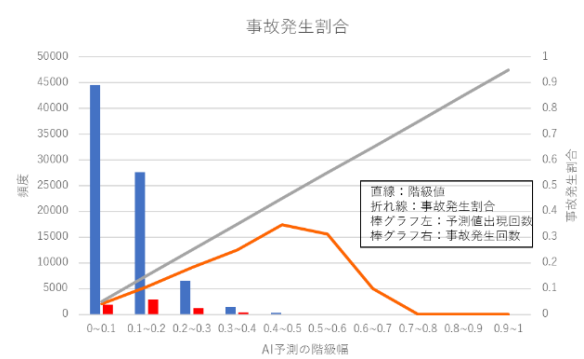


図-6-6 事故発生確率による検証結果 (5号池袋線下り)

5. おわりに

本研究では、予測時点から2時間先までの時間における事故リスクを推定するCNNモデルの構築を行い、モデル予測精度を検証した。その結果、多数の織り込み区間が存在する路線に於いて事故リスク予測精度が低下するとともに、一定のサンプルが獲得できれば高精度で事故発生確率を予測することができるとの結果が得られた。

今後は、不連続な交通流状態量の変化を示すショックウェーブや集計QKを用いたマクロ交通流状態量など、時々刻々と動的に変化する交通現象や交通状態量を入力値に取り込むことで予測精度の向上を図る。また、リカレントニューラルネットワークなどを用いたモデル構造、あるいは深層学習時におけるバッチサイズや最適化手法を見直すことでも予測精度の向上を図りたい。

謝辞:

貴重なデータを提供していただいた首都高速道路 (株)

の方々、モデル構造に関して貴重なご意見を頂いた (株) オリエンタルコンサルタンツならびに (株) エクサウィザーズの方々に感謝の意を表する。

参考文献

- 1) Shankar, V., Milton, J. and Mannering, F.: Modeling Accident Frequencies as Zero-Altered Probability Processes: An Empirical inquiry, Accident Analysis and Prevention, Vol. 29, pp.829-837, 1997.
- 2) 阪神高速道路公団：阪神高速道路の交通管制に関する研究報告書, 交通工学研究会, 1978.
- 3) Andrey, J., Yagar, S.: A temporal analysis of rain-related crash risk. Accident Analysis and Prevention, Vol.25 (4), pp.465-472, 1993.
- 4) 村上和宏, 倉内慎也, 吉井稔雄, 大西邦晃, 川原洋一, 高山雄貴, 兵頭知: 事故リスク情報がドライバーの選択行動に与える影響に関する研究, 第49回土木計画学研究発表会・講演集, CD-ROM, 2014

(?)

Prediction of Dynamic Traffic Accident Risk on Tokyo Metropolitan Expressway by Artificial Intelligence

Yuji Morimoto, Toshio Yoshii, tsubota takahiro, and Chun pang-jo

This study develops an on-line AI model for predicting the likelihood of the occurrence of an accident. The AI model with Deep Neural Networks can predict the likelihood on the next 2 hour. The input data consists of time-series traffic detector data. The model outputs the probability of an accident occurring during the next 2 hour. Then, the AI model is applied to the Tokyo Metropolitan Expressway network. Machine learning task is carried out on the training set of time-series data in 9 years. Then the performance of the training model is evaluated on the test set, which consists of them in a month outside the specific period above. As a result, the accident risks on the next 2 hours have been well predicted in a high accuracy.