

## 地球観測衛星画像上の地物認識

## Object recognition on Earth-observing satellite imagery

中村良介*1	石井智大*1,2	野里博和*1	坂無英徳*1
Ryosuke NAKAMURA	Tomohiro ISHII	Kiyokazu NOSATO	Hidenori SAKANASHI
シモセラ・エドガー*2	望月義彦*2	飯塚里志*2	石川博*2
Simo-Serra EDGAR	Yoshihiko MOCHIZUKI	Satoshi IIZUKA	Hiroshi ISHIKAWA

\*1 産業技術総合研究所  
National Institute of Advanced Industrial Science and Technology

\*2 早稲田大学  
Waseda University

Earth-observation satellites are producing tremendous amount of digital imagery in recent years. For the efficient analysis, automatic object recognition is an indispensable tool. We designed a convolutional neural network (CNN), which can locate any objects on the multiband images obtained by the most popular middle-resolution satellite Landsat-8. Our experiments demonstrated that photovoltaic power stations can be accurately identified through the spatial and spectral features with relatively small sets of training data.

## 1. はじめに

宇宙空間に打ち上げられる地球観測衛星の数は年々増加しており、そのペースは最近とくに加速している。たとえば 2014 年 1 月に初号機が打ち上げられた PlanetLabs 社の Dove 衛星の総数は、2016 年 3 月末には 133 機に達した [1]。近い将来、Dove 衛星は全陸域を 5 メートルの空間分解能で毎日観測し、地球上のあらゆる変化を継続的にモニタリングする予定である。こうした地球観測衛星群の 1 日あたりのデータ生成量は数百テラバイトに達するため、効率的な解析を行うためには画像上において物体(地物)を自動的に判別するアルゴリズムが不可欠である。一般物体認識の分野では、近年になって Convolutional Neural Network (CNN) が従来手法に比べて圧倒的に高い性能を示すようになってきた [2]。一般物体認識に用いられる多くの画像は、地上において水平方向にカメラを向けて取得されている。一方、人工衛星や航空機、ドローンは上空から下を向いて画像を取得する。しかしながら、空間分解能さえ十分に高ければ、両者の間に本質的な違いはない。実際、一般物体認識用に CNN が学習した特徴量は、空間分解能の高い航空写真や衛星画像においても既存の特徴量と同等以上の性能を示すことが知られている [3]。本研究では、中程度の分解能(～30m)の大規模な衛星画像データに対して CNN を適用することで、地物自動認識を試みた。

## 2. 対象データと解析手法

Landsat 8 号衛星は 2013 年の 2 月に打ち上げられ、地球上の任意の地点を 15~100m の空間分解能で 16 日ごとに観測し続けている [4]。Landsat シリーズのデータは、40 年以上にわたって地球の陸域を観測し続けているため、長期的なモニタリングに適している。また誰もが自由に使えるオープンデータとして公開されており、毎日 500GByte におよぶデータを取得しているため、将来的にさらに大規模な衛星画像処理を行う場合のべ

連絡先: 中村良介, 産業技術総合研究所人工知能研究センター, 茨城県つくば市梅園 1-1, 029-862-6702, r.nakamura@aist.go.jp

ンチマークとなる。つまり Landsat 画像上で実用的な自動地物認識ができれば、将来的な実利用への可能性が大きく拓ける。そこで本研究では、最初のステップとして、比較的大型の太陽光発電所すなわちメガソーラーの自動検知・識別を試みた。Landsat 8 号に搭載されたセンサは、紫外域から赤外域にわたる 11 の波長域(バンド)で一度に 180km 四方の領域を観測する。ここでは空間分解能が 30m のバンド 4/3/2 (可視域の RGB に相当)、および近赤外域をカバーする 5/6/7 を用いる。

従来の CNN を用いた地物認識の研究においては、画素単位で道路・建物・自動車といった地物の教師を作成し、学習・識別を行っていた [5]。一方、我々は元の Landsat 画像を 16x16 pixel のセルに分割し、このセル単位で教師付き学習を行う。その後、画像中の全セルについて対象地物かどうかの 2 値分類を行うことで、大規模な衛星画像データから効率的に対象地物を発見することが可能となる(図 2)。解析の基本単位として、画素ではなくセルを用いることで、教師データ作成のコストと学習時間を大幅に削減することができる [6]。

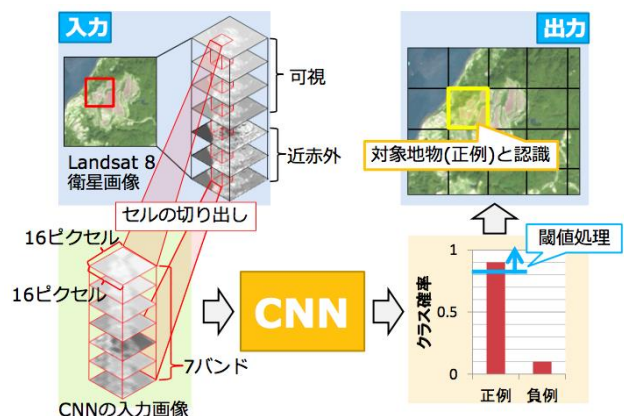


図2 CNNによるLandsat画像上の地物判読フロー

教師データの正例は、エレクトリカル・ジャパンが公開している日本全国のメガソーラーリストから、Landsatの分解能で確実に捉えられる出力 5MW 以上のもののみを選択することで作成した [7]。対象となるメガソーラーが複数セルにわたる場合は、メガソーラーの面積割合がおおよそ 20% 以上のセルのみを正例として設定した。一方、負例は正例に隣接しない領域からランダムに選択した。

### 3. 結果

図3に、認識結果の一例を示す。黄色い四角が Landsat 画像上でメガソーラーが存在すると判定されたセルである。拡大図として示されている Google Map などの高空間分解能衛星画像や、実際に現地で取得した画像から、ここでの検知例はいずれも太陽発電施設であることが確認された。

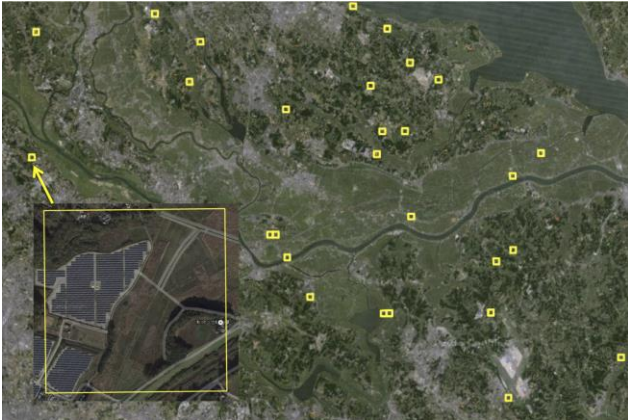


図3 利根側流域におけるメガソーラーの認識結果。  
背景図および拡大図は Google Map。

より定量的な評価を行うため、検証用データセットを用いて精度を見積もった(図4)。単独のバンドだけでは、ほとんど認識できていないが、利用するバンド数を増やすと精度が高くなるのがわかる。特に赤外域での反射率を捉える band 5~7 を加えることで認識精度が向上している。これは太陽電池パネルがこの波長域で特徴的な反射スペクトルを持ち、かつ CNN がメガソーラーの空間的な構造だけでなく反射スペクトルパターンも適切に学習しているためだと解釈できる。Recall に比べて Precision の値が低いが、False positive を個別にチェックしたところ、そのほとんどがエレクトリカル・ジャパンのリストに含まれていない、小規模な太陽光発電所であることが判明した。つまり実際の認識性能は、図4に示される値よりもはるかに高いことが期待される。一方で、より高分解能の衛星画像や航空写真あるいは現地写真などで存在が確認できる1~数 PIXEL の大きさの小規模な太陽光発電所は検出できていないものがある。これは前節で述べたとおり、セルサイズ(16x16)の 20%、すなわち 50pixel 以上の広がりを持つものしか、教師データとして採用していないことが原因だと考えられる。

### 4. まとめ

CNN を用いて Landsat 画像上のメガソーラーの自動認識を行った。どの波長域でも、単独のバンドだけでは、ほとんど認識できないが、利用するバンド数を増やすことで性能が向上する。これは CNN が対象地物の空間的な特徴だけでなく、波長方向の特徴も的確に学習しているためだと考えられる。全バンドを用

いた場合の IoU (Intersection of Union)は5割程度であるが、これは評価データに含まれていない小規模な太陽光発電設備の検知を False positive として判定しまっているためである。より高分解能の衛星画像や航空写真に基づいて、小規模な太陽光発電所を検証データに追加すれば、実際の精度はさらに高くなるはずである。

一般に CNN のような深層学習においては、膨大な教師データが必要となると考えられているが、本研究で学習に用いた正例の教師セル数は 300 程度である。ほかの地物についても同程度の教師データを与えることで、大規模な衛星画像からの全球自動探索が可能となる可能性が高い。また将来的に、空間分解能の高い衛星画像や航空写真を用いれば、より多様な地物を識別できるようになると期待される。

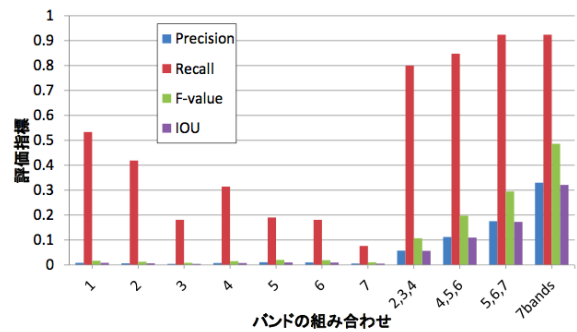


図4 異なるバンドを用いた時の認識精度の比較

### 参考文献

- [1] <https://www.planet.com/pulse/>
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1106–1114.
- [3] O. A. B. Penatti, K. Nogueira, and J. A. dos Santos, "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?" in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2015. □
- [4] <http://landsat8.geogrid.org/>
- [5] V. Mnih and G. E. Hinton, "Learning to detect roads in high-resolution aerial images," in *Computer Vision—ECCV 2010*. Springer, 2010, pp. 210–223.
- [6] T. Ishii, R. Nakamura, H. Nakada, Y. Mochizuki, and H. Ishikawa, "Surface object recognition with CNN and SVM in landsat 8 images," in *14th IAPR International Conference on Machine Vision Applications (MVA), 2015*. IEEE, 2015, pp. 341–344. □
- [7] <http://agora.ex.nii.ac.jp/earthquake/201103-eastjapan/energy/electrical-japan/>