

地図画像から交通事故危険度の高い小領域を抽出する CNN の学習 Learning of CNNs to Extract Small Areas with High Traffic Accident Risk from Map Images

大野 友暉* 右田 剛史† 高橋 規一†
Yuki ONO Tsuyoshi MIGITA Norikazu TAKAHASHI

1 はじめに

2023 年に日本国内で発生した交通事故は 307,911 件であり、19 年ぶりに増加に転じた。交通事故による死者も 2,678 人で 2022 年の 2,610 人から増加した [1]。長期的に見れば交通事故件数は減少しているものの、交通事故のない安全な社会を実現するには、さらなる対策が必要である。政府は第 11 次交通安全基本計画 [2] を定め、「世界一安全な道路交通の実現を目指し、24 時間死者数を 2,000 人以下とする」「重症者数を 22,000 人以下にする」の 2 つを目標を掲げている。また、具体的対策の 1 つとして「ビッグデータの活用による潜在的な危険箇所の解消」をあげている。

呉ら [3] は、OpenStreetMap (OSM) [4] のズームレベル 17 のタイル画像に、交通事故発生件数によって「危険」または「安全」のラベルを与え、それらを畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) [5, 6] に学習させ、学習済み CNN を用いて未知のタイル画像の交通事故危険性を判別する方法を提案した。さらに、岡山県内で過去 10 年間に発生した交通事故データを用いて岡山県内のタイル画像にラベル付けを行い、それらを CNN に学習させて性能を評価し、未知のタイル画像の交通事故危険性を約 90% の高い精度で判別できることを示した。一方、Arase ら [7] は、学習済み CNN がタイル画像の危険性を判別する際の着目箇所を、Gradient-weighted Class Activation Mapping (Grad-CAM) [8] を用いてヒートマップの形で可視化し、タイル画像と重ね合わせることによって、より細かな危険箇所の抽出ができる可能性を示した。その一例を図 1 に示す。オレンジ色が CNN の着目箇所を表している。

しかし、上記の先行研究にはいくつかの問題がある。呉ら [3] はズームレベル 17 のタイル画像の学習と評価を行っているが、それらのタイル画像は一辺が約 250m の正方領域に相当するため、危険と判定されたとしても交通事故抑止活動を効果的に行うのは難しい。Arase ら [7] は Grad-CAM を用いて CNN の着目箇所を抽出しているが、CNN の学習に用いるズームレベル 17 のタイル画像は、呉ら [3] と同じくその中で発生した交通事故件数のみでラベル付けされており、交通事故発生地点の情報は考慮されていないため、学習済み CNN がタイル画像の危険性を判別する際に着目する箇所が、交通事故が発生しやすい箇所と一致するかは明確でない。

本研究では、ズームレベル 17 のタイル画像を 8×8 の正方小領域 (以下ではブロックとよぶ) に分割し、交通

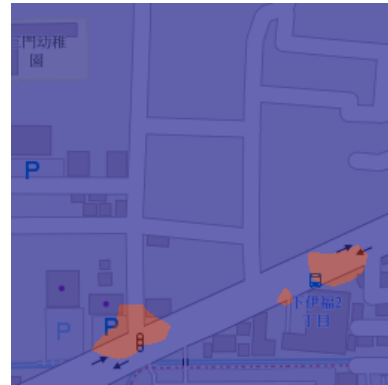


図 1: Grad-CAM による CNN の着目箇所の可視化

事故件数に基づき各ブロックにラベル付けを行った上で CNN に学習させ、学習済みの CNN を用いてブロックごとの交通事故危険性を判別する方法を提案する。

ブロックごとの交通事故件数を用いてラベル付けを行うため、タイル画像全体に 1 つのラベルを付与する先行研究とは異なり、交通事故発生箇所を考慮した学習が行われると期待される。また、Grad-CAM による可視化では CNN の着目箇所が複雑な形状をもつものに対し、提案手法では正方領域であるため性能評価が容易になる。

ラベル付けの方法に加え、学習に用いる画像や CNN モデルについても新たなアイデアを導入する。画像には、OSM 画像だけでなく、衛星画像や、衛星画像の道路部分を OSM 画像で置き換えたものも用いる。CNN モデルには、先行研究 [3, 7] で高性能を示した ResNet18 [6] と、本研究のために新たに設計したものを採用する。これらの効果を検証するため、画像と CNN モデルの組み合わせごとに学習を行い、PR (Precision-Recall) 曲線や PR-AUC (Area Under the Precision-Recall Curve) を用いて判別性能を評価する。

2 データセットの作成

2.1 ラベル付け方法

岡山県全域のズームレベル 17 のタイル画像 88,182 枚の各々を 8×8 のブロックに分割し、岡山県警察から提供された交通事故データを用いてブロックごとにラベル 1 または 0 を付ける。ラベル 1 は危険を表し、ラベル 0 は安全を表す。交通事故データは、2010 年から 2022 年までの 13 年間に岡山県内で発生した 12 万件を超える交通事故の情報からなり、各事故の発生地点の緯度と経度が含まれている。以下ではラベル付けの詳細を述べる。

まず、すべてのタイル画像を 8×8 個のブロックに分割する。各ブロックは一辺が約 30m の正方領域に相当する。交通事故データを用いて 13 年間に事故が 1 回以上発生したブロックを抽出し、それらを対象に 13 年間の事故件数の平均値を求めたところ約 2.7 であった。そこで、各ブロックに対し、13 年間の交通事故件数が 2.7

* 岡山大学大学院環境生命自然科学研究科 Graduate School of Environmental, Life, Natural Science and Technology, Okayama University

† 岡山大学学術研究院環境生命自然科学学域 Faculty of Environmental, Life, Natural Science and Technology, Okayama University



図 2: ラベル付けの結果

より大きければ危険とみなしてラベル 1 を付与し、そうでなければ安全とみなして 0 を付与する。ある OSM タイル画像の 64 個のブロックに、この方法でラベル付けを行った結果を図 2 に示す。赤色を重ねたブロックにはラベル 1 が付与され、それ以外のブロックにはラベル 0 が付与されている。

2.2 学習・評価用画像の作成

CNN の学習および評価に 3 種類の画像を用いる。以下ではそれらの作成方法を述べる。

1 つ目は先行研究 [3, 7] で用いられた OSM 画像である。ただし、OSM 画像をそのまま用いると、CNN が交通事故危険性を判別する際に地図中の文字や地図記号の影響を受けてしまうことがある [7]。そこで本研究では文字も地図記号も含まない OSM 画像を用いる。また、ズームレベル 17 のタイル画像 (256×256 画素) を CNN の学習および評価に用いると、境界にあるブロックを判別する際に周辺ブロックの情報が不足して精度が低下するおそれがあるため、タイル画像の周囲を 1 ブロック分 (32×32 画素) 拡張した 320×320 画素の画像を用いる。

2 つ目は Mapbox [9] から取得した衛星画像である。衛星画像には OSM 画像では失われる建物や樹木の情報が含まれているため、それを利用することによって CNN の判別精度が向上する可能性がある。OSM 画像の場合と同様に、ズームレベル 17 の衛星タイル画像 (256×256 画素) の周囲を 1 ブロック分拡張した 320×320 画素の画像を用いる。

3 つ目は、2 つ目の 320×320 画素の衛星画像の道路部分を OSM 画像の道路で置き換えたものである。衛星画像には道路周辺の建物や樹木の情報が含まれているものの、肝心の道路部分が不鮮明な場合がある。そこで衛星画像の道路部分のみを OSM 画像で置き換える。道路が鮮明になり、かつ道路周辺の情報も利用できるため、衛星画像をそのまま用いる場合に比べて CNN の判別精度が向上する可能性がある。この画像の作成手順を図 3 に示す。

2.3 データ拡張

CNN に画像を学習させる際には、元の画像データに移動、回転、拡大などの人工的な変換を施して拡張することによって、学習済み CNN の精度が向上する可能性がある。本研究でも画像データの拡張を行う。具体的には、各画像に左右反転、上下反転、右 90 度回転、左 90 度回転、180 度回転の処理を行う。これによって学習に

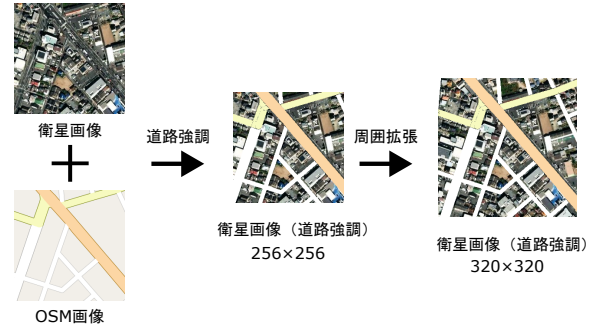


図 3: 画像の作成手順

用いる画像データ数が 6 倍になる。

3 CNN モデル

画像内の 8×8 個のブロックごとに交通事故危険性を判別する CNN モデルを提案する。提案モデルの構造を図 4 に示す。提案モデルは、320×320 画素の画像を入力として受け取ると、最初に 5 つの ResLayer に通す。図 5 に ResLayer の構造を示す。ResLayer は 2 つの ResBlock をつなげた構造になっている。最初の ResBlock では畳み込みと同時にダウンサンプリングを行い、2 番目の ResBlock では畳み込みのみを行う。5 つの ResLayer を通したら 10×10×256 の特徴マップを得る。この特徴マップにカーネルサイズ 3×3 の畳み込みを行い、8×8×512 の特徴マップを得る。この特徴マップは 8×8 個のブロックに対応し、各ブロックの周辺を畳み込んだ情報も持っていると考えられる。特徴マップから各ブロックに対応する部分を抽出して 1×1×512 の特徴マップを得る。画像全体の特徴も交通事故危険箇所の抽出に有用と考えられるため、10×10×256 の特徴マップに大域平均プーリングを行い、1×1×512 の特徴マップと統合して 1×1×768 の特徴マップを得る。1×1×768 の特徴マップの次元削減を行い、768 の大きさの 1 次元特徴マップに変換する。そして全結合層を 2 つ通してブロックの危険性を判別する。全結合層に入力する直前には過学習を抑えるため 2 割の Dropout [10] を行う。出力を Sigmoid 関数に通して 0 から 1 の範囲にする。これらの処理を 64 個のブロックすべてに行う。

4 性能評価実験

4.1 実験設定

提案した枠組みの有効性を検証するため、2 種類の CNN モデルと 3 種類の画像データを用いて判別性能評価実験を行った。CNN モデルと画像データの組み合わせは以下の 6 通りである。

1. ResNet18 + OSM 画像
2. ResNet18 + 衛星画像
3. ResNet18 + 衛星画像 (道路強調)
4. 提案モデル + OSM 画像
5. 提案モデル + 衛星画像
6. 提案モデル + 衛星画像 (道路強調)

表 1 にはデータセット内の画像種別ごとの画像数を示す。安全画像はすべてのブロックが安全とラベル付けされた画像であり、危険画像は危険とラベル付けされたブロックを 1 個以上含む画像である。ラベル付けされた画

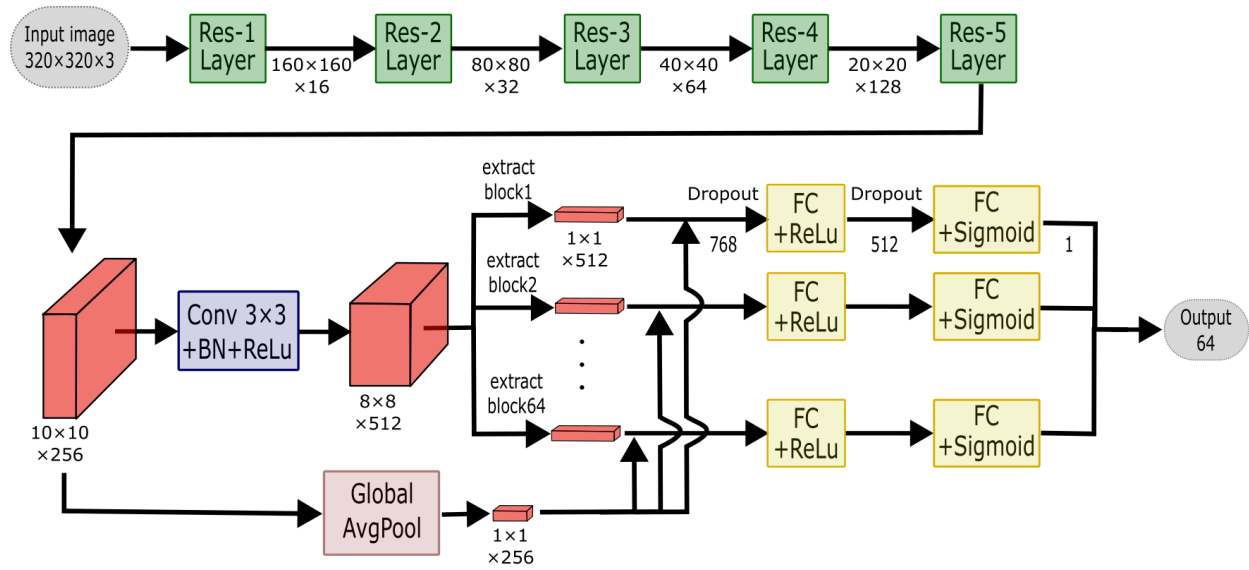


図 4: 提案モデル

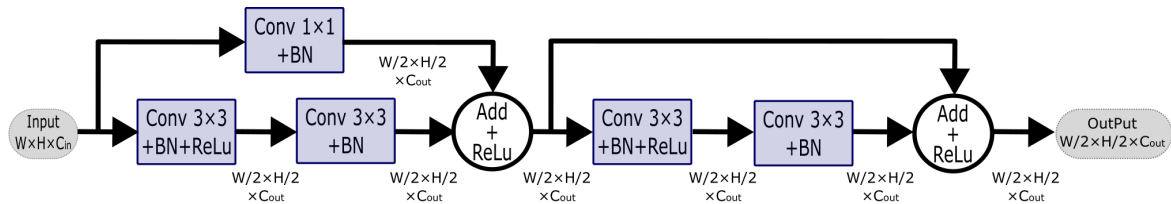


図 5: ResLayer

表 1: データセット内の画像数

画像種別	枚数
全画像	7,454
安全画像	3,727
危険画像	3,727
訓練画像 (データ拡張前)	5,962
訓練画像 (データ拡張後)	35,772
訓練画像 (安全)	17,886
訓練画像 (危険)	17,886
テスト画像	1,492
テスト画像 (安全)	746
テスト画像 (危険)	746

像の中から、安全画像と危険画像の枚数が等しくなるように 7,454 枚を選んだ。選んだ画像を 8 対 2 の割合で訓練用とテスト用に分け、訓練用にはデータ拡張を行う。データ拡張後の訓練画像とテスト画像をそれぞれ CNN の学習と評価に用いる。

CNN の学習には最適化アルゴリズム Adam [11] を用いる。ハイパーパラメータである学習率と重み減衰係数の値は、Optuna [12] とよばれるブラックボックス最適化フレームワークを用いて探索した。その際、学習率の探索範囲を 10^{-7} から 10^{-1} までに設定し、重み減衰係数の探索範囲を 10^{-10} から 10^{-3} までに設定した。探索結果を表 2 に示す。CNN の学習時には表 2 に示される値を用いる。

損失関数には交差エントロピー誤差を用い、ミニバッチごとの損失を最小にするように CNN のパラメータ値を更新する。ただし、作成した画像データセットにおいてはラベル 1 とラベル 0 のサンプル数に大きな差がある

表 2: Optuna で得られた学習率と重み減衰係数の値

CNN モデル	画像	学習率	重み減衰
ResNet18	OSM	4.392×10^{-4}	2.874×10^{-9}
ResNet18	衛星	1.523×10^{-3}	1.949×10^{-7}
ResNet18	衛星 (道路強調)	1.539×10^{-3}	4.916×10^{-5}
提案モデル	OSM	4.468×10^{-4}	3.101×10^{-7}
提案モデル	衛星	1.135×10^{-3}	7.219×10^{-7}
提案モデル	衛星 (道路強調)	1.092×10^{-3}	5.032×10^{-4}

ため、ラベル 0 と判定する傾向が強くなり、ラベル 1 と判定することがほとんどないという問題が起こる。この問題を回避するにはミニバッチごとの損失の算出法を適切に決める必要がある。本研究では以下の方法を用いる。ミニバッチ内の画像の中でブロック b にラベル 1 が付与されているものの枚数を N_1^b とし、それらの画像のラベル 1 に対する損失の和を L_1^b とする。同様に、ブロック b にラベル 0 が付与されている画像の枚数を N_0^b とし、それらの画像のラベル 0 に対する損失の和を L_0^b とする。バッチサイズは $N_1^b + N_0^b$ に等しい。ブロック b に対する損失を

$$L^b = \left(\frac{L_0^b}{N_0^b} + \frac{L_1^b}{N_1^b} \right) \times \frac{1}{2}$$

で求める。この方法ですべてのブロックの損失を求め、その平均値をパラメータ値の更新に使用する。

本研究では CNN の判別性能を向上させるために図 6 に示すような 5 分割交差検証を行う。訓練画像を 5 分割して 1 つを検証用とし、残りの 4 つを訓練用とする。検証用画像はモデルの各エポックでの精度を推定するために使用する。検証用画像の選び方は 5 通りあるため 5 種

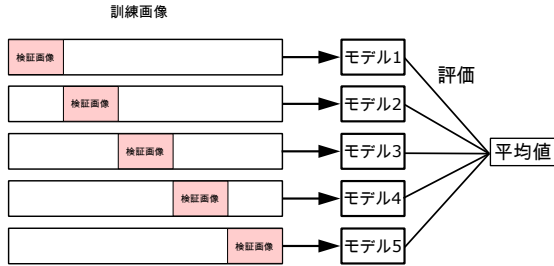


図 6: 交差検証

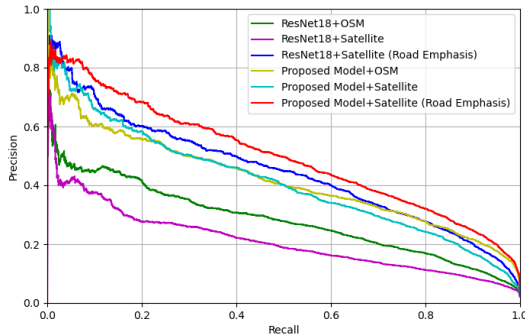


図 7: PR (Precision-Recall) 曲線

類のデータセットが得られ、それぞれを用いて CNN の学習を行うと 5 種類の学習済み CNN モデルが得られる。それら 5 種類のモデルにテスト画像を入力して得られる出力の平均値から交通事故危険性を判別する。

4.2 実験結果

CNN モデルと画像の組み合わせごとの PR (Precision-Recall) 曲線を図 7 に示す。Recall (再現率) は危険とラベル付けされたブロックのうち、学習済み CNN が危険と判定したブロックの割合であり、Precision (適合率) は学習済み CNN が危険と判定したブロックのうち、危険とラベル付けされているブロックの割合である。表 3 には各組み合わせの PR-AUC (Area Under the Precision-Recall Curve) の値を示す。PR-AUC は PR 曲線の下の面積であり、0 から 1 までの値をとる。

表 3 から、衛星画像をそのまま用いると OSM を用いる場合よりも PR-AUC が低くなり、衛星画像の道路部分を強調した画像を用いると OSM を用いる場合よりも高くなるのがわかる。衛星画像をそのまま用いたときの CNN の判別結果を見ると、道路以外の部分を危険箇所と判定している。衛星画像は道路部分が不鮮明なために学習や判別がうまくできていないと考えられる。

表 3 において ResNet18 と提案モデルを比較すると、後者の方が PR-AUC が高いことがわかる。ResNet18 では最終結合層の前にプーリング層をおき、認識する物体の位置のずれに対して影響を受けないようにしている。しかし、ブロックごとに交通事故危険性を判別する本研究の問題では、画像から抽出した特徴の位置情報が重要であるため、それを考慮して学習と判別を行う提案モデルの方が PR-AUC が高くなったと考えられる。

PR-AUC が最も高かった提案モデルと衛星画像 (道路強調) の組み合わせで学習した CNN と、2 番目に高かった ResNet18 と衛星画像 (道路強調) の組み合わせで学

表 3: PR-AUC

モデル	画像	PR-AUC
ResNet18	OSM	0.2828
ResNet18	衛星	0.2107
ResNet18	衛星 (道路強調)	0.4473
提案モデル	OSM	0.4164
提案モデル	衛星	0.4114
提案モデル	衛星 (道路強調)	0.4968

習した CNN による危険箇所抽出結果の一例を図 8 に示す。右端の列が提案モデルと衛星画像 (道路強調) の組み合わせで学習した CNN による抽出結果であり、右から 2 番目の列が ResNet18 と衛星画像 (道路強調) の組み合わせで学習した CNN による抽出結果である。この図では、CNN の出力が 0.9 を超えたブロックを危険とし、そうでないブロックを安全としている。また、危険と判定されたブロックには赤色を重ねている。どちらの組み合わせにおいても、国道や主要地方道路を表す色付きの道路や交差点を含むブロックを危険と判定し、周辺に田畑や自然が多いブロックは安全と判定する傾向がある。その一方で、CNN が安全と判定する交差点もあるし、安全ラベルが付与されているのに危険と判定しているブロックも多くある。

5 おわりに

交通事故データを用いて地図画像や衛星画像にラベルを付与し、それらのラベル付き画像を CNN モデルに学習させ、交通事故危険度の高い小領域を抽出する方法を提案した。3 種類の画像と 2 種類の CNN モデルの組み合わせごとに CNN の判別性能を評価する実験を行ったところ、衛星画像の道路部分のみを OSM 画像に置き換えた画像を、本研究で設計した CNN に学習させる場合が最も高い判別精度を示した。しかし、この組み合わせでも誤った判別を行うことが少なくない。今後の課題は、ラベル付与方法の見直しや CNN モデルの修正によって、判別精度向上を実現することである。

謝辞

本研究は、岡山県警察、株式会社トスコ、岡山大学による STRADA (Studies on TRaffic Accident Data Analysis) プロジェクトの支援および JSPS 科研費 JP23K21703 の助成を受けたものである。ここに謝意を表す。

参考文献

- [1] 警察庁交通局交通企画課, 令和 5 年中の交通事故死者数について, <https://www.npa.go.jp/news/release/2024/20240104001jiko.html> (2024 年 6 月 13 日アクセス)。
- [2] 中央交通安全対策会議, 第 11 次交通安全基本計画, 2021。
- [3] 呉志堅, 右田剛史, 高橋規一, “交通事故危険箇所予測のための地図画像の深層学習,” 2021 年度 (第 72 回) 電気・情報関連学会中国支部連合大会公演論文集, 2 pages, 2021。
- [4] OpenStreetMap, <https://www.openstreetmap.org> (2024 年 6 月 13 日アクセス)。
- [5] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv:1409.1556, 2014.
- [6] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.
- [7] K. Arase, Z. Wu, T. Migita and N. Takahashi, “Deep learning of OpenStreetMap images labeled using road traffic accident data,”

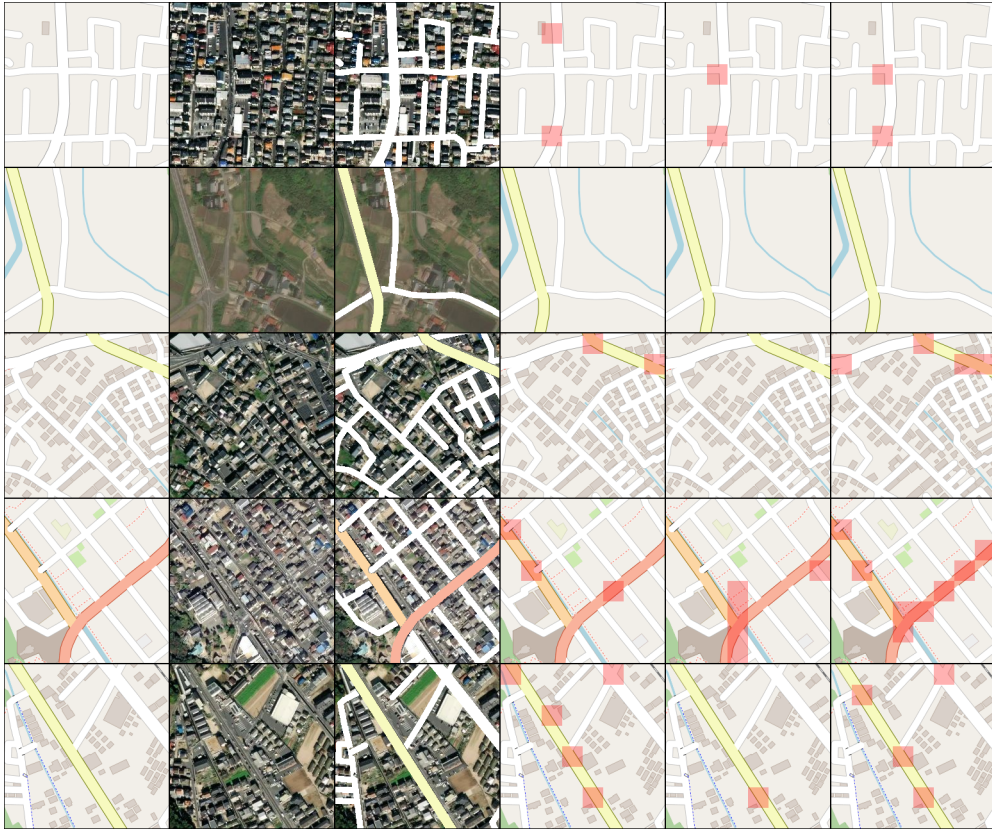


図 8: 交通事故危険箇所抽出結果 (左から OSM 画像, 衛星画像, 衛星画像 (道路強調), 正解ラベル, ResNet18 + 衛星画像 (道路強調) の抽出結果, 提案モデル+衛星画像 (道路強調) の抽出結果)

Proceedings of the 2022 IEEE Region 10 Conference, 6 pages, 2022.

- [8] R.R. Selvarahu, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, “Grad CAM: Visual explanations from deep networks via gradient-based localization,” *Proceedings of the IEEE Conference on Computer Vision*, pp.618–626, 2017.
- [9] Mapbox, <https://www.mapbox.com> (2024 年 6 月 13 日アクセス) .
- [10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol.15, no.1, pp.1929–1958, 2014.
- [11] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv:1412:6980, 2015.
- [12] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework,” *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.2623–2631, 2019.