

データフュージョンによる交通状況予測システム

2020SC070 榊原桃花

指導教員：河野浩之

1 はじめに

第 11 次愛知県交通安全計画 [1] によると、2021 年から 2025 年の交通事故による死者数の目標は 125 人以下である。2021 年では死者数が 117 人となり、目標の 125 人以下を達成した。しかし、翌年の 2022 年には死者数が 137 人となり、目標を大幅に下回った結果となった。その目標を達成するためには、こういった場所で交通事故が発生しやすいかを予測するためのシステムの確立と全体の意識改革が必要である。

本研究では、過去に発生した交通事故の緯度・経度の情報と衛星画像を用いてデータフュージョンを行うことで 1 つのデータセットを作成し、そのデータセットを 2 つのモデルに学習させることで交通事故発生場所の予測システムの一部を構築を行い、モデルの比較も同時に行った。交通事故発生場所の予測によって、交通事故の削減と予防を死者数の減少に繋げる。

2 データフュージョンに関する先行研究

データフュージョンを用いた先行研究について、次の表 1 に示す。鳥居ら [2] は道路、交通量、企業数などの数値データと Google Map から得た衛星画像データを融合して用いて交通事故発生場所の予測を行い、77% の精度を達成した。衛星画像データは訓練済み CNN モデルの Dense-121 を用いて特徴抽出を行うことで特徴を数値化した。その後の数値データと数値化された衛星画像データを lightGBM モデルで学習させた。

また、Songtao He ら [3] は、高解像度の衛星画像データ、GPS の軌跡、道路の地図や過去の交通事故情報を用いて End-to-End モデルに学習させたところ、ロサンゼルス・ニューヨーク・シカゴ・ボストンの 4 都市の平均適合率は 47.67 % であった。

表 1: データフュージョンに関する先行研究

研究者	手法
鳥居ら [2]	Dense-121(CNN), lightGBM
Songtao He ら [3]	End-to-End

3 交通状況予測システムの提案

本研究では、作成したデータセットにおける道路の形状や大きさ・色彩を中心に分析を行うことで入力された画像から交通事故の発生しやすさを 3 段階で予想することを可能にすることで交通状況の予測システムの 1 部を構築する。

3.1 交通事故の発生しやすさの定義

交通事故の発生しやすさを過去の交通事故の緯度・経度情報を用いて作成したデータセットをもとに 3 段階に分類した。はじめに、交通事故が発生していない地点の画像を「交通事故なし」とした。次に、交通事故が 1 件から 3 件の再現性の少ない地点の画像を「交通事故 1 件から 3 件」とした。最後に、交通事故が 4 件以上の頻発して発生している地点の画像を「交通事故 4 件以上」として学習させる際のラベル付けを行った。

3.2 CNN と ViT の環境構築

本研究では精度向上の手立てを見つけやすくするため、2 つのモデルで学習させたのち精度を比較した。1 つ目のモデルは CNN(畳み込みニューラルネットワーク) である。CNN は主に画像データに特化したモデルであることから高い精度を見込むことができる。2 つ目のモデルは ViT(Vision Transformer) である。ViT は自然言語系に特化した Transformer を画像に応用したもので畳み込みを使用しないモデルである。

どちらのモデルも機械学習に便利なライブラリが既に環境構築がされている google colabatory で Python3.10.12 で構築を行った。

4 CNN と ViT による 3 値分類

4.1 データフュージョンによるデータセットの作成

本研究では、過去の交通事故の緯度・経度の情報とその交通事故発生地点の衛星画像を融合して新たなデータセットを作成する。過去の交通事故情報は 2019 年度から 2022 年度の交通事故統計情報のオープンデータセット [4] から CSV 形式のファイルをダウンロードした。また、衛星画像については地理空間情報ソフトウェアである QGIS を用いて投影した。同時にダウンロードした交通事故情報を含むファイルを重ねて投影し、地図帳印刷という機能を用いて画像データをフォルダ内に作成することができる。

また、交通事故が発生していない画像データも学習とテストに必要であるため、緯度・経度をランダムに発生させた CSV 形式のファイルを作成し、同様に QGIS 上で衛星画像と作成したファイルを重ねて投影することで画像データを作成することができる。下の図 1 に交通事故が発生したときと発生していないときの画像データの一部を示す。

このとき、図 1 の (a) のように交通事故発生地点には丸い印をつけているが交通事故の程度別にラベル付けを行う際に利用するものとする。実際にデータセットを使用する際には丸を消去したものを利用する。



(a) 交通事故あり (b) 交通事故なし

図 1: 取得した画像データの一部

4.2 作成したデータセットの内訳

作成した画像データを「交通事故なし」、「交通事故 1 件から 3 件」、「交通事故 4 件以上」の 3 つラベルに分けてデータセットを作成した。ラベル毎の画像データの枚数は次の表 2 のようになる。

表 2: ラベル別データセットの枚数

ラベル (交通事故)	なし	1 件から 3 件	4 件以上
枚数	446	634	475

4.3 CNN による 3 値分類の精度

google colabatory 上で機械学習向けのライブラリである Keras と Tensorflow を用いて構築した CNN に作成したデータセットを入力して学習を行った。テストデータに対する予測ラベルと正解ラベルを可視化したヒートマップを次の図 2 に示す。

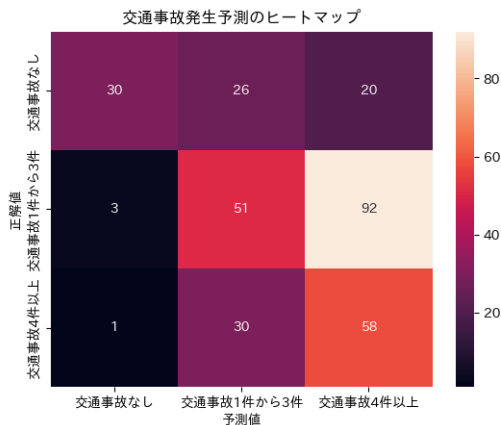


図 2: CNN による 3 値分類のヒートマップ

図 2 から精度は約 44.7 %であった。また、交通事故発生の再現率は正解値が交通事故が発生したものの 235 枚のうち、予測値が交通事故が発生したものが 231 枚だったため、約 98.3 %と高くなっている。

4.4 ViT による 3 値分類の精度

google colabatory 上で ViT のパッチサイズが 16 である事前学習済みモデルを構築し、作成したデータセット

を入力して学習を行った。テストデータに対する予測ラベルと正解ラベルを可視化したヒートマップを次の図 3 に示す。

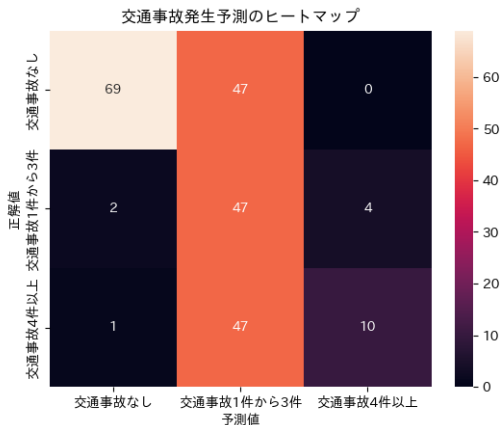


図 3: ViT による 3 値分類のヒートマップ

図 3 から精度は約 55.5 %であった。交通事故発生の再現率は正解値が交通事故が発生したものの 111 枚のうち、予測値が交通事故が発生したものが 108 枚だったため、約 97.3 %と高くなっている。

5 むすび

本研究では、第 11 次愛知県交通安全計画の目標を達成するために交通状況予測システムの一部として画像による交通事故発生しやすさの予測を 3 値分類で 2 つのモデルを用いて行った。その結果、ViT の精度が CNN を約 10 %を上回り、約 55.5 %となった。その理由としては ViT は事前学習モデルを使用しているため、データ数の少ない学習でも対応できたと考えられる。データセットに対してセマンティックセグメンテーションを用いることでの物体認識を行うことで精度向上が期待できる。

参考文献

- [1] 愛知県 . “第 11 次愛知県交通安全計画”. <https://www.pref.aichi.jp/uploaded/attachment/386150.pdf>, (参照 2023.11.27)
- [2] 鳥居 和史, 水野 由裕, 遠山 一則, 清水 茂樹, 向後 颯太, “構造化データと衛星画像のマルチモーダル AI モデルによる交通事故発生確率の予測手法の開発”, 自動車技術会論文集, vol.53, no.2, pp.404-409, 2022.
- [3] Songtao He . Mohammad Amin Sadeghi . Sanjay Chawla . Mohammad Alizadeh . Hari Balakrishnan . Samuel Madden . “Inferring high-resolution traffic accident risk maps based on satellite imagery and GPS trajectories”. IEEE , October 2021 .
- [4] 警視庁 . “交通事故統計情報のオープンデータ”. https://www.npa.go.jp/publications/statistics/koutsuu/opendata/2022/codebook_2022.pdf, (参照 2023.01.06)