衛星画像と機械学習を用いた土地利用分類に基づく開発適地の自動選定手法の構築

久保田涼太*・寺田遣都・秋山祐樹*****

Development of an Automated Method for Selecting Suitable Development Sites through Land Use Classification with Satellite Imagery and Machine Learning

Ryota Kubota*, Kento Terada**, Yuki Akiyama***

Abstract: In recent years, the tendency toward long working hours has been cited as a problem in Japan's construction industry, necessitating the promotion of Digital Transformation（DX）to streamline operations. However, the selection of suitable development sites is still carried out through human visual inspection, requiring significant costs and time. Meanwhile, the development of detailed land use mesh data—which divides all of Japan into 100m mesh units and classifies them by land use purpose—faces similar challenges. This study predicted and evaluated land use classifications in the Tohoku region using satellite imagery and machine learning. It then combined this with an evaluation metric called development potential, created based on factory site selection criteria, to extract suitable development sites. As a result, high prediction accuracy was achieved in two prefectures of Tohoku, successfully identifying suitable development sites. We aim to further advance this technology to achieve full-scale implementation in practical operations and significantly reduce the time required for this task.

Keywords: 土地利用分類（land use classification），衛星画像（satellite imagery），機械学習（machine learning），開発適地（suitable development site），画像分類（image classification）

１．はじめに

１.１　研究背景

近年，日本の建設業界における問題の 1 つに，「長時間労働」が挙げられる．業界全体ではデジタル・トランスフォーメーション（DX）を推進して業務を効率化させ，労働時間を削減していく努力が続けられている．

しかし，依然として DX 化が進んでいない業務領域も数多く存在する．その一つとして，開発可能な土地（以下，「開発適地」）の選定業務が挙げられる．開発適地とは，まとまった広さの田畑・森林など，将来的に大規模施設の建設に適した土地を指す．同業務は，現状ではオペレーターによる航空写真の目視判読に依存しており，膨大な時間と労力を要することが課題となっている．開発適地の選定は，都市計画や防災計画の基盤ともなっている土地利用細分メッシュデータ（以下，「土地利用データ」）が重要な判断材料となっている．土地利用データは，全国を 100m メッシュに区分し，各メッシュを土地利用目的別に分類した大規模な地理空間データである．

現在，土地利用データの整備は，開発適地の選定と同様に，オペレーターが衛星画像や航空写真を目視で判読し，手作業で分類するという労働集約的な手法に依存している．その結果，日本全国の約 3,400 万メッシュの整備には多大な時間とコストを要する上，オペレーターの判断に依存するため品質の不均一性も生じる．また，データの更新頻度が限られており，都市開発や災害等による土地利用の最新の変化に迅速に対応できず，データの鮮度が損なわれるという課題も生じている．

こうした背景から，AI 等の先進技術を活用した効率的な土地利用データの整備手法を確立し，開発適地を自動的に選定する技術が強く求められている．

１．２　既往研究と本研究の位置づけ

AI を用いて画像から土地被覆や地物を自動分類する手法は，近年数多く報告されている．越智（2009）は，画像オブジェクトに基づく高分解能衛星画像での土地被覆分類手法の提案を行っている．また，有水・毛綱（2020）は，深層畳み込みニューラルネットワークを用いたセグメン

*　学生会員　東京都市大学建築都市デザイン学部（Tokyo City University）
〒158-0087　東京都世田谷区玉堤 1-28-1　E-mail：g22320333@tcu.ac.jp
**　学生会員　東京都市大学大学院総合理工学研究科（Tokyo City University）
〒158-0087　東京都世田谷区玉堤 1-28-1　E-mail：g2581638@tcu.ac.jp
***　正会員　東京都市大学建築都市デザイン学部（Tokyo City University）
〒158-0087　東京都世田谷区玉堤 1-28-1　E-mail：akiyamay@tcu.ac.jp

テーション手法によって，森林作業道の自動走行における走行可能領域を判別する手法を開発し，車道クラスの高精度な分類方法の提案を行っている．

しかし，本研究が目的とする日本全国という規模を対象に，100m メッシュ単位で土地利用を自動的に分類するとともに，その結果を開発適地の選定作業の自動化に活用しようとする手法は未だに開発されていない．

１．３　本研究の目的

本研究の目的は以下の 2 点である．まず 1 点目として，衛星画像と機械学習を用いて，日本全国の土地利用を 100m メッシュ単位で自動分類し，土地利用データを高頻度かつ効率的に整備する技術を開発することである．具体的には，建物用地，その他可住地，水域，森林・荒地の 4 クラスに分類する．また，2 点目として，同技術を活用して得られた土地利用データを用いて，開発適地の候補となる領域を自動選定する手法を提案することである．

２．研究手法

２.１　対象地域

本研究では，日本全国への展開を見据えたパイロットスタディとして，都市部と農村部が存在し，土地利用分類の精度評価に適している東北地方全域を対象とした．

２.２　使用データ

モデル構築の学習データおよび精度検証の評価データには，国土数値情報が提供する 2020 年度の土地利用データを過去の土地利用判読結果として使用した．また，本研究では，衛星画像として欧州宇宙機関（ESA）が運用しており，無償での利用が可能な Sentinel-2 を採用した．Sentinel-2 は最大 10m の高い空間分解能，5 日間隔の高頻度な観測，13 種類のスペクトルバンドを特徴とする．

本研究で使用する画像については，土地利用データの判読時期である 2020 年 10 月から 12 月を対象としたモザイク画像を取得した．

２.３　特徴量の抽出

取得した衛星画像に対して，土地利用分類のための特徴量の抽出を実施した．分類モデルの入力特徴量として，土地被覆の識別に有効とされる複数のスペクトルバンドおよび指標を選定した．

具体的には，可視光で光の三原色である Blue（Band2），Green（Band3），Red（Band4）の 3 種類と，近赤外線に対応する NIR（Band8），短波長赤外線に対応する SWIR（Band11）の 2 種類，合計 5 種類のバンドを使用した．加えて，これらのバンドから式(1)～(3)で定義される 3 種類の指標（Hall, D. K. et al.，1995），（Tucker, C. J.，1979），（Xu, H.，2006）を作成し，特徴量として追加した．表 1 に各特徴量の解像度と概要を示す．

２.４　教師データの作成

機械学習モデルに入力可能な形式とするため，100m メッシュごとの特徴量データと，正解ラベルである土地利用判読結果を結合し，教師データセットを作成した．

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \quad (1)$$

$$MNDWI = \frac{(Green - SWIR)}{(Green + SWIR)} \quad (2)$$

$$NDSI = \frac{(SWIR - NIR)}{(SWIR + NIR)} \quad (3)$$

分類ラベルは，国土数値情報の土地利用データをもとに表 2 に示す 4 クラスとした．まず第 1 分類として，土地を可住地と非可住地に大別する．次に第 2 分類として，可住地を建物用地とその他可住地の 2 クラスに，非可住地を水域と森林・荒地のクラスに分類する．その結果，最終的に合計 4 クラスに分類される．

２.５　機械学習モデルの構築

2.4 で作成した教師データを用い，東北地方全域の土地利用分類を自動化することを目的に，Random Forest に基づく機械学習モデルを構築した．なお，過学習を抑制して汎化性能を確保するため，5 分割交差検証を実施した．分類処理には，2.4 で述べた段階的な二値分類手法を採用した．なお，機械学習モデルは県単位で構築した．

２.６　予測と評価

2.5 で構築した機械学習モデルを用いて東北地方各県を対象として 100m メッシュ単位で土地利用分類の予測を実施した．得られた予測結果は，正解データである土地利用判読結果と比較することで，モデルの分類精度を定量的に評価する．なお，モデルの精度を評価するために採用する指標は表 3 の通りである．

表 1　本研究で用いた特徴量

特徴量	解像度(m)	概要
Blue (Band2)	10	青色可視光スペクトルバンド(490nm)
Green (Band3)	10	緑色可視光スペクトルバンド(560nm)
Red (Band4)	10	赤色可視光スペクトルバンド(665nm)
NIR (Band8)	10	近赤外スペクトルバンド(842nm)
SWIR (Band11)	20	短波赤外スペクトルバンド(1610nm)
NDVI	10	緑植生の茂り具合を示す指標.
MNDWI	20	地表面における雪を含めた水域や，植生に含まる水分量の存在を示す指標.
NDSI	20	地表面における砂やコンクリートなどの分布を示す指標.

表 2　土地利用分類項目の対応表

第一分類	第二分類	土地利用データの分類項目
可住地	建物用地	-
	その他可住地	田・その他の農用地・道路・鉄道・その他の用地・ゴルフ場
非可住地	水域	河川地および湖沼・海浜・海水域
	森林・荒地	-

また，本研究でモデルの精度を多角的に評価するために採用する指標について表 3 に示す．

2.7　開発ポテンシャルの評価

2.6 で予測した 4 分類の土地利用データを用いて，開発適地の選定を実施した．具体的には，主に田畑を中心とした農用地を占める，その他可住地と森林・荒地に分類されたメッシュを対象に，開発適地の絞り込みを定量的に実施するため，新たな評価指標として，開発ポテンシャルを定義し評価を行った．

開発ポテンシャルとは，経済産業省の「工場適地の選定基準について」で定められている基準に基づいて，開発適地としての適性を定量的に評価する指標である．具体的な算出項目と配点を表 4 に，算出に使用したデータを表 5 に示す．評価項目の多くは産業施設やインフラ施設までの距離となっており，距離が近いほど高い得点が与えられる．

表 3　モデルの精度評価のための指標

指標	概要	評価の観点
正解率 (Accuracy)	全てのデータのうち，正しく分類されたものの割合.	モデル全体の正しさ
適合率 (Precision)	あるクラスだと予測されたデータの中で，実際に正解だったデータの割合.	予測の的確さ
再現率 (Recall)	実際に特定のクラスに属するデータ内で，正しく予測されたデータの割合.	見逃しの少なさ
F値 (F-measure)	適合率と再現率の調和平均であり，両者のバランスを考慮した指標.	適合率と再現率の調和

表 4　開発ポテンシャルの算出項目と配点

実装された調査項目	得点
国道までの距離	0～5
高速道路のI.Cまでの距離	0～7
新幹線到着駅までの距離	0～5
空港までの距離	0～5
市町村の人口規模	0～5
人口5万人都市までの距離	0～3
人口15万人都市までの距離	0～5
工科系大学までの距離	0～3
公設試験場までの距離	0～3

表 5　開発ポテンシャルの算出に使用したデータ

データ元	詳細情報
国土数値情報	全国道路データ(1995年)
	全国高速自動車道データ(2020年)
	全国駅データ(2020年)
	全国空港データ(2020年)
	全国行政区域データ(2020年)
	全国学校データ(2023年)
国勢調査(e-Stat)	人口等基本集計(2020年)
経済省公設試験場	工業試験場(2023年)

本研究では，合計得点（41 点）のうち，6 割（25 点）を満たすメッシュを抽出することで，開発適地の候補となる領域の絞り込みと評価を行った．

3．結果

3.1　土地利用分類の予測結果

本稿では，宮城県および秋田県の結果を例として示す．2 県の土地利用分類の予測結果を図 1 と図 2 に，各クラスに対する評価指標と推定精度を表 6 と表 7 に示す．宮城県と秋田県のいずれにおいても，高い精度で土地利用分類予測が行われたことが確認された．なお，いずれの県においても，モデル構築から予測終了までに要した時間は約 2 時間であり，人間による目視判読に比べて大幅に短時間で処理が可能であり，作業負担の大幅な軽減を実現することができた．

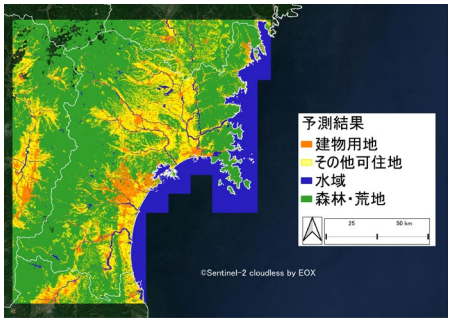


図 1　宮城県における予測結果

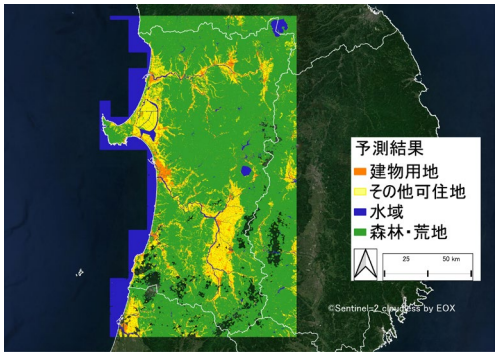


図 2　秋田県における予測結果

表 6　宮城県における各クラスに対する評価指標と推定精度

宮城県	正解率(%)	適合率(%)	再現率(%)	F値(%)
建物用地	—	92.71	99.99	96.22
その他可住地	—	98.51	98.09	98.3
水域	—	99.43	99.3	99.37
森林・荒地	—	99.99	99.31	99.65
全体	99.08	97.66	99.17	98.38

表 7　秋田県における各クラスに対する評価指標と推定精度

秋田県	正解率(%)	適合率(%)	再現率(%)	F値(%)
建物用地	—	77.17	99.99	87.11
その他可住地	—	92.84	97.03	94.89
水域	—	98.58	98.22	98.4
森林・荒地	—	99.99	98.12	99.05
全体	98.02	92.15	98.34	94.86

３.２　開発適地の抽出

2.7 で定義した開発適地を抽出した結果を図 3，図 4 に示す．開発ポテンシャルの得点が 25 点未満の領域を半透明で，25 点以上の開発適地は得点が大きくなるにつれて濃い色で示している．2 県ともに，主に大都市の周辺部で開発適地が連続的に広がっていることが確認された．なお，開発ポテンシャルの付与に要した時間は 1 県につき約 2 時間であった．

４．考察

表 7 によると，秋田県の建物用地における適合率が 80%に満たず，他の結果と比較して低いことがわかる．一方，表 6 より宮城県における建物用地では，90%を超える適合率となっていた．総務省統計局の「住宅及び世帯に関する基本集計」（2023）によると，秋田県は 2023 年時点での総住宅数が約 44 万戸と全国的にも少ないことがわかる．2018 年と比較すると 1.1%減少しており，この減少率は全国で最も高い値である．また，居住世帯のある住宅に関するデータでは，一戸建の割合が 79.4%を占めており，これも全国で最も高い値である．さらに，政府統計が公開している「都道府県・市区町村のすがた（社会・人口統計体系）」の 2020 年のデータによると，秋田県には秋田市を除いて，人口 10 万人規模の市町村がないこともわかる．以上より，秋田県では衛星画像内の建物データが少ないことや，建物用地と判断できる領域が小さいことが，適合率の低さに影響していると考えられる．

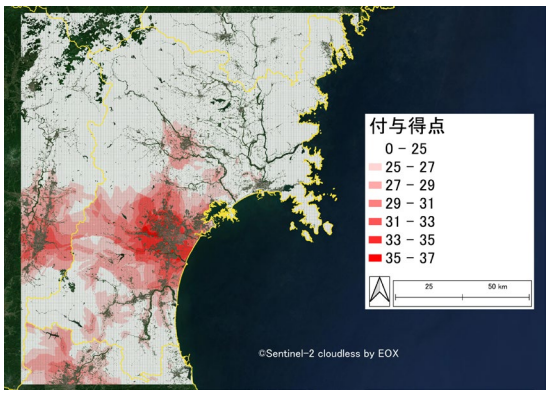


図 3　宮城県における開発適地

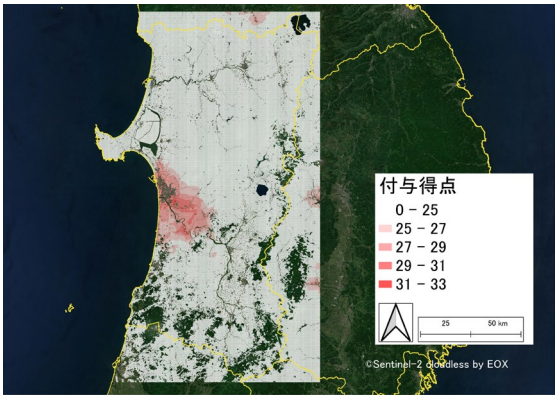


図 4　秋田県における開発適地

５．おわりに

本研究では，衛星画像に機械学習を適用して土地利用分類を予測し，その結果を開発ポテンシャルの評価値と重ね合わせることで，開発適地を自動的に選定する技術を構築した．その成果として，東北地方全域を対象とした土地利用分類の自動化を実現し，宮城県および秋田県における精度検証により高い分類性能を確認した．さらに，開発ポテンシャルと土地利用分類を組み合わせることで，開発適地の自動選定が可能であることを示した．また，従来の目視による選定と比較して，業務時間やコストを大幅に削減できることも明らかとなった．

今後は，日本全国を対象とした開発適地の自動選定を進めることを目標とする．また，本研究では土地利用データの判読時期に合わせて 2020 年の衛星画像を使用したが，今後はより新しい時点の画像を活用し，既存モデルによる外挿的な適用を検討する予定である．

さらに，開発ポテンシャルの評価項目は現状，産業施設やインフラ施設までの距離に依存している．実務的な適用を見据えると，土地の形状や工事および運営に要するコストなど，新たな実務的観点を取り入れた指標の拡充が不可欠である．これらの展望を踏まえ，本技術を実務における本格的な運用へと発展させることを目指す．

謝辞

本研究は，JSPS 科研費（24K00243）の助成を受けて実施した．ここに記して謝意を表す．

参考文献

越智士郎（2009）画像オブジェクトに基づく高分解能衛星画像での土地被覆分類手法の検討.「東南アジア研究」, **46**（4）, 578-592.

有水賢吾・毛綱昌弘（2020）深層畳み込みニューラルネットワークを用いたセマンティックセグメンテーションによる森林作業道抽出.「森林利用学会誌」, **35**（1）, 7-13.

Hall, D. K., Riggs, G. A., Salomonson, V. V.（1995）Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer data. *Remote Sensing of Environment*, **54**（2）, 127–140.

Tucker, C. J.（1979）Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, **8**（2）, 127-150.

Xu, H.（2006）Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, **27**（14）, 3025–3033.

経済産業省（2017）工場適地の選定基準について. <https://www.meti.go.jp/shingikai/sankoshin/chiiki_keizai/kojo_ritchi/pdf/035_s03_02.pdf>.

総務省統計局（2024）住宅及び世帯に関する基本集計. <https://www.stat.go.jp/data/jyutaku/2023/pdf/kihon_gaiyou.pdf>.