

## 平成 30 年度プロジェクト研究実績報告書

【研究課題名】	空間情報分野における AI を応用し作成した画像分類結果の評価手法の開発
【研究代表者】	富田 瑞樹（東京情報大学・准教授）
【研究分担者】	原 慶太郎（東京情報大学・教授） 平山 英毅（東京情報大学・総合情報学研究科後期課程 2 年）

### 【研究の目的】

衛星画像から土地被覆分類図を作成する研究は数多くなされてきた。近年では、機械学習を分類処理に用いることで、従来の手法と比べると高精度な分類器の作成が可能となってきた。また、より高精度の分類器を選択する際には、Kappa 係数や、Kappa 係数を用いた Z 検定による分類器の比較がなされてきている。しかし、Z 検定による比較では、Kappa 係数の差が非常に小さい場合は分類器間の有意差を確認することが困難であった。さらに、分類器間の Kappa 係数が非常に近い場合であっても、衛星画像を土地被覆分類した結果を確認すると相違点がある。たとえば、異なる分類器で処理した同じ衛星画像のそれぞれの分類結果の Kappa 係数がほぼ同じ値を示しても、ある土地被覆パッチの空間分布は異なることが多い。すなわち、機械学習によって、高精度な分類図を作成できるようになったものの、実際の土地被覆パッチの分布を再現可能な、より高精度な分類器の選択に関しては課題が残る。また、分類器の比較には、同じ検証データが用いられることが多く、これは、統計学的には対応のあるデータである。したがって、対応のあるデータを対象とした McNemar 検定を分類結果の差の判定に用いることが望ましい。そこで本研究では、同一の衛星画像を異なる分類器を用いて処理した土地被覆分類図間の分類精度の差を、Z 検定と McNemar 検定を用いて判定し、両者の結果を比較することを目的とした。

### 【研究報告】

#### 1. データと方法

##### 1.1. 研究全体の流れ

6 つの機械学習手法により各分類器を作成した。すべての分類器の組み合わせについて、Z 検定と McNemar 検定による比較をそれぞれ実施した（図 1）。

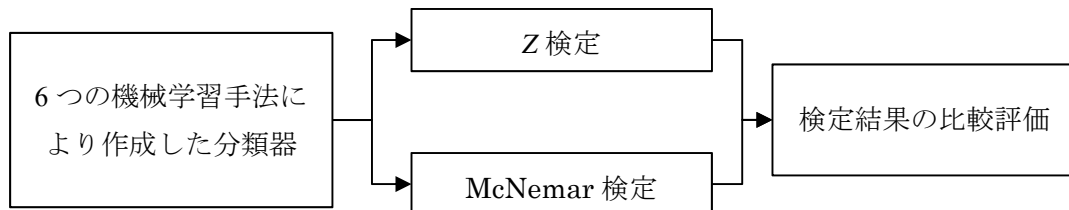


図 1. 研究全体の流れ図

## 1.2. 分類器の作成

6つの機械学習手法 (Random forests (RF), Bagging (BAG), eXtreme Gradient Boosting (XGB), Support vector machine (SVM), Feed-forward neural networks (NNET), K-nearest neighbor (KNN)) により各分類器を作成した。なお、各分類器のパラメータには、グリッドサーチによって Kappa 係数が最高値となる組み合わせを選択した。

解析に用いたデータは、RapidEye 衛星により得られた画像と、国土地理院が提供する基盤地図情報 5 m DEM, さらに、当該 DEM から計算した傾斜度と傾斜率である。教師データと検証データには、複数回の現地調査と Google Earth による情報から整備したグランドトゥルスデータを用いた。グランドトゥルスデータとして 6つの分類クラス (森林, 低木/草地, 耕作地, 都市域, 水域, 裸地) を設定し、分類クラスごとに 2076 ポイントを整備した。検証には合計 1495 ポイントのデータを用いた。

## 1.3. Z検定

Z 検定は、標本の平均と母集団の平均とが統計学的にみて有意に異なるかどうかを検定する手法である。本研究では、各分類器の分類結果をもとに Kappa 係数を計算した。さらに、分類器のすべての組み合わせについて、Kappa 係数を用いた Z 検定により Z 値を求め、 $p$  値を計算した。多重比較には Bonferroni 法を用いた。

## 1.4. McNemar 検定

McNemar 検定は、対応のある 2 組の標本の比率の差を検定する手法である。本研究では、比較する 2 つの分類器 A と分類器 B について、検証データに対する一致・不一致の数を集計した。そして、分類器 A のみが一致した数と分類器 B のみが一致した数について、カイ 2 乗値を求めた。最後に、Z 検定の結果と比較するため、カイ 2 乗値から  $p$  値を計算した。多重比較には、Z 検定と同様に、Bonferroni 法を用いた。

## 2. 結果と考察

### 2.1. Kappa 係数と Z 検定による分類結果の比較

XGB, RF, BAG の Kappa 係数は非常に近く、その差は最大で 0.023 であった。また、RF, BAG および XGB の組み合わせには Z 検定による有意差はみられなかった ( $p > 0.003$ , Bonferroni の補正)。一方で、SVM, NNET, KNN は、Kappa 係数の値に最大で 0.124 の差があり、分類器間に有意差がみられた ( $p < 0.003$ , Bonferroni の補正)。近年の機械学習手法の進歩により、非常に高精度かつ近い精度の分類器を作成できるようになってきた状況から、Z 検定を分類結果の比較に用いることは適切ではないと考えられる。

表 1. 分類器ごとの Kappa 係数 ( $\kappa$ )。

	RF	BAG	XGB	SVM	NNET	KNN
$\kappa$	0.961	0.943	0.966	0.861	0.737	0.794

表 2. 分類器ごとの Z 検定の結果.

	RF	BAG	XGB	SVM	NNET	KNN
RF	-	0.03	0.28	$1.6 \times 10^{-16}$	$9.8 \times 10^{-52}$	$1.0 \times 10^{-33}$
BAG		-	0.01	$7.1 \times 10^{-11}$	$1.6 \times 10^{-41}$	$1.6 \times 10^{-25}$
XGB			-	$8.6 \times 10^{-18}$	$8.6 \times 10^{-55}$	$3.0 \times 10^{-36}$
SVM				-	$4.8 \times 10^{-13}$	$2.8 \times 10^{-5}$
NNET					-	$9.6 \times 10^{-4}$
KNN						-

### 2.2. McNemar 検定により有意差が認められた要因

Z 検定では分類器間に差が見られなかった RF と BAG, BAG と XGB の組み合わせについて McNemar 検定を用いたところ有意差が認められた ( $p < 0.003$ , Bonferroni の補正). この要因としては, 対応のあるデータに用いる適切な検定手法を実施したことにより, 分類器による分類結果の差を検出できたためだと考えられる. また, 分類器間に差が認められなかった RF と XGB の組み合わせについては, 土地被覆分類結果のパッチの分布がほぼ同様であることが確認できた (図 2). すなわち, 対応のあるデータを用いた分類結果の比較には, McNemar 検定を用いることが適していると示唆された.

表 3. 分類器ごとの McNemar 検定の結果.

	RF	BAG	XGB	SVM	NNET	KNN
RF	-	$2.5 \times 10^{-4}$	$3.1 \times 10^{-1}$	$2.7 \times 10^{-21}$	$7.0 \times 10^{-53}$	$1.1 \times 10^{-39}$
BAG		-	$2.9 \times 10^{-4}$	$9.8 \times 10^{-14}$	$5.0 \times 10^{-43}$	$1.6 \times 10^{-31}$
XGB			-	$2.5 \times 10^{-23}$	$3.6 \times 10^{-53}$	$2.2 \times 10^{-40}$
SVM				-	$3.4 \times 10^{-20}$	$1.2 \times 10^{-8}$
NNET					-	$3.0 \times 10^{-4}$
KNN						-

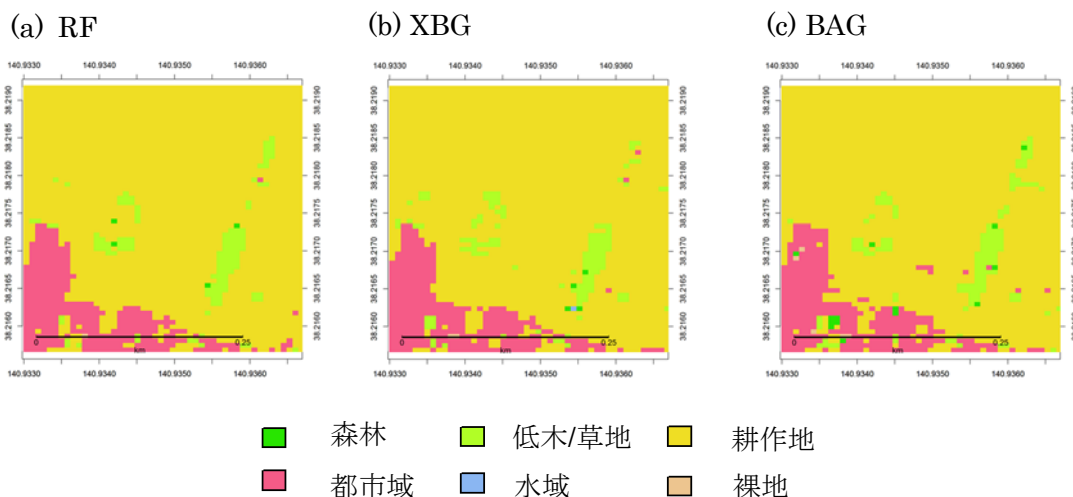


図 2. 土地被覆分類図の拡大図.

### 3. 結論

本研究では、複数の機械学習による分類器を作成し、分類器間の差について、 $Z$  検定と McNemar 検定を用いて比較した。その結果、同じ検証データを用いた分類結果の差の判定には、McNemar 検定を用いることが適していると示唆された。

#### 【成果の公表】

国際会議プロシーディング

Hidetake Hirayama, Mizuki Tomita, and Keitarou Hara. 2018. An Examination of Evaluation Method for the Classifiers with A Very High Accuracy in Land Cover Classifiers. Proceedings of the 39th ACRS (Asian Conference on Remote Sensing) Vol.5: 3155-3159.

国際会議

Hidetake Hirayama, Mizuki Tomita, and Keitarou Hara. 2018. An Examination of Evaluation Method for the Classifiers with A Very High Accuracy in Land Cover Classifiers. The 39th ACRS (Asian Conference on Remote Sensing) Kuala Lumpur, Malaysia.

国内学会

平山英毅・富田瑞樹・原慶太郎. 2018. 機械学習を用いた高精度な分類器により作成した土地被覆分類図の評価方法の検討, 平成 30 年度 写真測量学会, アオーレ長岡.