

人と協調する半自己学習に基づく衛星画像上の地物検出

上原 和樹^{1,a)} 野里 博和¹ 村川 正宏¹ 坂無 英徳¹

概要：衛星画像上の地物検出において、畳み込みニューラルネットワーク（CNN）は非常に有効である。しかしながら、リモートセンシングの領域においては、専門性やコストの観点から CNN を学習するのに十分なデータセットを用意することが困難である。そこで本研究では、限られたデータセットで CNN を効率よく学習するために、半自己学習手法を提案する。これは、半教師あり学習の一つである self-training 法を活用し、ユーザと協調しながらラベル付きデータを収集し、自己学習により学習効率を向上する手法である。衛星画像を用いた地物検出タスクに適用したところ、提案手法はすべてのデータを用いて学習した CNN と同等の性能を約 18% のデータ量で得ることができた。

キーワード：半自己学習, Self-training 法, 能動学習, 衛星画像, 地物検出

1. はじめに

衛星画像上の自動的な地物検出は土地計画 [1] [2] や災害の被害推定 [3] といった様々な応用の観点から非常に重要なタスクである。畳み込みニューラルネットワーク (CNN) が様々な画像認識タスクにおいて高い成績を収めるなど注目を集めており [4] [5], 衛星画像の自動認識に関する研究も盛んに行われている [6] [7] [8]。

一方、そのような性能を達成するためには多量のラベル付きデータを要するが、リモートセンシングの領域では学習に活用できるラベル付きデータを多量に用意することが困難である。その理由として、衛星画像へのラベル付加に高い専門性が要求されることや、一つの画像が対象地物に対し非常に大きく、対象を探し出すことに集中力を要することなどが挙げられる。

本研究では、少ないデータから CNN を効率よく学習するために、半自己学習手法を提案する。本手法の狙いは、識別器による自己学習と要所でのユーザの介入により、少ないコストで CNN の学習効率を向上することである。識別器の自己学習をするために、半教師あり学習の一つである self-training 法 [9] を導入する。self-training 法は、ラベル付きデータから学習したモデルを用いて、ラベルなしデータに仮ラベルを付与する。その後、モデルが出力したクラス確率（確信度）が一定の閾値を超えた仮ラベルデータと

既存のラベル付きデータをあわせてデータ数を増やし、再度学習することでモデルの性能を向上する手法である [9]。しかしながら、モデルの確信度が高い場合においても、その仮ラベルが真に正しいとは限らないという問題点があり、場合によっては学習を妨げる恐れがある。特に、学習データ数が少ないほどこの傾向は顕著であると考えられる。

そこで本研究では、self-training 法により付けられた仮ラベルが正しいかどうかを交差検定により自動的に評価し、学習に有用そうなデータのみを仮ラベル候補として選出する。これにより、学習を妨げる可能性のあるサンプルが学習データセットに混入することを防ぐ。さらに、自己学習で性能が改善されない場合は、能動学習 [10] を用いたユーザの介入により学習効率を高める。能動学習は、まだラベルのないデータに対し、ユーザと識別器で対話的にラベルを作成しながら学習する手法である。モデルを効率よく改善するようなサンプルの選択方法に関する研究が多くなされており [10], この手法を活用することで、ラベル付きデータを補強し学習モデルの性能を向上をさせる。つまり、自己学習と能動学習を効果的に組み合わせ、少ないユーザの介入でモデルの性能を改善することを目指す。

提案手法の有効性評価をするため、Landsat 8 号による衛星画像を用いたゴルフ場検出タスクを実施した。その結果、約 18% のラベル付きデータで、全データを用いた場合と同等の性能を達成した。さらに、従来の能動学習との比較においても、早期から安定して良い性能を示し、また、約 65% 程度のデータ量で同等の性能を得ることができた。

¹ 国立研究開発法人 産業技術総合研究所 人工知能研究センター
National Institute of Advanced Industrial Science and Technology (AIST) 1-1-1 Umezono, Tsukuba, Ibaraki 305-8560, Japan

^{a)} k-uehara@aist.go.jp

2. 関連研究

本研究に関連する研究として、Wang ら [11] は、能動学習が学習に有効なサンプルだけを選択することから、学習に大量のデータを要する CNN には不十分であることを指摘し、self-training 法によりラベルなしデータから自動的に学習データを収集する手法 (Cost-Effective Active Learning, CEAL) を提案した。CEAL は、ラベル付きデータの少ない学習の初期段階では分類モデルの精度が低いことから、学習段階に応じた閾値の調整を導入しており、確信度が高いサンプルだけを学習に用いる。しかしながら、自動的に追加されるサンプルが学習に有効かどうかの評価は行わないため、追加したサンプルによりモデルの性能が下がる可能性がある点で本研究と異なる。

3. 提案手法

3.1 問題設定

データ集合 $D = \{x_i\}_{i=1}^{l+u}$ は、ラベル付きデータ集合 $D^L = \{x_i, y_i\}_{i=1}^l$ とラベルなしデータ集合 $D^U = \{x_i\}_{i=l+1}^{l+u}$ から構成される。本研究においては、 x_i は衛星画像を表し、 y_i は対象地物であるか否かを示すラベルを表す。CNN を f_θ とし、既存のラベル付きデータ D^L により、モデルパラメータ θ を学習する。学習後は、ラベルなしデータ集合 D^U から、そのモデル f_θ の改善が期待できるサンプル $Q = \{x_i\}_{i=l+1}^{l+q}$ をあらかじめ定めた基準により q 個選択し、ユーザに提示する。ユーザはこれらのサンプルにラベルを与え $\tilde{Q} = \{x_i, y_i\}_{i=l+1}^{l+q}$ とし、ラベル付きデータ集合 D^L にラベル付けしたデータ $\tilde{Q}$ を加える。この学習とラベル付けの繰り返しにより、可能な限り少ない試行回数で汎化性能の高いモデルの学習を目指す。

3.2 半自己学習手法

提案手法は、ユーザと識別器間で対話的にサンプルを収集する対話部と、自動的なラベルの作成により学習する自己学習部から構成される。基本的なアイデアは、既存のラベル付きデータに基づき学習モデルを構築し、そのモデルによる自己学習でモデルの性能向上を狙う。ここで、モデルが自己学習時に性能を向上できなくなった場合に限り、ユーザにラベル付けを要求して学習モデルを強化する。これにより、極力少ないユーザの介入で学習モデルの性能を改善する。

自己学習部では、半教師あり学習手法である self-training 法をベースとする。self-training 方では、ラベル付きデータから学習したモデルの識別結果に基づいて、ラベルなしデータに仮ラベルを付与する。しかし、モデルの確信度が高い場合であってもその仮ラベルが真に正しいとは限らず、場合によっては学習を妨げる可能性がある。そこで本

研究では、self-training 法において得られる仮ラベル付きデータを評価し、学習に有効なサンプルだけを学習に用いることでこの問題に対処する。このとき、一度に多量の仮ラベルデータを扱うと、学習に有効なサンプルの選出が困難となるため、少量ずつ評価する。

具体的には、既存のラベル付きデータだけで学習したモデルと、少量の仮ラベル付きデータ候補と既存のラベル付きデータをあわせて学習したモデルの性能を比較し、性能を向上できる場合は、その仮ラベル付きデータ候補は学習に有用であるとし仮ラベル付きデータセットに蓄積する。仮ラベル付きデータを蓄積し続け、モデルの性能向上が見込めない場合や、モデルの性能が悪化した場合は、ユーザによるラベル付きデータの補強でモデルの性能向上を促す。

以降の節では、対話部と自己学習部について詳細を述べる。

Algorithm 1 :対話部アルゴリズム

Input:

Initial training set $D^L = \{x_i, y_i\}_{i=1}^l$.
Unlabeled samples $D^U = \{x_i\}_{i=l+1}^{l+u}$.
Number of samples q queried in each iteration.
A classifier f_θ .

Initialization : $D^{PL} = \phi$

```
1: repeat
2:   Perform semi-self-training.
      $f_\theta, D^L, D^U, D^{PL} \leftarrow \text{SemiSelfTrain}(D^L, D^U, D^{PL})$ 
3:   for ( $i = 1$  to  $u$ ) do
4:     Rank unlabeled sample  $x_i \in D^U$  based on query strategy.
5:   end for
6:   Select  $q$  samples to be labeled, and display the samples to the user.
7:   The user assigns a label to the queried samples.
      $\tilde{Q} = \{x_j, y_j\}_{j=l+1}^{l+q}$ 
8:   Add the labeled samples to the labeled set.  $D^L \leftarrow D^L \cup \tilde{Q}$ 
9:   Remove the labeled samples from the unlabeled pool.
      $D^U \leftarrow D^U \setminus \tilde{Q}$ 
10: until Stopping criterion is met.
11: return Trained classifier  $f_\theta$ 
```

3.2.1 対話部アルゴリズム

対話部アルゴリズムを、Algorithm 1 に示す。対話部においては、(1) まずラベル付きデータセット D^L 、ラベルなしデータセット D^U 、仮ラベル付きデータセット D^{PL} (初期状態では空集合) を用いて 3.2.2 節で述べる方法により自己学習を行う。(2) 自己学習が終わると、 D^U 、 D^{PL} がそれぞれ更新され、その学習結果によるモデル f_θ を得る。(3) その後、 f_θ に対し、モデルの性能向上が期待できるサンプルを D^U から q 個選択し、ユーザにラベル付けを要求する。(4) ラベルが付けられた集合 $\tilde{Q} = \{x_i, y_i\}_{i=l+1}^{l+q}$ を D^L に統合し、 D^U から取り除く。以上、(1)~(4) について終了基準を満たすまで繰り返す。

モデルの性能向上が期待できるサンプルの選出方法は

これまで様々な方法が考案されている [10] [12]. 本研究では, 検出対象であるか否かの二値分類であるため, モデルの確信度が 50%に近い順からサンプルを取得する least confidence 戦略による選出方法を用いた.

Algorithm 2 : 自己学習部アルゴリズム

Input:

Labeled samples D^L .
Unlabeled samples D^U .
Pseudo-labeled samples D^{PL} .

```

1: repeat
2:   Train a classifier  $f_\theta$  using labeled dataset  $D^L$  and pseudo-labeled dataset  $D^{PL}$ .
3:   Assign pseudo-labels to unlabeled samples  $D^U$  using the classifier  $f_\theta$ .
4:   Select  $n$  samples from pseudo-labeled samples.  $d^{PL} = \{x_i, \tilde{y}_i\}_{i=1}^n$ .
5:   for ( $k = 1$  to  $K$ ) do
6:     Divide labeled dataset to  $D_k^T$  and  $D_k^E$ .
7:     Train classifier  $f_\theta^T$  using dataset  $D_k^T$ .
8:     Train classifier  $f_\theta^{T \cup PL}$  using dataset  $D_k^T$  and  $d^{PL}$ .
9:     Evaluate classifier  $f_\theta^T$  using dataset  $D_k^E$  as  $e_k^T$ .
10:    Evaluate classifier  $f_\theta^{T \cup PL}$  using dataset  $D_k^T$  as  $e_k^{T \cup PL}$ .
11:   end for
12:   if  $E^T \leq E^{T \cup PL}$  then
13:     Add the small set of pseudo-labeled samples  $d^{PL}$  to the pseudo-labeled set  $D^{PL}$ .
14:     Remove the small set of pseudo-labeled samples  $d^{PL}$  from unlabeled pool  $D^U$ .
15:   else
16:     Check pseudo-labeled set  $D^{PL}$ .
17:      $D^{PL}, D^R \leftarrow \text{check}(D^{PL})$ 
18:     Add rejected samples  $D^R$  to unlabeled pool  $D^U$ .
19:   end if
20: until Stopping criterion is met
21: return Trained classifier  $f_\theta$ , unlabeled pool  $D^U$ , pseudo-labeled set  $D^{PL}$ .

```

3.2.2 自己学習部アルゴリズム

自己学習部のアルゴリズムを Algorithm 2 に示す. 自己学習部においては, 入力された D^L と D^{PL} を用いて, モデル f_θ を学習し, そのモデルの D^U に対する予測に基づき仮ラベル $\tilde{y}$ を与える. 仮ラベル候補をユーザが予め定めた基準に従って n 個選出し, $d^{PL} = \{x_i, \tilde{y}_i\}_{i=1}^n$ とする. ここで, 仮ラベル候補 d^{PL} が学習に有効であるか否かを評価するため, K-fold 交差検定を用いる.

K-fold 交差検定の各試行では, まず D^L を割合 (K-1):1 でそれぞれ学習用データ D_k^T と評価用データ D_k^E に分割する. 新たに D_k^T だけで学習したモデル f_θ^T , ならびに D_k^T と d^{PL} をあわせて学習したモデル $f_\theta^{T \cup PL}$ を用意する. K 回分それらのモデルを評価し, その評価値に基づき自己学習を継続するか否かを決定する. モデル $f_\theta^{T \cup PL}$ による評価値 $E^{T \cup PL}$ が, モデル f_θ^T の評価値 E^T よりも上回れば, このとき学習に用いた仮ラベル候補 d^{PL} を D^{PL} に加えて

自己学習を継続する. 一方, $E^{T \cup PL}$ が E^T より低い場合は, これまで蓄積してきた仮ラベルデータ D^{PL} の内容を後述の方法により見直す. 仮ラベルデータの見直しにより, 保持するデータ D^{PL} と破棄するデータ D^R に分類する. D^{PL} から除かれたデータ D^R は, ラベルなしデータとして再度 D^U に統合される. この時点で自己学習を終了し, f_θ , D^U , D^{PL} を返す.

仮ラベルデータの見直し方法として様々な方法が考えられるが, 本研究では K-近傍法を用いた. 具体的には, 現在得られている CNN モデル f_θ の識別層直前の出力を特徴ベクトルとし, ラベル付きデータだけを用いて K-近傍の予測モデルを構築する. その後, 仮ラベル付きデータから f_θ を用いて特徴抽出をし, K-近傍によりラベルの予測を行う. このとき, 仮ラベル $\tilde{y}$ と K-近傍による予測結果が一致しないサンプルは仮ラベルデータから排除する (D^R). これにより, 学習を妨げる可能性のあるデータが仮ラベル付きデータセットに混入することを防ぐ.

4. 実験

提案手法の有効性を確認するため, 衛星画像の地物検出タスクにて比較実験を行った. 本稿ではゴルフ場を検出タスクの対象地物とした.

4.1 Landsat 画像

本稿では, 衛星画像として Landsat 8 号の画像を採用した. Landsat プロジェクトでは 40 年以上もの間, 地球を観測し続けており, Landsat 8 号はそのプロジェクトにおける最新の衛星である. また, Landsat 画像は Web 上で公開されており, 無償でダウンロード可能である^{*1}. したがって, Landsat 画像を用いた物体検出は, 様々な応用の観点から有用である.

Landsat8 号は, 表 1 に示すとおり 11 バンドの異なる周波数帯, 空間解像度で画像を取得しているため, 共通の空間解像度 (30 m) を持ったバンド 1 から 7 までは入力画像として用いた.

4.2 データセット

物体検出用のデータセットを構築するため, 衛星画像を 16×16 ピクセル (480×480 m² 相当) のパッチ画像に分割した. データセットはパッチ画像中, 20%以上のゴルフ場領域を有するものを正例, それ以外を負例として作成し, これらをグラウンドトゥルース (正解) データとする. モデルの学習には関東地方のシーン画像を用い, 評価には東海地方のシーン画像を用いた. それぞれのシーン画像が含む正例, 負例数を表 2 にまとめた.

^{*1} <https://earthexplorer.usgs.gov>

表 1 Landsat8 号における各バンドの波長帯 (μm) と空間解像度 (m).

バンド	波長 [μm]	空間解像度 [m]
1 (紫外)	0.43-0.45	30
2 (可視光, 青)	0.45-0.51	30
3 (可視光, 緑)	0.53-0.59	30
4 (可視光, 赤)	0.64-0.67	30
5 (近赤外)	0.85-0.88	30
6 (短波長赤外)	1.57-1.65	30
7 (短波長赤外)	2.11-2.29	30
8 (パンクロマチック)	0.50-0.68	15
9 (雲検知)	1.39-1.38	30
10 (熱赤外)	10.60-11.19	100
11 (熱赤外)	11.50-12.51	100

表 2 各シーンにおける正解データの正例数と負例数.

シーン ID	正例数	負例数
LC81070352015218LGN00 (関東)	3,418	158,056
LC81080362016196LGN00 (東海)	570	161,188

表 3 ネットワーク構成.

layer	kernel size	output
input	-	$7 \times 16 \times 16$
convolution	3×3	$32 \times 14 \times 14$
convolution	3×3	$32 \times 12 \times 12$
convolution	3×3	$32 \times 10 \times 10$
fully connected	-	1024
fully connected	-	2

4.3 CNN のネットワーク構成

本研究では, Landsat 画像を用いたメガソーラーの検出タスクにおいて良い性能を示した CNN[6] に微修正を加えたネットワーク構成を用いた. その構成の詳細を表 3 に示す. ネットワーク構成は, 3 層の畳み込み層と 2 つの全結合層とした. 本実験で用いる入力画像は 16×16 と小さいため, プーリングは採用しなかった. また, スクラッチから効率よくモデルを学習するために, 各畳み込み層の出力に batch normalization, ReLU 関数を順に適用した. さらに過学習を防ぐため, 3 層目の畳み込み層直後にドロップアウトを適用し, その確率を 50% とした.

4.4 実験条件

提案手法は, ユーザとの対話的な学習を繰り返すことで性能を向上する手法であるため, 本実験では, 比較対象として通常の能動学習のみの手法 (AL) と能動学習に加えて, self-training 法により学習する手法 (AL-ST) を用いた. AL は, ユーザに問い合わせた追加したラベル付きデータを用いて学習する. AL-ST は, ユーザに問い合わせた追加したラベルに加え, self-training 法で仮ラベルをさらに

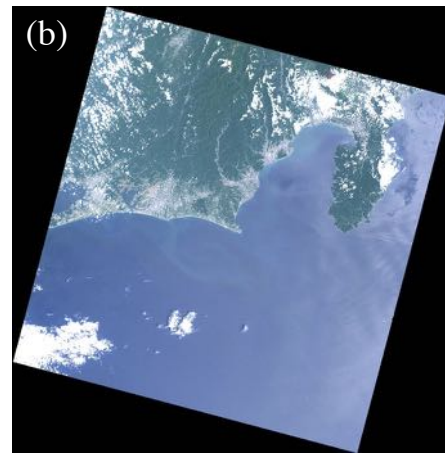
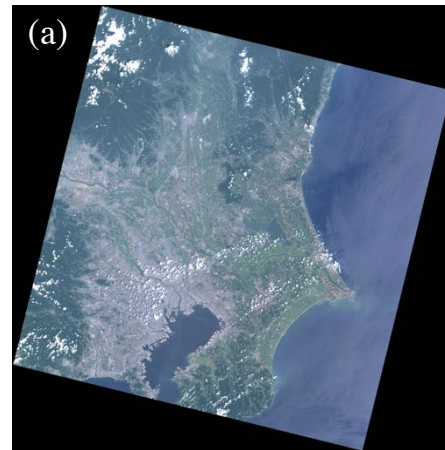


図 1 モデルの学習と評価に用いたシーン画像. (a):LC81070352015218LGN00 (関東地方), モデルの学習に用いた. (b):LC81080362016196LGN00 (東海地方), (a) において収集したデータで学習したモデルの評価に用いた.

追加して学習を行う. このとき, 提案手法のような仮ラベルの評価は行わない. これらの手法においてユーザに問い合わせるサンプル選択の戦略は, いずれも提案手法と同様に least confidence 戦略とした.

少数データからの性能の改善度合いを確認するため, 初期の正例数を 20, 負例数を 400 とし, 各ステップにおいてユーザにラベル付けを要求する数 (3.1 節の q) は 200 と設定した. また, 提案手法において仮ラベル候補を選出する際に追加するデータ数 (3.2.2 節の n) は, ラベル付きデータセットの正例数と負例数それぞれの 2 倍ずつ追加するものとし, 仮ラベルの確信度が高いものから順に選出した.

ユーザがラベル付けを行う場合, そのユーザの状態, 例えば疲労度や集中力, 判断基準等によりラベル付けの結果が変わる可能性が懸念される. そこで, 実験を公正に行うために, 4.2 節で述べた正解データを用いて自動的にラベルを与えた. 各試行実験は 10 回行い, その平均値を示す.

提案手法は, 仮ラベルの評価状況に応じて学習データに付加する仮ラベル数が変動する. 実験条件を同等とするため, 比較手法である AL-ST については, 各ステップで付

加する仮ラベル数を提案手法で付加した仮ラベルと同数とした。各ステップで付加したラベル数をそれぞれ表 4 に示す。

衛星画像からの地物検出タスクでは、画像全体に対する対象地物の割合は非常に小さくなるため、F 値を評価指標として採用した。F 値は適合率と再現率の調和平均として算出され、正例と負例のバランスが偏ったデータセットの評価において有効であることが示されている [13]。

4.5 実験結果

図 2 は各ステップにおける F 値の推移を示す。実線はそれぞれ、提案手法 (青)、AL (黄)、AL.ST (緑) を表し、赤の破線は正解データの正例すべて (3,418) と、負例の一部 (16,000) を用いて学習したモデル (Full_data) の 10 試行における平均性能を示す。また、提案手法と AL.ST においてはエラーバーにより性能の標準偏差を示す。提案手法における各ステップでユーザがラベル付けしたデータ数の平均値は表 5 にまとめた。

AL.ST の結果は、2 ステップ目と 10 ステップ目を除くすべてにおいて他の戦略より下回る性能を示していることから、仮ラベルを単に追加しただけではかえって性能が低くなることを示している。また、この手法はモデルの学習が十分に行えない初期段階では、性能のばらつきが非常に大きいことがわかる。エラーバーの上端は AL の性能の平均値を超えていることから、仮ラベルとして学習に良いデータを追加できれば、通常の能動学習よりも性能を改善できる可能性を示唆している。しかし、追加するデータが悪い場合は精度が極端に落ちることもわかる。

一方、提案手法においては、早い段階からばらつきが少なくなるとともに、いずれのステップにおいても AL より性能を向上できている。したがって、自己学習における提案した仮ラベルの選出は有効であると考えられる。また、図 2 中 Full_data における性能の 90% (F 値 0.7) を達成しているステップ数に着目すると、提案手法は 5 ステップ目 (ラベル付け数は 800)、AL は 7 ステップ目 (1,200)、AL.ST は 10 ステップ目 (1,800) となる。これより、提案手法は AL、AL.ST に対し、それぞれおよそ 35% と 56% のラベル作成コストを削減できている。さらに、10 ステップ目の性能において提案手法は、Full_data と同等の性能を示している。表 5 より、このとき提案手法で要した正例の数は 629 (およそ 18%) であったことから、本手法はユーザの負荷を軽減できる。

5. おわりに

本研究では、ラベル付きデータが少ない状況を想定し、少数のラベル付きデータから効果的に CNN を学習するため、半自己学習手法を提案した。提案手法では、要所でのユーザの介入によるラベル付きデータの作成と、self-training

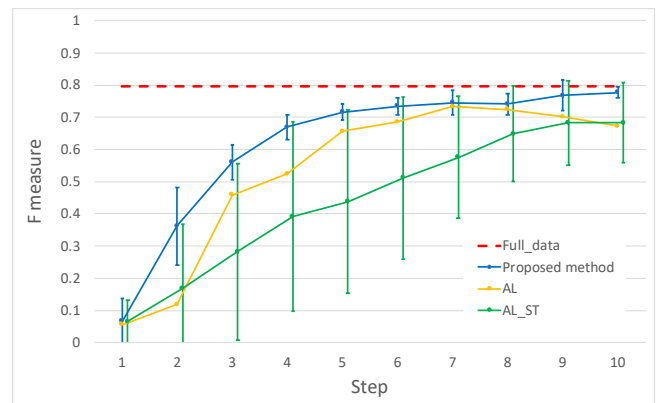


図 2 各ステップにおける F 値 (10 試行平均値) の推移。

に基づいた自己学習により CNN の学習効率を高める。評価実験として、Landsat8 号により撮影された衛星画像を用いたゴルフ場の検出タスクを実施し、従来手法との比較を行った。その結果、提案手法は従来の能動学習と比べて早く性能を向上でき、さらに、すべてのデータを用いて学習した CNN と同等の性能を、約 18% のデータ量で達成できた。

謝辞 本研究の実施にあたり、産業技術総合研究所 人工知能研究センター 地理情報科学研究チーム 研究チーム長の中村良介氏よりご助言いただきました。また、この成果は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務の結果得られたものです。深く感謝致します。

参考文献

- [1] Durieux, L., Lagabriele, E. and Nelson, A.: A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data, *Journal of Photogrammetry and Remote Sensing*, Vol. 63, No. 4, pp. 399–408 (2008).
- [2] Chen, G., Hay, G. J., Carvalho, L. M. T. and Wulder, M. A.: Object-based change detection, *International Journal of Remote Sensing*, Vol. 33, pp. 4434–4457 (2012).
- [3] Fujita, A., Sakurada, K., Imaizumi, T., Ito, R., Hikosaka, S. and Nakamura, R.: Damage detection from aerial images via convolutional neural networks, *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)* (2017).
- [4] LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, *Nature*, Vol. 521, pp. 436–444 (2015).
- [5] Rawat, W. and Wang, Z.: Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review, *Neural Computation*, Vol. 29, No. 9, pp. 2352–2449 (2017).
- [6] Ishii, T., Simo-Serra, E., Iizuka, S., Mochizuki, Y., Sugimoto, A., Ishikawa, H. and Nakamura, R.: Detection by Classification of Buildings in Multispectral Satellite Imagery, *International Conference on Pattern Recognition (ICPR)* (2016).
- [7] Zhong, Y., Fei, F., Liu, Y., Zhao, B., Jiao, H. and Zhang, L.: SatCNN: satellite image dataset classification using agile convolutional neural networks, *Remote Sensing*

表 4 各学習段階における仮ラベル数 (10 試行の平均値).

ステップ	1	2	3	4	5	6	7	8	9	10
仮正例数	-	43	92	320	810	1,816	2,571	2,812	2,904	2,909
仮負例数	-	1,077	3,819	6,576	9,818	14,641	19,292	23,679	28,039	33,333

表 5 提案手法における各学習段階のラベル数 (10 試行の平均値).

ステップ	1	2	3	4	5	6	7	8	9	10
正例数	20	28	80	159	249	330	409	482	557	629
負例数	400	592	740	861	971	1,090	1,211	1,338	1,463	1,590

- Letters*, Vol. 8, No. 2, pp. 136–145 (2017).
- [8] Deng, Z., Sun, H., Zhou, S., Zhao, J., Lei, L. and Zou, H.: Multi-scale object detection in remote sensing imagery with convolutional neural networks, *Journal of Photogrammetry and Remote Sensing*, Vol. 145, pp. 3–22 (2018).
- [9] Zhu, X.: Semi-Supervised Learning Literature Survey, Technical Report 1530, Computer Sciences, University of Wisconsin-Madison (2005).
- [10] Settles, B.: Active Learning Literature Survey, Technical Report 1648, University of Wisconsin-Madison (2010).
- [11] Wang, K., Zhang, D., Li, Y., Zhang, R. and Lin, L.: Cost-Effective Active Learning for Deep Image Classification, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 27, No. 12, pp. 2591–2600 (2017).
- [12] Tuia, D., Volpi, M., Copa, L., Kanevski, M. and Muñoz-Mari, J.: A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 3, pp. 606–617 (2011).
- [13] Gu, Q., Zhu, L. and Cai, Z.: Evaluation Measures of the Classification Performance of Imbalanced Data Sets, *International Symposium on Intelligence Computation and Applications, Computational Intelligence and Intelligent Systems*, pp. 461–471 (2009).