AIを活用した地物自動抽出に関する研究（第4年次）

実施期間　　平成30年度～令和4年度
地理地殻活動研究センター
地理情報解析研究室　大野　裕幸　浦　郁子

1. はじめに

本研究は，労働集約的な業務構造となっている地図作成に，近年画像認識の分野にブレークスルーをもたらした畳み込みニューラルネットワーク（CNN）と深層学習の技術（以下総称して「AI」という.）を導入することで，長期的には地図作成の自動化を，短期的には地図作成における「判読」及び「図化」の工程の生産性向上を図ることを目的としている．これまでに抽出性能検証用のデータセットを作成（大野ほか，2019a）するとともに，使用するCNNとハイパーパラメータを決定（大野ほか，2020）し，40 項目以上の地物を対象として学習データの構築を進めてきた（大野ほか，2021）．本年度は，学習データの構築を進め，抽出性能の数値目標をクリアする項目数の増加を図るとともに，項目毎に本研究の手法の適合可否の判定，抽出性能検証用データセットの拡充，CNNの改良を実施した．

2. 研究内容及び得られた成果

2.1 研究開発の概要

研究内容は，①深層学習に用いる教師データ（学習データと検証用データ）の作成，②深層学習で用いるハイパーパラメータの探索及びCNNの構成の検討と，③AIで自動抽出した地物を自動的に地図データベースに反映させるアルゴリズムの検討，に大別される．本年度は，①を中心としつつ，①の性能向上に寄与する点に関して②を実施した．研究全体の方針として，学習データは2007年度から2017年度までに撮影された空中写真に限定し，検証用データは2007年度から2019年度（一部2020年度と 2021 年度撮影地区を含む）までに撮影された全ての地区から選択ルールを決めて機械的に選定した空中写真画像を用いた．

2.2 学習データの作成

学習データは，深層学習において欠かせない基礎データであり，かつその品質が抽出性能の良否に直結する．ただし，単に枚数を多く作成しても抽出性能向上への寄与は大きくなく（大野ほか，2020），作成する箇所を適切に選定した良質なものを構築する必要がある．学習データは572×572ピクセルのサイズで作成したものを286×286ピクセルの4枚に分割し，さらに572×572の範囲全体を286×286サイズに縮小した1枚を追加した5枚の学習データとする仕様で構築を進めた．縮小した1枚は，推論時に入力する空中写真のスケーリングにある程度対応させることを意図して追加したものである．学習データは，入力用の画像と，対象地物を指定色で塗分けたラベリング画像のセットで構成される．

令和4年3月31日時点の項目毎の作成枚数及び抽出性能評価の最高値（F-Score：F値）を表-1に示す．表-1に記載した枚数は，286×286ピクセルサイズの学習データとしての数であり，入力用画像とラベリング画像のセットで1枚とカウントしている．また，最高値は，必ずしも枚数欄に記載の枚数で深層学習したモデルで算出された値とは限らない．F値の算出には，基本的に2.3の検証用データを用いたが，一部の項目は当該項目専用のデータをさらに追加した検証用データを用いて算出した値となっている．

表-1　主要な項目別学習データの整備状況と抽出性能評価値（F 値）

項目名	枚数	F 値	項目名	枚数	F 値	項目名	枚数	F 値
02 舗装道路	9600	0.812	19 竹林	4100	0.448	43 不透過水制	300	-
04 駐車場	2800	0.605	20 針葉樹林	2000	0.657	44 防波堤	4800	0.725
05 中央分離帯	3200	0.336	21 茶畑	1300	-	46 ダム	200	-
06 軌道の中心線	4000	0.555	22 果樹園	1200	0.175	54 芝地	900	-
07 水部	5500	0.880	23 荒地	1000	0.174	58 砂礫地	2500	0.324
08 堅牢建物	2300	0.340	26 しの地	1000	-	61 たたき	100	-
09 普通建物	5800	0.807	27 太陽光発電設備	3200	0.844	62 岩崖・岩	1000	0.064
10 無壁建物	1200	0.624	28 輸送管	600	-	63 高塔	5500	0.607
11 被覆	1400	0.053	34 石段	450	-	65 湿地	1200	-
12 土崖	1500	0.023	36 雪覆い等	800	-	66 材料置き場	300	-
13 透過水制	3400	0.804	38 プラットホーム	3600	0.512	67 園庭	300	0.147
14 横断歩道橋	1100	-	39 タンク	2200	0.873	68 墓地	4100	0.802
16　畑	6500	0.445	40 広葉樹林	1300	0.389	※F 値欄，下線は当該項目専用の検証データも加えた評価値		
17・18 水田	4900	0.677	41 風車	2200	0.384			

本研究では，各項目の F 値の数値目標を 0.800 以上と定めている．本年度中に新たにこの数値目標を達成した項目は，13 透過水制，27 太陽光発電設備，39 タンク，68 墓地の 4 項目である．この他に，変化情報抽出に特化した項目として 48 変化情報用建物が目標を達成している（3200 枚時点で F=0.802）．

学習データ構築の際は，区切りの良い枚数に到達した時点で深層学習を行い，F 値を算出して性能向上の傾向を分析し，学習データに用いる画像の選択方針を決定する．この時の F 値が意図せず悪化（F 値は，Precision と Recall の調和平均で算出される．Precision と Recall は通常トレードオフの関係にあり，枚数を増やしたからと言って F 値が必ず向上するものではないが，Precision 又は Recall のどちらかを向上させる方針で構築に取り組む．しかし，その方針に反して Precision，Recall ともに低下する場合がある.）した場合は追加した学習データの再検討を行って一部のデータを差し替えたり，追加分を新規に再構築したりすることがある．その傾向から，11 被覆，12 土崖，62 岩崖・岩については，比高が関連する地物であることから本研究の手法によって抽出することは困難と判断し，本年度をもって対象から除外した．また，34 石段も比高が関連するうえ，地上画素寸法 20cm の画像では階段部分がぼやけて階段としての特徴量を呈さないケースが多く見られたことから，対象から除外した．さらに，25 はい松地は，深層学習に必要な数を充足するほど空中写真の撮影事例が無かったことから，そもそも深層学習に適さない地物と判断して，対象から除外した．

また，この傾向分析の結果から，フィルム式カメラの画像を含めると性能が悪化すること，地上画素寸法 40cm の画像を含めると，例えば 02 舗装道路と 14 横断歩道橋の特徴量の区別ができなくなる，といった致命的な傾向等が認められたことから，過去に作成した学習データから両者を除外して地上画素寸法 20cm のデジタル航空カメラの撮影画像に入れ替えるという作業も実施した．

2.3 検証用データの作成

検証用データは，本研究初年度に 2007 年度から 2017 年度にデジタル航空カメラで撮影された全ての空中写真撮影地区からルールを決めて 1 枚ずつ空中写真画像を選定し，その中央部 572×572 ピクセ

ルの範囲を 36 項目の指定色で塗分けたラベリング画像として作成していた（大野ほか，2019b）．その後，2018 年度以降に実施された空中写真撮影地区が増加してきたことから，2018 年度から 2021 年度に撮影された全ての地区も検証用データに含めることとし，検証用データの追加を行った．また，それに合わせて，従来は 572×572 ピクセルであった検証用データのサイズを，1000×1000 ピクセルの範囲に増加させることとしたため，全データを判読し直して検証用データの再構築を実施している．これによって，従来よりも検証用データ 1 枚あたりのピクセル数が約 3 倍に増加することになり，検証用データセットのピクセル数不足の問題をある程度解消することができると考えている．

また，自然植生は，従来は針葉樹林や広葉樹林といった区別なく，15 自然植生 1 項目として扱っていたが，検証用データの再構築に合わせて，19 竹林，20 針葉樹林，26 しの地，40 広葉樹林の 4 項目に分類し直している．表-1 で 19 竹林や 20 針葉樹林等の F 値が記載されているのはこの効果による．

一方，学習データ構築時の分析の結果，地上画素寸法 40cm で撮影された画像が混在することに関する課題が生じたことから，地上画素寸法 20cm で撮影された画像のみに対象を絞って，本研究の対象から地上画素寸法 40cm の画像を除外することとし，検証用データからもそれらの画像を除外した．図-1 に新たな検証用データの例を示す．再構築における本年度末時点での進捗率は約 35%である．

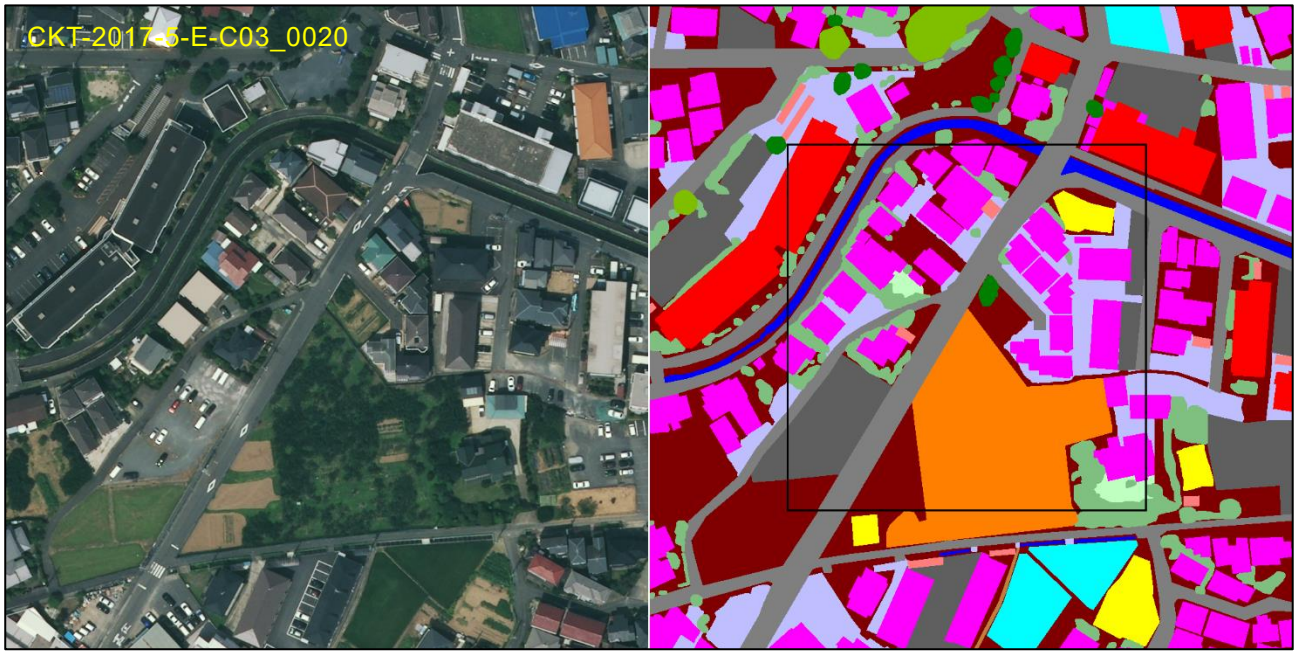

図-1　新しい検証用データ．左が入力画像，右がラベリング画像．黒枠内が従来の検証用データの範囲．

2.4 ハイパーパラメータの探索と CNN の構成の検討

深層学習において，あらかじめ決定することができるパラメータのことをハイパーパラメータという．本研究では，項目毎に指定色を設定し，その色で対象地物を塗分けたラベリング画像（単項目）を作成してきた．その指定色の色合いが抽出性能に与える影響を検証してみたところ，8 ビットの RGB 値が取り得る 0～255 の範囲のうち，中間値を指定色とする場合と 0 又は 255 を指定色とする場合で有意な性能差が確認された（表-2）．これは，本研究では損失関数にこのようなタスクで一般的に用いられるクロスエントロピーではなく L1 距離を用いているためと思われる．L1 距離では，正解値に対してオーバーとアンダーの両側で差が生じ得る中間値より，そのどちらか片方にしか差が生じない値の方が効率よく収束したものと考えられたため，指定色を RGB 値 0 と 255 を組み合わせた 8 色のい

ずれかの色に変更して，学習データを再構築することとした．

表-2　ラベリング色の違いによる抽出性能の相違.

項目	学習枚数（枚）	ラベリング色	Precision	Recall	F 値
10 無壁建物	1,100	#FF7F7F	0.793	0.194	0.311
	1,100	#FFFF00	0.755	0.531	0.624
44 防波堤	3,300	#FFBFBF	0.596	0.366	0.454
	3,300	#FFFF00	0.670	0.714	0.691

また，深層学習に用いる CNN は，第 2 年次以降 16 層の Encoder-Decoder の構造を持った U-Net 系のネットワークを使用しているが，例えば PSP-Net（Zhao et al., 2017）といった，より性能の高い CNN では，異なるサイズでより広い範囲の特徴量を捉えた層を挿入することで性能向上を図っている．これは，判読を行う際にいったん画像の表示倍率を小さくし，広い範囲を見てみると地物を把握しやすいこととも符合するため, CNN を U-Net の構造を保持しつつ入力画像サイズを 512×512 ピクセルに拡大した 18 層に拡張し，抽出性能を比較してみた．入力サイズが拡大したことに伴い，学習データ枚数が減少したにもかかわらず，ヘリコプターが撮影した映像から濁水の範囲を抽出する学習済モデルのケース（この地物は表-1 には含まれない）で，表-3 のように F 値が約 3.9%向上した．深層学習のステップ数や目的関数等のハイパーパラメータは同じである．このことから，F 値が頭打ちの傾向を見せている 16 畑や 17 水田等の地物にこの新しい CNN を適用することを検討している．

表-3　入力画像サイズの異なる CNN の抽出性能の比較.

入力サイズ	学習枚数（枚）	Precision	Recall	F 値
256×256	57,989	0.948	0.621	0.750
512×512	14,592	0.947	0.662	0.779

4. まとめ

本年度は，本研究が目標とする地物抽出性能 F 値≧0.8 が非常に高いことを認識させられた 1 年であった．今後は，表-1 において F 値≧0.6 に到達している項目を中心に，新しい CNN の効果を見極めながら, 数値目標の達成に向けてさらなる学習データの構築とハイパーパラメータの探索を継続する．

参考文献

大野裕幸，中埜貴元，遠藤涼（2019a）: AI を活用した地物自動抽出に関する研究（第 1 年次），国土地理院平成 30 年度調査研究年報，180-183.

大野裕幸，遠藤涼，中埜貴元，篠田昌子（2019b）: 空中写真を用いた地物抽出性能評価用データセットの構築，2019 年度人工知能学会全国大会（第 33 回）講演論文集，3Q3-J-13-01.

大野裕幸，白石喬久，遠藤涼，中埜貴元，岸本紀子（2020）: AI を活用した地物自動抽出に関する研究（第 2 年次），国土地理院令和元年度調査研究年報，166-169.

大野裕幸，遠藤涼，白石喬久，浦郁子（2021）: AI を活用した地物自動抽出に関する研究（第 3 年次），国土地理院令和 2 年度調査研究年報，180-183.

ZHAO, H., SHI, J., QI, X., WANG, X., and JIA, J. (2017): Pyramid Scene Parsing Network, CVPR2017.