
愛知県の犯罪オープンデータと地理的特徴量を利用した 機械学習による犯罪種別の学習と予測

向 直 人

概要

愛知県警察は、防犯活動の促進を目的として、2018年から愛知県内で発生した犯罪の情報を犯罪オープンデータ¹⁾として公開している。犯罪オープンデータは、罪名、手口、発生年月日、発生場所の属性、市区町村コードなどの情報で構成されており、愛知県内の犯罪のおおまかな傾向を把握することが可能である。そこで、本研究では、機械学習の分類アルゴリズムを利用して、犯罪オープンデータから抽出された特徴量を基に、犯罪種別の予測を試みる。犯罪発生場所の地図から独自に算出した「道路面積比」「建物面積比」などを地理的特徴量として用いる。分類アルゴリズムのランダム・フォレストを採用し、「自転車盗」「車上ねらい」「自動車盗」の3種類の犯罪種別を学習させた結果、52.74%の精度で正しく分類することができた。犯罪の「発生時」や「道路面積比」などが、犯罪種別の分類に強く寄与していることがわかった。

1 はじめに

欧米では犯罪が発生した時間・場所などの履歴データを活用することで、パトロール業務などを最適化するなど、防犯を目的としたシステムの構築が1980年代から進められてきた。近年、米国立司法省研究所(National Institute of Justice: NIJ)²⁾は、2016年の「Crime Forecasting Challenge(犯罪予測チャレンジ)」、2021年の「Recidivism Forecasting Challenge(再犯予測チャレンジ)」などにおいて、過去の犯罪データを一般に公開し、犯罪予測の精度を競うコンペティションを開催している。コンペティションの結果を、警察組織や地域コミュニティに還元することで、犯罪を防止し、より安全な社会の構築を試みている。一方、

日本においては、京都府警が2016年から「予測型犯罪防御システム」の運用を開始したと発表している[1]。このシステムでは、犯罪の発生情報を集約・分析して、将来の犯罪を予測することを可能としており、パトロール地域の選定や、犯罪の未然防止に活用されている。また、各都道府県は犯罪履歴をオープンデータとして積極的に公開することを進めている。愛知県警察は、2018年からの犯罪発生情報をオープンデータとして公開しており、犯罪傾向の把握や、地域の防犯力の向上に注力している。

従来の犯罪予測では、犯罪が発生した地域や頻度に応じて確率密度分布を推定するカーネル密度推定や、駅・駐車場・金融機関などの環境要因から犯罪リスクを推定するRisk Terrain Modeling(RTM)が主に用いられる[2]。これらの手法では、過去の犯罪発生地点と環境要因の重ね合わせで、

犯罪リスクが算出される。一方で、犯罪が発生した地域の道路や建物など地理的な特徴も、犯罪発生に寄与していると考えられる。例えば、建物が多い都市の中心部では放置自転車が多く「自転車盗」、また、郊外の人目につきにくい駐車場では自動車を盗む「自動車盗」が発生しやすいと考えられる。

そこで、本研究では、愛知県警察が公開している犯罪オープンデータに含まれる時刻や住所など犯罪に関する基本的な特徴量に加え、犯罪発生地域の道路や建物などの地理的な特徴量も考慮して、機械学習による犯罪種別の分類を試みる。地理的な特徴量は、地図開発プラットフォームの Mapbox³⁾ を利用して、犯罪発生地域の地図画像を生成し、地図画像を画像処理することにより抽出する。分類アルゴリズムとして、決定木のアンサンブル・モデルであるランダム・フォレストを採用した。発生件数の多い「自転車盗」「車上ねらい」「自動車盗」の3種類の犯罪種別に着目し、サンプル数 $N=3000$ の独自のデータセットで評価実験を行った。評価実験では、分割数 $k=20$ の k -分割交差検証により、犯罪種別の分類の正解率を算出した。

本稿の構成は以下である。第2章で、犯罪予測に関する既存研究について示し、本研究の独自性について述べる。第3章で、愛知県警察が公開している犯罪オープンデータの詳細を明らかにする。第4章で、ランダム・フォレストを利用した犯罪種別の予測に用いる特徴量について述べる。特徴量は犯罪オープンデータに含まれる特徴量と、犯罪発生地域の地図画像から抽出した特徴量で構成される。第5章で、独自に構築したデータセットを用いた評価実験について述べ、犯罪種別の予測精度について報告する。最後に、第6で本稿をまとめる。

2 関連研究

本章では、犯罪予測に関する既存研究について述べる。大山らは、ある地域における犯罪の発生リスクを予測する研究について広くまとめている [2]。現在の研究は「時空間クラスタの検出」「犯罪の時空間相互作用を考慮した犯罪発生強度推定」「環境要因からの犯罪発生リスク推定」「回帰分析を利用した犯罪発生件数・確率の推定」の4つのカテゴリに分類できると述べている。本稿で提案する手法は、過去の犯罪データから犯罪種別を分類することを目的とするため、上記のいずれのカテゴリにも分類されないが、地理的な特徴量を用いることから「環境要因からの犯罪発生リスク推定」の考え方を包含している。

犯罪の時空間クラスタや時空間相互作用を推定するために主に用いられる手法がカーネル密度推定である。中川らは、パトロール業務支援を目的として、時空間のカーネル密度推定を利用して、高い頻度で犯罪が発生する地域(ホットスポット)を予測する手法を提案した [3]。この手法では、カーネル関数を、時間成分と空間成分の積でモデル化し、パトロール可能な面積を考慮してパラメータを最適化した。この結果、提案手法が特に侵入盗に対して有効であることを明らかにした。また、野らは、同様に効果的なパトロール経路を導出することを目的として、カーネル密度推定とRTMを組み合わせた手法を採用した [4]。RTMは環境要因から犯罪を予測する手法であり、駅、基幹道路、駐車場など10種類の要因を考慮している。導出されたパトロール経路による犯罪数の有意な減少は確認されなかったが、アンケート結果から、提案手法は効果的であると感じる被験者が多いことが明らかとなった。これらの手法は、過去の犯罪データから、将来、犯罪が発生しやすい地域や時間帯を予測するものである。本研究に

においても、犯罪の発生日時や住所などの空間的特徴量を利用するが、特定の地域において、発生確率が高い犯罪種別を推定することを目的としている。

近年は深層学習（ディープラーニング）を犯罪予測に応用した研究が報告されている。山下らは、Google Street Viewで取得した画像から、畳み込みニューラルネットワーク（Convolutional Neural Network: CNN）を利用して特徴量を抽出し、犯罪発生の予測に用いる手法を提案した [5]。抽出された特徴量を、サポートベクタマシン（Support Vector Machine: SVM）によって学習させることで、2%程度の正解率の改善が確認できたと述べている。また、張らは、犯罪発生の時系列変化に注目し、LSTMネットワーク（Long Short Term Memory Network: LSTM Network）を利用した予測モデルを構築した [6]。東京都が公開している犯罪オープンデータを利用し、「空き巣」「自転車盗」「万引き」を学習させたLSTMネットワークは、シンプルなりカレント・ニューラルネットワークよりも高い精度で予測できることを明らかにした。本研究では、山下らのアプローチと同様に、犯罪発生地域の地図画像をMapboxを利用して生成する。生成された地図画像を画像処理することで、地理的な特徴量を抽出し、ランダム・フォレストの学習に用いる。

3 愛知県警察の犯罪オープンデータ

愛知県警察は2018年から犯罪オープンデータを公開している。ここでは、犯罪オープンデータの統計的な特徴を確認し、データの全体的な傾向を明らかにする。犯罪オープンデータは、「ひったくり」「車上ねらい」「部品ねらい」「自動販売機ねらい」「自動車盗」「オートバイ盗」「自転車盗」の7種類の犯罪種別（手口）で構成されており、2018年から2020年までの3年間を合わせると45,456件のレコードが記録されている。各レコードは、罪名、犯罪種別（手口）、管轄警察署、管轄交番・駐在所、市区町村コード、都道府県、市区町村、町丁目、発生年月日、発生時、発生場所の11種類の属性で構成されている⁴⁾。愛知県が公開しているデータの「罪名」は全て「窃盗」である。また、犯罪の発生場所は、事件関係者のプライバシーの保護するため、町丁目までの情報しか含まれない（緯度・経度ではないことに注意されたい）。表1が、犯罪オープンデータのサンプルである（罪名、管轄警察署、管轄交番・駐在所、市区町村コード、都道府県を除く）。

犯罪種別の内訳を表2に示す。「自転車盗」の件数が最も多く、次いで車内の荷物や現金を盗む「車上ねらい」、車に取り付けてある部品を盗む「部品ねらい」の順に多い。一方で、路上で持ち物を奪う「ひったくり」は最も件数が少ないことがわかる。次に、犯罪件数が多い市区町村を表3に示

表1 犯罪オープンデータのサンプル

犯罪種別	市区町村	町丁目	発生年月日	発生時	発生場所
ひったくり	名古屋市西区	上名古屋3丁目	2019-01-18	21	道路上
車上ねらい	名古屋市西区	香呑町4丁目	2019-05-23	20	駐車（輪）場
部品ねらい	名古屋市西区	上堀越町3丁目	2019-05-23	8	駐車（輪）場
自動販売機ねらい	名古屋市西区	稲生町3丁目	2019-05-21	23	その他
自動車盗	名古屋市東区	泉1丁目	2019-08-11	12	道路上

す。名古屋の繁華街である「名古屋市中区」の犯罪件数が最も多い。次いで名古屋市ベッドタウンである「一宮市」、交通拠点である名古屋駅がある「名古屋市中村区」の順に多い。ここで、上述の3地域で発生した犯罪種別に注目する。表4に3地域の犯罪種別ごとの件数と割合をまとめた。名古屋市中区と名古屋市中村区では「自動車盗」が75%以上を占めているのに対し、一宮市では「自動車盗」は約55%に留まっていることがわかる。また、名古屋市中区と名古屋市中村区では「部品ねらい」が6%以下であるのに対し、一宮市では13%に達しており、2倍以上の差がある。このことは、犯罪種別には地域差が存在することを示しており、地域に応じて注力すべき犯罪

種別が異なると言える。この地域差の原因は様々なと考えられるが、本稿では、駅や駐車場などの環境要因や、道路や建物などの地理的要因が関係していると仮説を立て、機械学習の特徴量として用いる。

4 分類に用いる特徴量の生成

本章では、犯罪種別の分類を実現するランダム・フォレストの学習に関して述べる。ここでは、7種類ある犯罪種別のうち、発生頻度の高い「自転車盗」「車上ねらい」「自動車盗」の3種類に絞って分類を試みる。各犯罪種別ごとに1000件のレ

表2 犯罪種別と件数（2018年～2020年）

犯罪種別	犯罪件数
ひったくり	267
車上ねらい	7070
部品ねらい	5685
自動販売機ねらい	1620
自動車盗	2008
オートバイ盗	1847
自転車盗	28059

表3 犯罪件数が上位の市区町村（2018年～2020年）

市区町村	犯罪件数
名古屋市中区	2749
一宮市	2019
名古屋市中村区	1922
春日井市	1896
豊橋市	1790
豊田市	1692
名古屋市中川区	1686
岡崎市	1557
名古屋市北区	1374
名古屋市西区	1362

表4 市区町村の犯罪種別の割合

犯罪種別	名古屋市中区	一宮市	名古屋市中村区
自転車盗	2204 (80.44%)	1101 (54.53%)	1442 (75.03%)
車上ねらい	209 (7.63%)	349 (17.29%)	211 (10.98%)
自動販売機ねらい	105 (3.83%)	96 (4.75%)	52 (2.71%)
部品ねらい	80 (2.92%)	269 (13.32%)	103 (5.36%)
オートバイ盗	61 (2.23%)	66 (3.27%)	49 (2.55%)
自動車盗	53 (1.93%)	129 (6.39%)	40 (2.08%)
ひったくり	37 (1.35%)	9 (0.45%)	25 (1.30%)

コードを無作為に抽出し、サンプル数 $N=3000$ のデータセットを構築する。学習に用いる特徴量は、犯罪オープンデータに含まれる基本的な特徴量と、Mapboxで生成した地図画像から抽出した特徴量を用いる。下記に特徴量の詳細を述べる。

4.1 犯罪オープンデータに含まれる特徴量

犯罪オープンデータから、「都道府県」「市区町村」「町丁目」「発生時」「発生場所」を特徴量として用いる。「発生時」が不明となっている22件のレコードはデータセットから除外する。「発生時」のみは、そのまま学習に用いるが、他の4つの特徴量は次のように変換して学習に用いる。

4.1.1 位置を表す特徴量への変換

「都道府県」「市区町村」「町丁目」は結合して、住所を表す文字列に変換する。この住所を表す文字列を、Mapbox Geocoding API⁵⁾に送信し、対応する緯度・経度を取得する(ジオ・コーディング)。取得した緯度・経度のサンプルを表5に示す。上述したように住所は「町丁目」までの情報しか含まれていない。このため、同一の緯度・経度であっ

ても、異なる犯罪種別のレコードが含まれており、緯度・経度だけでは唯一に同定できない。例えば、「愛知県名古屋市中区栄3丁目(緯度:35.16285、経度:136.907699)」で記録されているレコードは23件存在し、その犯罪種別は16件が「自転車盗」、6件が「車上ねらい」、1件が「自動車盗」である。また、取得される緯度・経度には誤差があるため、信頼度(relevance)が0.8以上のレコードのみをサンプルに採用した。

さらに、緯度・経度を平面直角座標に変換する。全レコードの緯度・経度の平均を原点(0, 0)とし、各レコードに対する東西方向と南北方向のユークリッド距離(単位はメートル)を座標(x , y)として表す。平面直角座標に変換したサンプルを表6に示す。この変換により、犯罪発生地域の位置関係が直感的に理解しやすくなることに加え、機械学習を用いた分類においても効果的となる。

4.1.2 発生場所をダミー変数に変換

犯罪オープンデータに含まれる「発生場所」は、質的変数であり、「道路上」「駐車(輪)場」「一

表5 住所から緯度経度への変換

都道府県	市区町村	町丁目	緯度	経度
愛知県	名古屋市西区	貴生町	35.223248	136.875789
愛知県	名古屋市西区	城西2丁目	35.186802	136.894798
愛知県	名古屋市西区	名駅2丁目	35.176948	136.885773
愛知県	名古屋市西区	栄生1丁目	35.185583	136.878758
愛知県	名古屋市西区	栄生2丁目	35.186094	136.874471

表6 緯度・経度から平面直交座標への変換

都道府県	市区町村	町丁目	X	Y
愛知県	名古屋市西区	貴生町	11320.125496	11320.125496
愛知県	名古屋市西区	城西2丁目	7272.093674	7272.093674
愛知県	名古屋市西区	名駅2丁目	6181.266758	6181.266758
愛知県	名古屋市西区	栄生1丁目	7141.023289	7141.023289
愛知県	名古屋市西区	栄生2丁目	7198.846970	7198.846970

戸建住宅」「4階建以上共同住宅」「その他の住宅(3階建以下共同住宅等)」「その他」の6種類のいずれかの値をとる。質的変数をそのまま機械学習に用いることはできないため、表7のようにダミー変数に変換する。各値に応じた属性を設け、「1」または「0」で、対応する発生場所を表現する。

4.2 地図画像に基づく特徴量

上述した犯罪発生位置を表す緯度・経度から、Mapbox Static Tiles API⁶⁾を利用して地図画像を生成する。地図画像のサイズは256×256ピクセルであり、1辺が約234メートルを表す(0.915 meters/pixel)。住所は町丁目までしか情報がないため、正確な犯罪発生位置は不明であり、生成した地図画像内で発生したとは限らない。しかし、地図画像は犯罪が発生した周辺のおおまかな地理的特徴を含むと考えることができる。図1が生成された地図画像のサンプルであり、白、黒、緑、

赤、青の5色で塗り分けされている。各色は表8に示す領域(駐車場のみは点)を表している。Mapboxの地図データにおいて、駐車場は、面積のある領域データではなく、座標を表す点データとして記録されている。このため、各座標に同一半径の円を表示することで、駐車場を表すことにした。また、一部の駐車場しか地図データには含まれておらず、不完全な情報であることに注意されたい。

次に生成された地図画像から、学習に用いる特徴量を抽出する。抽出する特徴量は「ノード数」「エッジ数」「密度」「平均次数」「道路面積比」「建物面積比」「駅」「駐車場」の8種類である。

4.2.1 道路ネットワークに基づく特徴量

地図画像から道路ネットワークを抽出し、「ノード数」「エッジ数」「密度」「平均次数」を算出する。道路ネットワークの抽出結果を図2に示す。最初

表7 発生場所をダミー変数に変換

道路上	駐車場	一戸建住宅	4階建以上共同住宅	3階建以下共同住宅	その他
0	1	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
0	1	0	0	0	0
0	0	0	0	0	1

表8 地図画像の色

色	対象
白	道路領域
黒	背景領域
緑	建物領域
赤	駅領域
青	駐車場(点)



図1 地図画像のサンプル

(© Mapbox, © OpenStreetMap)

に地図画像の建物領域、駅領域、駐車場を背景領域と同じ黒で塗りつぶし、道路と背景領域の2値画像に変換する。この2値画像にZhangらの細線化アルゴリズム [7] を適用し、幅が1ピクセルのスケルトン画像に変換する。さらに、Boeingらのネットワークの抽出アルゴリズム [8] を適用し、ノードとエッジで構成される道路ネットワークを抽出する。

図2から算出された特徴量を表9にまとめる。ノード数 N とエッジ数 E は、ネットワークを構成しているノードとエッジの数である。ノード数は交差点数を表すが、エッジの端点も交差点としてカウントされる。密度 D は、可能なエッジの組み合わせに対する実際のエッジの比率を表す。ここで $N=33$ であり、可能なエッジの組み合わせは $N(N-1)/2=528$ 通りある⁷⁾。実際に存在するエッジ数は $E=43$ であることから、密度 D は $43/528 \approx 0.08$ となる。密度が高いネットワークは、道路が

網目状に相互に接続された状態であり、幅員の狭い生活道路であることが考えられる。平均次数 k は、ノードに接続されるエッジ数の平均であり、ここでは $(43 \times 2)/33 \approx 2.61$ となる。平均次数は交差点の分岐数を表し、交差点の少ない直線的な道路では小さな値となる。

4.2.2 地図画像の面積比に基づく特徴量

地図画像に占める道路や建物の面積から「道路面積比」「建物面積比」を算出する。また、駅や駐車場は、犯罪を誘引する環境要因であると考えられ、野らの研究 [4] らの研究においても利用されている。そこで、駅と駐車場の存在の有無を表す「駅」「駐車場」も特徴量とする（存在するときは1、そうでなければ0で表す）。図3に示す「名古屋市西区上名古屋4丁目」と「小牧市中央1丁目」の特徴量を表10にまとめる。両地図を比較すると、「名古屋市西区上名古屋4丁目」の方

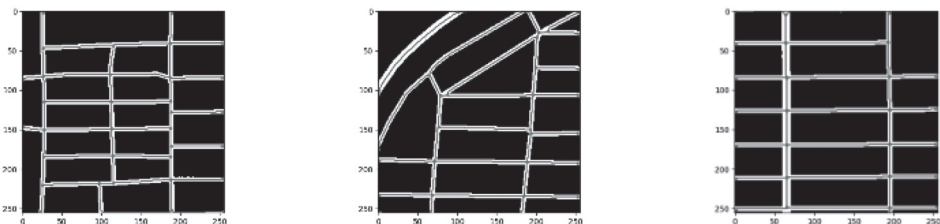


図2 道路ネットワークの抽出

表9 道路ネットワークの特徴量

特徴量	意味	値
ノード数 N	ネットワークを構成するノード数	$N=33$
エッジ数 E	ネットワークを構成するエッジ数	$E=43$
密度 D	可能なエッジの組み合わせに対する実際のエッジの比率	$D \approx 0.08$
平均次数 k	ノードに接続されるエッジ数の平均	$k \approx 2.61$

表10 面積比に基づく特徴量

住所	緯度	経度	道路面積比	建物面積比	駅	駐車場
名古屋市西区上名古屋4丁目	35.196853	136.899306	0.241302	0.280624	0	1
小牧市中央1丁目	35.289302	136.927114	0.135635	0.244675	1	0

が道路と建物の面積が広いことがわかる。また、「小牧市中央1丁目」には駅が存在し、「名古屋市西区上名古屋4丁目」には駐車場が存在することがわかる。

5 評価実験

上述した12種類の特徴量をランダム・フォレストで学習し、3種類の犯罪種別（自転車盗、車上ねらい、自動車盗）の分類精度を評価する。あ

らためて12種類の特徴量を表11にまとめる。このうち、「X」「Y」「ノード数」「エッジ数」「密度」「平均次数」は、前処理として平均0、分散1に標準化した。データセットは $N=3000$ のサンプルで構成され、 $K=20$ のk-分割交差検証で評価した。ランダム・フォレストのパラメータはハイパーパラメータの最適化フレームワークであるOptuna⁸⁾を利用して最適化した。最適化したパラメータは「最大の深さ」「特徴量の数」「中間ノードの最小分割数」「葉ノードの最小分割数」の4種類であり、表12の結果となった。

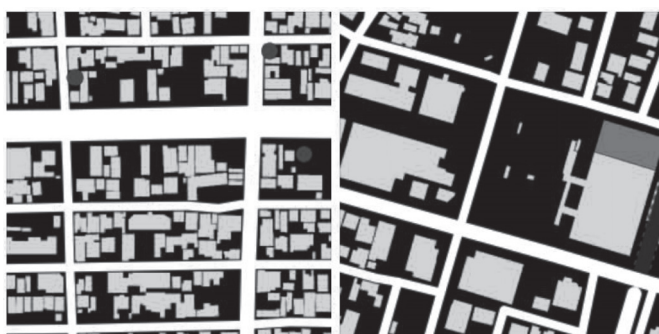


図3 面積比に基づく特徴量のサンプル（左が名古屋市西区上名古屋4丁目、右が小牧市中央1丁目）

(© Mapbox, © OpenStreetMap)

表11 学習に用いる特徴量

特徴量	意味
発生時	発生時刻
X	平面直角座標のX
Y	平面直角座標のY
ノード数	道路ネットワークのノード数
エッジ数	道路ネットワークのエッジ数
密度	全ての組み合わせに対する実際のエッジの比率
平均次数	ノードに接続されるエッジ数の平均
発生場所	発生場所（ダミー変数）
道路面積比	道路が占める面積の割合
建物面積比	建物が占める面積の割合
駅	駅の有無
駐車場	駐車場の有無

表12 ランダム・フォレストのパラメータ

特徴量	値
最大の深さ	6
特徴量の数	11
中間ノードの最小分割数	7
葉ノードの最小分割数	4

ランダム・フォレストを構成する決定木の数
を1から30まで変化させたときの、正解率の推移を
図4に示す。また、各特徴量の不純度に基づく重
要度は表13となった。正解率が最大となるのは、
決定木の数が増えると正解率も増加するが、その数が
15に到達する頃に収束し、その後は大幅な改善
は確認できない。3種類の分類であることから、
学習に用いた12種類の特徴量は約20%程度の正
解率の向上に寄与したと考えることができる。最
も重要度が高かった特徴量は「発生時」である。
「発生時」を軸とした犯罪数のヒストグラムを図
5に示す。「自転車盗」は午前中、「車上ねらい」
と「自動車盗」は深夜に発生しやすいことが確認
できる。次に、「道路上」と「一戸建住宅」の発
生場所の重要性が高い。さらに、平面直交座標の
「X」「Y」、また、「建物面積比」「道路面積比」が
続き、地理的な位置関係が分類に寄与しているこ
とが確認できる。特に建物面積比が高い都市部で、
「自転車盗」が発生しやすい傾向があるようだ。
一方で、「駐車場」や「駅」などの環境要因は、
分類に殆ど寄与していなかった。これは、生成し
た地図の範囲が狭いことに加え、「駐車場」の情

報が不十分であるためと考えられる。正解率は5
割を少し超える程度であるため、実用化に向ける
信頼性は不十分であるが、地理的な特徴量が分
類に寄与していることは確認できた。今後は、新
たな特徴量の追加や、分類アルゴリズムの改良に
よって、精度の向上を狙うと共に、防犯を目的と
したシステムの実用化を目指したい。

6 まとめ

本研究では、地域の防犯を目的として、過去の
犯罪の時刻や住所などの特徴量と、犯罪発生地域
の道路や建物などの地理的特徴量を考慮した、機
械学習による犯罪種別の分類を試みた。時刻や住
所などの基本的な特徴量は愛知県警察が公開して
いる犯罪オープンデータから抽出し、地理的特徴
量は地図画像に画像処理を適用することで抽出し
た。独自にデータセットを構築し、ランダム・フォ
レストによる学習の結果、「自転車盗」「車上ねら
い」「自動車盗」の3種類の犯罪種別に対して、最
大で52.74%の精度で正解を導出することができ
た。正解率の向上には、地図画像から生成した建

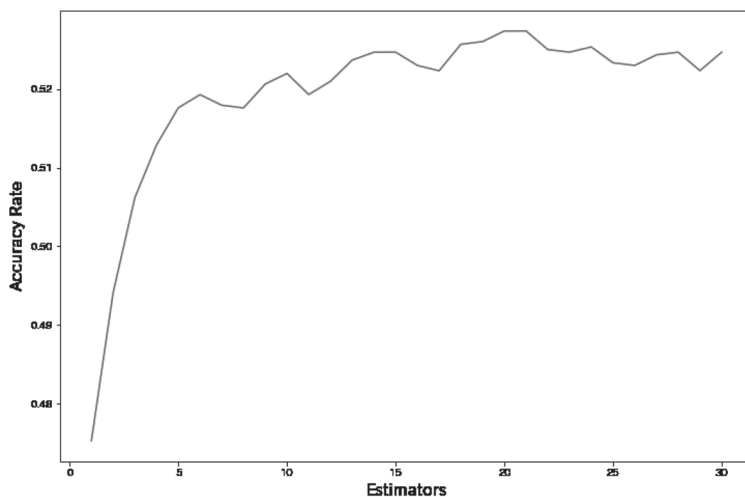


図4 ランダム・フォレストの正解率

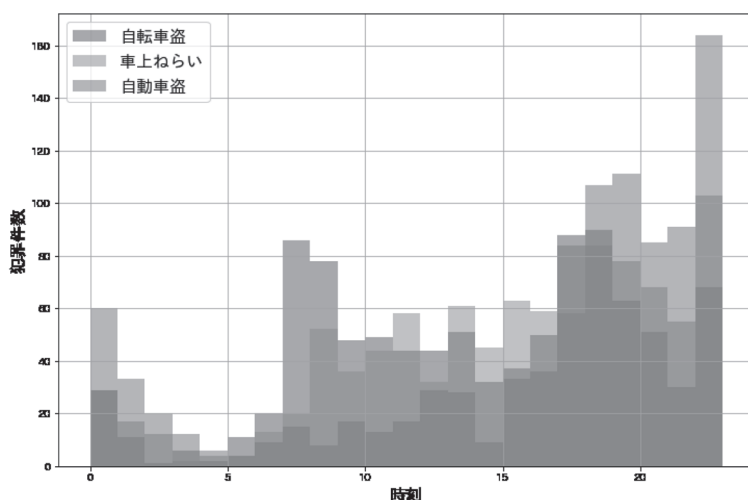


図5 発生時を軸とした犯罪数のヒストグラム

表 13 特徴量の重要度

順位	特徴量	重要性
1	発生時	0.226243
2	道路上（発生場所）	0.120355
3	一戸建住宅（発生場所）	0.104505
4	Y	0.100482
5	建物面積比	0.087125
6	X	0.085422
7	道路面積比	0.073741
8	その他（発生場所）	0.055741
9	平均次数	0.042924
10	エッジ数	0.033852
11	ノード数	0.022188
12	密度	0.021862
13	駐車場（発生場所）	0.012209
14	4階以上共同住宅（発生場所）	0.008736
15	駐車場	0.001968
16	駅	0.001469
17	3階以下共同住宅（発生場所）	0.001177

物面積比などの特徴量が強く寄与していることがわかった。今後は、駅や駐車場だけでなく、金融機関やコンビニエンスストアなど他の環境要因を

考慮したいと考えている。このためには、生成する地図の適切なサイズを検討すると共に、日本二輪車普及安全協会⁹⁾などが提供している精度の高

い駐車場の情報を取り込む必要があると考えている。また、研究成果をパトロールの効率向上などに応用し、地域社会の安全に貢献したいと考えている。

むかい・なおと / 文化情報学部准教授
E-mail : nmukai@sugiyama-u.ac.jp

注

- 1) 愛知県警察犯罪オープンデータサイト (<https://www.pref.aichi.jp/police/anzen/toukei/opendata/seians/crimeopendata.html>)
- 2) NIJ (<https://nij.ojp.gov/>)
- 3) Mapbox (<https://www.mapbox.jp/>)
- 4) 「ひったくり」のみは、追加で被害者の情報が含まれる
- 5) Mapbox Geocoding API (<https://docs.mapbox.com/api/search/geocoding/>)
- 6) Mapbox Static Tiles API (<https://docs.mapbox.com/api/maps/static-tiles/>)
- 7) 無向グラフであることに注意されたい
- 8) <https://www.preferred.jp/ja/projects/optuna/>
- 9) 日本二輪車普及安全協会 (<https://www.jmpsa.or.jp/>)

参考文献

- [1] 工藤郁子. 行政におけるaiの活用と信頼. 人工知能学会全国大会論文集, Vol. JSAI2021, No. 0, pp. 2C4OS9b01-2C4OS9b01, 2021.
- [2] 大山智也、雨宮護、島田貴仁、中谷友樹. 地理的犯罪予測研究の潮流. GIS-理論と応用, Vol. 25, No. 1, pp. 33-43, 2017.
- [3] 中川淳子、西村祥治、宮野博義. 犯罪発生履歴データの機械学習による時空間カーネル密度推定型犯罪予測の最適化. 情報処理学会論文誌, Vol. 61, No. 7, pp. 1255-1266, jul 2020.
- [4] 野貴泰、糸井川栄一. 犯罪多発地点の予測に基づく防犯パトロール経路に関する提案. 地域安全学会論文集, Vol. 31, No. 0, pp. 195-204, 2017.
- [5] 山下雄大、森純一郎. 市街地画像を用いた犯罪地域の特徴抽出. 人工知能学会全国大会論文集, Vol. JSAI2017, No. 0, pp. 2M12-2M12, 2017.
- [6] 張曉齊、米澤剛、吉田大介. オープンデータと1stmを用いた犯罪発生の予測及び時間的近接性における考察. 情報学, Vol. 16, No. 1, pp. 56-70, jun 2019.
- [7] T. Y. Zhang and C. Y. Suen. A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, Vol. 27, No. 3, 1984.
- [8] Geoff Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers Environment and Urban Systems*, Vol. 65, pp. 126-39, 07 2017.

