

# ディープラーニングを適用した衛星画像からの土地被覆分類手法の評価

## Assessing the Land Cover Classification Technique of the Satellite Imagery using Deep Learning

伊東 里保<sup>\*1</sup>   飯野 翔太<sup>\*1</sup>   藤田 藍斗<sup>\*1</sup>   今泉 友之<sup>\*1</sup>   彦坂 修平<sup>\*1</sup>  
 Riho Ito   Shota Iino   Aito Fujita   Tomoyuki Imaizumi   Shuheik Hikosaka

<sup>\*1</sup> 株式会社パスコ 衛星事業部  
 PASCO CORPORATION Satellite Business Division

For satellite imagery land cover classification, different classification models are used depending on several factors such as sensor property or utilization purposes. There are so many types of satellites and several new satellites will be launched. As the number of types increases, it takes longer time and higher cost to make individual classification models for each kind of satellites. The current research aims to develop versatile classification models efficiently which can be used for multi-purposes. Land cover classification with Deep learning and the assessments were performed using open-access Landsat 5 satellite images and the Global Land Cover by National Mapping Organizations (GLCNMO) at the spatial resolution of 500 m. The Overall accuracy of the classification was 77.9% comparing with the high-resolution land cover classification map, showing the applicability of deep learning techniques in land-cover classification. Training a classifier with more satellite images, taken at different places and in different seasons, will enable to establish versatile classification models.

### 1. はじめに

衛星画像は画像そのものとしても価値はあるが画像を解析し付加価値を付けたプロダクトやサービスを創出することが衛星ビジネスの普及、市場拡大の上で非常に重要である。その中の 1 つの付加価値プロダクトとして土地被覆分類図がある。土地被覆図とは地表の物理形態や利用形態ごとに分類したものである。衛星画像からの土地被覆分類には様々な統計的解析手法が用いられており、近年ではディープラーニングを用いた研究も発表されているが、実利用という観点では検討されていない。

衛星画像からの土地被覆分類では、個々のセンサ特性(分解能、観測波長帯等)の違いや利用用途に応じて分類クラスが異なり、衛星種別ごとに都度設計しモデルを作成している。しかしながら、ベンチャー企業による超小型衛星の台頭など、衛星の種類は多岐にわたり、今後も衛星が増える中、個別に分類モデルを設計するのはコストと時間が掛かる。

一方で、ディープラーニングは教師データさえあれば個別にモデルを設計する必要が無く、教師データから自動的に特徴を設計できるため、モデル作成コストと時間を削減でき、ビジネスチャンスを広げる可能性がある。

そこで本研究では、ディープラーニングとフリーで入手可能な教師データを利用して土地被覆分類する手法を構築し、以下の i), ii) の観点で衛星画像を用いた土地被覆分類に対するディープラーニングの有効性の評価を行うことを目的とする。

- i) 簡単に(コストと時間を掛けずに)モデル作成が出来る。
- ii) 作成したモデルの精度が実利用に資する(土地被覆分類への適用が可能)か評価する。

### 2. 教師データ

#### 2.1 土地被覆分類図および衛星画像

本研究では土地被覆分類の手法として、画像認識の分野で高い性能を達成している Convolutional Neural Network (CNN) を用いた。CNN で教師付き学習を行うには、衛星画像の他に

衛星画像の各画素がどの土地被覆のクラスに属しているかを示した Ground Truth (GT) 画像が必要となる。本研究では、GT 画像作成コスト削減のため、フリーで入手可能である土地被覆分類図を GT 画像として使用した。国や地域によっては、個別に高解像度の土地被覆分類図は存在する。全球を対象とした土地被覆分類図作成を検討するため、本研究では、学習時の GT として、全球をカバーしている Global Land Cover National Mapping Organization 2008 (GLCNMO 2008) [Tateishi 14] を用いた。GLCNMO 2008 は、NASA の地球観測衛星 Terra/Aqua に搭載された光学センサ Moderate Resolution Imaging Spectroradiometer (MODIS)により 2008 年に観測されたデータ等を用いて作成されている[Tateishi 14]。GLCNMO 2008 は、解像度が 500m, 分類クラスが 20 個, 総合精度 (Overall Accuracy; OA)が 77.9%である。全球の土地被覆分類図については、GLCNMO 2008 の他にも、解像度が 300m, 分類クラスが 22 個, OA が 67.5%である GlobCover[Bontemps 11] や、解像度 1 km, 分類クラス 11 個, OA 80.2%である GLC SHARE [Latham 14]等がある。GLCNMO 2008 はフリーで公開されている全球の土地被覆分類図の中でも、比較的、解像度と OA の両方が高い。

また、解像度 500m の GT 画像から作成された分類モデルが、どれだけ精度よく(細かく)分類が可能であるかを評価するため、評価時の GT 画像として、GLCNMO 2008 よりも解像度が高い、AVNIR-2 高解像度土地利用被覆図 ver. 16.02 [JAXA 16] を用いた。この土地被覆分類図は、陸域観測技術衛星「だいち」に搭載されている光学センサ(AVNIR-2)より、2006 - 2011 年の期間に観測されたデータ等を用いて作成されており、OA が 76.9%, 解像度 10m, 分類クラス 10 個の日本全域の高解像度土地被覆分類図である。OA のみに着目すると、GLCNMO 2008 の方が AVNIR-2 高解像度土地利用被覆図 ver. 16.02 よりも OA が高いが、土地被覆分類図の OA は一般的に分類クラスが多く、解像度が高いほど低くなるため、分類クラスや解像度の条件が違うもとで単純に比較することはできない。

衛星画像は、GLCNMO 2008 に用いられているデータの観測年度に合わせて、2008 年前後に撮像された Landsat 5 の画像を使用した。Landsat5 は可視域 3 band, 近赤外域 2 band, 中間赤外域 1 band, 熱赤外域 1band の計 7 band の観測波長帯

連絡先:伊東里保, roithi7072@pasco.co.jp

株式会社パスコ

〒164-0001 東京都中野区中野 4-10-1 中野 CP イースト 3F

を持ち、解像度が 30m の中分解能衛星である。Landsat 5 は、リモートセンシングの分野においては、一般的に都市・森林・海洋等の環境変化モニタリング等、様々な解析に用いられている。本研究では、衛星画像を用いた土地被覆分類に対するディープラーニングの有効性を評価するにあたり、対象領域を、解像度の高い土地被覆分類図が存在する日本とした。対象領域の詳細については、表 1 に示す。地域や時期による汎用性を考慮するため、複数地域・多時期の画像を選択した。

表 1: 訓練・検証・評価に用いる衛星画像

(a) 訓練・検証用画像		
対象領域	撮像年月	画像サイズ
関東(東京都周辺)	2007年 5月	4985 x 4595
関東(東京都周辺)	2008年 10月	4988 x 4599
近畿(大阪府周辺)	2008年 9月	5697 x 3514
中部(愛知県周辺)	2007年 4月	4592 x 4588
(b) 評価用画像		
対象領域	撮像年月	画像サイズ
関東(東京都周辺)	2007年 11月	4987 x 4597
中部(静岡県周辺)	2008年 5月	5565 x 3290
九州(福岡県周辺)	2008年 10月	5202 x 4485

(a) 衛星画像



© Landsat Surface Reflectance products courtesy of the U.S. Geological Survey Earth Resources Observation and Science Center

(b) 土地被覆分類図

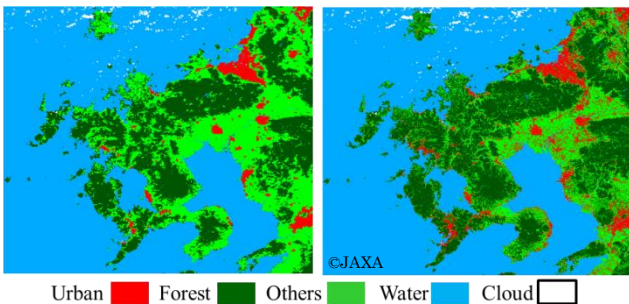


図 1: 衛星画像と土地被覆分類図の例。(a)SR 補正後の九州(福岡県周辺)の衛星画像。(b)左図は GLCNMO 2008 [国土地理院 14], 右図は AVNIR-2 高解像度土地利用被覆図 ver. 16.02 の土地被覆分類図。Urban, Forest, Others, Water の 4 クラスに統合(2.2 参照)後、雲域をマスクした。

本研究ではアメリカ地質調査所(USGS)がフリーで公開している、Landsat の画像に対し地上反射率補正を行った Higher Level Science Data Product [USGS 16]を用いた。地上反射率補正を行うことで、撮像条件(季節・太陽高度等)の違いによる画像の輝度値のばらつきを抑えることができる。また、このプロダクトには、各画像に対応した水、雪、雲、雲の影のマスク情報が含

まれている。GT 画像の高精度化のため、水域に関しては、GLCNMO 2008 と AVNIR-2 高解像度土地利用被覆図 ver. 16.02 に Higher Level Science Data Product の水域のマスク情報を加えたものを GT 画像として用いた(図 1)。

## 2.2 分類クラスの検討および訓練・検証・評価用パッチ

解像度 30m の Landsat 5 の画像から精度良く分類可能であり、実用に資するような分類クラスの検討を行った。そのうえで、本研究では、分類クラスを Urban, Forest, Water, Others の 4 クラスとした。本研究における分類クラスと、GLCNMO 2008 のクラス、AVNIR-2 高解像度土地利用被覆図 ver. 16.02 のクラスの対応を表 2 に示す。雲に覆われているエリアは 2.1 で述べた Higher Level Science Data Product の、雲のマスク情報を用いて訓練・検証・評価に使用するデータから除外した。GLCNMO 2008 は解像度が 500m、AVNIR-2 高解像度土地利用被覆図 ver. 16.02 は解像度が 10m であるが、衛星画像に合わせて解像度が 30m となるよう、Nearest Neighbor 法を用いてリサンプリングを行い、衛星画像と GT 画像の各画素を対応付けたうえで、衛星画像の各画素がどのクラスに分類されているかのラベル付けを行った。

本研究では CNN への入力を、衛星画像からランダムに切り出した 31 x 31 画素のパッチと、その中心 1 画素に対する正解ラベルの組とする。訓練・検証用の 4 枚の衛星画像(表 1 (a))から、ランダムに各クラス 40,000 枚(計 160,000 枚)切り出したパッチを訓練用パッチセット、各クラス 30,000 枚(計 120,000 枚)切り出したパッチを検証用パッチセットとする。また、評価用の衛星画像 3 枚(表 1 (b))から、同様に 31 x 31 画素のパッチを各クラス 15,000 枚(計 60,000 枚)切り出したものを評価用パッチセットとする。この時、各パッチの中心 1 画素は重複しないものとする。

表 2: 分類クラスの対応

本研究における 分類クラス	GLCNMO 2008 [国土地理院 14]	高解像度土地利用 被覆図 ver. 16.02 [JAXA 16]
Urban	都市部	都市部
Forest	常用広葉樹	常用広葉樹
	落葉広葉樹	落葉広葉樹
	常用針葉樹	常用針葉樹
	落葉針葉樹	落葉針葉樹
	混合樹林	
Water	水域	水域
Others	その他 13 クラス	その他 4 クラス

## 3. 分類・評価手法

### 3.1 CNN による学習

2.2 で述べた 31 x 31 画素のパッチを CNN への入力とした。また、最終層の活性化関数は Softmax 関数とし、入力パッチが各クラスである予測確率を出力とした。学習率や正則化パラメータ等の、学習によって自動的に決めることのできない超パラメータは、ベイズ的最適化を利用して検証用パッチセットの分類精度が最も高くなるような組み合わせを探索した。層の構造(組み合わせ、数等)は、超パラメータを固定したうえで幾通りか設定し、学習に要する時間や、検証用パッチセットに対する精度を考慮し、図 2 のように設定した。



### 3.2 評価

本評価では、i) 学習時の GT 画像として使用した GLCNMO 2008 に対するモデルの分類精度の評価と、ii) 解像度の高い AVNIR-2 高解像度土地利用被覆図 ver. 16.02 に対する、モデルと GLCNMO 2008 の分類精度の相対的な比較を行う。ii) では粗い GT 画像から更に精度の良い(細かい)分類が可能であるかを評価する。

作成したモデルの分類精度の評価には、2.2 で述べた評価用パッチセット 60,000 枚に対するモデルの分類結果の OA と、各クラスの Recall, Precision, F 値を用いた。Recall は実際の土地被覆をどの程度正しく再現できているか(再現率)を示し、Precision は分類結果がどの程度正しく土地被覆を表しているか(精度)を示す。分類方法によっては、Recall は高いが Precision は低い、またはその逆のパターンも起こり得るため、分類精度の評価指標として F 値にも着目する。モデルの実用性に対する評価指標は、OA  $\geq$  85.0 %、クラス毎の分類精度(本研究では Recall)  $\geq$  70.0 % [Thomlinson 99]を採用した。

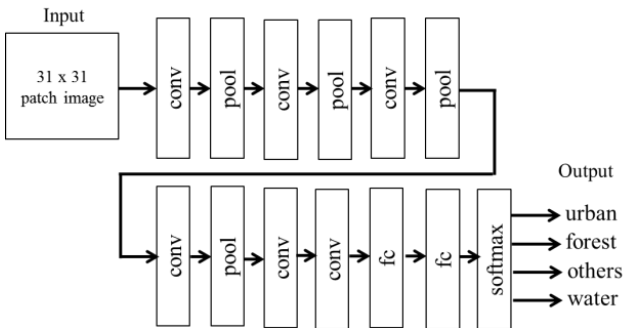


図 2: CNN の層構造の略図。conv:畳み込み層, pool:プーリング層, fc:全結合層。最終層の活性化関数は softmax 関数とし、出力は入力パッチに対する各クラスの予測確率となる。

### 4. 結果

図 1 に対する、CNN の土地被覆分類の予測結果を図 3 に示す。衛星画像において雲である領域は、Landsat の Higher Level Science Data Product を用いてマスクしている。3.2 で述べた i), ii) の結果はそれぞれ 4.1, 4.2 に示す。

#### 4.1 GLCNMO 2008 を GT 画像とした場合の精度評価

評価用パッチセット 60,000 枚の分類結果の Error Matrix を表 3 に示す。Urban, Forest, Water の 3 クラスについては、Recall がそれぞれ 95.9 %, 86.3 %, 90.6 %, F 値がそれぞれ 84.9 %, 86.1 %, 94.4%と、クラス毎の分類精度の実用評価基準を満たした。一方で、Others については Recall が 58.2 %, F 値が 64.0 %と、クラス毎の実用評価基準を下回る結果となり、OA も 82.8 %と実用評価基準には届かなかった。その理由として、GT 画像の解像度の粗さが考えられる。衛星画像の解像度が 30m であるのに対し、GT 画像の元の解像度が 500m であるため、衛星画像に対して誤った正解ラベルを割り当てている可能性がある。こうしたデータを学習に使用すると、本来は正解であるデータにもペナルティが課され、分類モデルの精度が低下すると考えられる。

#### 4.2 AVNIR-2 高解像度土地利用被覆図 ver. 16.02 を GT 画像とした場合の精度評価

評価用パッチセット 60,000 枚の分類結果の Error Matrix を表 4 (a)に示す。Urban, Forest, Water の 3 クラスについては、

Recall がそれぞれ 79.0 %, 75.3 %, 97.6 %, F 値がそれぞれ 75.0 %, 79.3 %, 97.3%と、高解像度な土地被覆図と比較してもクラス毎の分類精度の実用評価基準を満たした。一方で、Others に関しては Recall が 59.9 %, F 値が 60.1 %とクラス毎の実用評価基準を満たさず、OA も 77.9 %であり実用評価基準には届かなかった。

また、作成したモデルと、GLCNMO 2008 の分類精度を比較するため、AVNIR-2 高解像度土地利用被覆図 ver. 16.02 を GT 画像とした場合の GLCNMO 2008 の分類精度を算出した(表 4 (b))。GLCNMO 2008 の評価地点は、評価用パッチセットと同様の位置座標とした。ただし、Water については 2.1 で述べたように、同様の GT を用いているため、比較の対象から除外した。その結果 GLCNMO 2008 の 3 クラスの OA は 66.6 %, Urban, Forest, Others の Recall はそれぞれ 54.6 %, 71.0 %, 74.0 %, F 値はそれぞれ 65.6 %, 74.3 %, 61.1 %となった。作成したモデルの 3 クラスの OA は 72.1 %であり、GLCNMO 2008 よりも 5.5 %高く、AVNIR-2 高解像度土地利用被覆図 ver. 16.02 に対しては、作成したモデルの方が精度よく分類できていることがわかる。

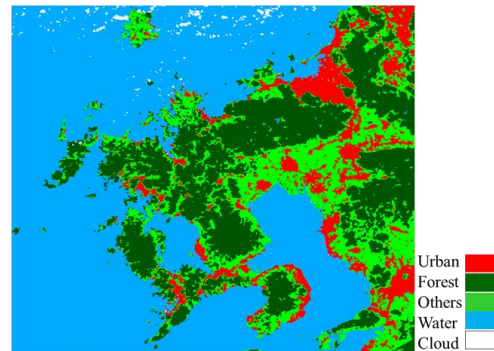


図 3: 図 1 の衛星画像に対する分類モデルの予測結果。衛星画像の雲域は、Landsat の Higher Level Science Data Product を用いてマスクしている。

### 5. おわりに

本研究では土地被覆分類に対するディープラーニングの有効性を、モデル作成に掛かるコストおよび時間の削減と、実用に資する精度で分類が可能であるかの 2 つの観点で評価を行った。コストと時間の削減という観点においては、訓練・検証用パッチ数や、分類を行う画像の数やサイズ、計算機の環境等によって変動する。2.2 で述べたパッチセットと、3.1 で述べたモデル設定の条件下では、CNN の超パラメータの設定に始まり、学習(モデルの作成)、各評価領域(表 1 (b))の土地被覆分類を約 4 日間で行うことができた。

また、Urban, Forest, Others の 3 クラスについて、AVNIR-2 高解像度土地利用被覆図 ver. 16.02 に対する、作成したモデルと GLCNMO 2008 の分類精度を、相対的に比較した。その結果、GLCNMO 2008 よりも、作成した分類モデルの OA の方が 5.5 %高かった。一方で、作成した分類モデルの OA は、実用評価基準である 85.0 %には届かなかったが、2.1 で述べたように、他の土地被覆分類図も OA 85.0 %には到達していない。

更なる精度向上のため、以下のような対策が考えられる。まず一つ目は GT 画像として使用する土地被覆分類図の精度向上である。本研究では、衛星画像と比べて粗い解像度の GT 画像を教師データとして用いたため、訓練・検証正解ラベルが誤っているパッチが使用された可能性がある。衛星画像と比べて解像度が粗い土地被覆分類図から、より精度の高い GT 画像を

表 3 : GLCNMO 2008 に対する Error Matrix

		Prediction				Total	Recall (%)
		Urban	Forest	Others	Water		
Reference	Urban	14383	14	532	71	15000	95.9
	Forest	138	12950	1846	66	15000	86.3
	Others	4183	2033	8724	60	15000	58.2
	Water	167	80	1163	13590	15000	90.6
	Total	18871	15077	12265	13787	60000	
Precision (%)		76.2	85.9	71.1	98.6		
F-value (%)		84.9	86.1	64.0	94.4		OA = 82.8 %

表 4: AVNIR-2 高解像度土地利用被覆図 ver. 16.02 に対する Error Matrix

(a)本研究で作成した分類モデル

		Prediction				Total	Recall (%)
		Urban	Forest	Others	Water		
Reference	Urban	11844	267	2722	167	15000	79.0
	Forest	591	11289	3051	69	15000	75.3
	Others	3988	1820	8980	212	15000	59.9
	Water	145	83	135	14637	15000	97.6
	Total	16568	13459	14888	15085	60000	
Precision (%)		71.5	83.9	60.3	97.0		OA=77.9 %
F-value (%)		75.0	79.3	60.1	97.3		OA (Urban, Forest, Others) =72.1 %

(b) GLCNMO 2008

		Prediction			Total	Recall (%)
		Urban	Forest	Others		
Reference	Urban	8112	680	6058	14850	54.6%
	Forest	273	10615	4054	14942	71.0%
	Others	1512	2354	10983	14849	74.0%
	Total	9897	13649	21095	44641	
Precision (%)		82.0	77.8	52.1		
F-value (%)		65.6	74.3	61.1		OA (Urban, Forest, Others) = 66.6%

作成する手法を検討する必要がある。加えて、衛星画像の解像度に合った土地被覆分類図を、GT 画像として選択する必要がある。二つ目は、学習に様々な地域や多時期の画像を使用することである。今回は土地被覆分類に対するディープラーニングの有用性の検証のため、使用する衛星画像や対象領域を絞った。様々な地域・時期の画像を使用することで、地域性や季節性の異なる画像を精度よく分類できるような、より汎用性の高いモデルが作成可能であると考えられる。三つ目は、分類クラスの検討である。クラスの分け方は分類の精度に影響を与えるため、土地被覆分類図の利用目的に応じた適切な分類クラスの検討が今後の課題となる。

本手法により解像度の粗い GT 画像からより細かい分類ができるモデルを作成できたことから、ディープラーニングは衛星画像に対する土地被覆分類に対して有効であり、上記のような対策を検討することによって実利用に資する精度を出すことが可能であると考えられる。

参考文献

[Bontemps 11] Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., Ramos Perez, J.: GLOBCOVER 2009—Products Description and Validation Report (2011)  
 [JAXA 16] 宇宙航空研究開発機構 (JAXA) 地球観測研究センター: 高解像度土地利用土地被覆図ホームページ, アクセス日時:2016年3月22日  
 <[http://www.eorc.jaxa.jp/ALOS/lulc/lulc\\_jindex.htm](http://www.eorc.jaxa.jp/ALOS/lulc/lulc_jindex.htm)>  
 [Latham 14] Latham, J., Cumani, R., Rosati, I., and Bloise, M.: Global Land Cover SHARE (GLC-SHARE) Database Beta-Release Version 1.0-2014: FAO, Rome, Italy (2014)  
 [Tateishi 14] Tateishi, R., Hoan, N., T., Kobayashi, T., Alsaaidh, B., Tana, G., and Phong, D. X.: Production of Global Land

Cover Data - GLCNMO 2008: Journal of Geography and Geology, Vol. 6, No. 3, Canadian Center of Science and Education (2014)  
 [Thomlinson 99] Thomlinson, J., R., Bolstad, P., V., and Cohen, W., B.: Coordinating Methodologies for Scaling Landcover Classifications from Site-Specific to Global: Steps toward Validating Global Map Products: Remote Sensing of Environment 70.1, pp. 16-28 (1999)  
 [USGS 16] United States Geological Survey: Landsat Surface Reflectance High Level Data Products, Accessed on 22 Mar. 2016  
 <[http://landsat.usgs.gov/CDR\\_LSR.php](http://landsat.usgs.gov/CDR_LSR.php)>  
 [国土地理院 14] 国土地理院, 千葉大学, 協働機関: 地球地図全球版, アクセス日時:2016年3月22日  
 <[http://www.gsi.go.jp/kankyochiri/gm\\_global.html](http://www.gsi.go.jp/kankyochiri/gm_global.html)>