

5K-07

衛星画像の時系列データを用いた再帰的土地被覆分類手法の検討

大屋 優[†] 金盛 克俊[‡] 大和田 勇人[‡]

東京理科大学大学院 理工学研究科[†] 東京理科大学 理工学部[‡]

1. 序論

多くの環境問題で土地利用情報が必要とされている。現在、リモートセンシングデータに対して土地被覆分類として教師付き分類を行い、どこに何があるのか明確にすることができる。しかしながら、既存の手法は高精度ではあるものの多くの教師データを必要としており[1]、詳細の不明な未利用森林を対象とするのに不適切である。

そこで本研究では、既存の手法を用いる前段階として、僅かな教師データから土地被覆分類を行い、その分類結果から正解である可能性の高い結果だけを探し推薦することを目的とする。

従来、土地被覆分類はデータの入出力が1対1であるが[2]、本研究では、異なる時系列データに注目することで同じ教師データから複数の異なる分類結果を得ることと、それら結果を教師データとみなして複数回分類することを再帰的に行うことで、分類結果が正しいと推定した結果のみを抽出していく。

2. 提案手法

本提案手法は以下のように2段階で構成することにより対象地の被覆状態を分類する。

1. 初回時、季節ごとに最尤法で分類し、かつ、バンドリングデータをSVMで分類する。
2. 以降、バンドリングデータのみを用いて最尤法とSVMによる分類を行う。

図1は本提案手法の流れを示している。ここで、本研究では時系列データとして被覆情報の差が一番大きい組み合わせであろう春夏秋冬の4つを利用する。なおバンドリングデータとは各季節の衛星データを束ね1つのデータセットとみなしたものである。

2.1 初回時の季節ごとの分類

四季データごとに最尤法によって土地被覆分類を行う。本研究では状態の不明な未利用森林を対象とするため、教師データは分類クラスにつき1つのみとする。ここで衛星画像とは、物質の持つ光の反射特性を利用し、複数のバンド（種類）に分光されたスペクトルにおける反射率を示したデータのことである。そのため衛星画像は人間の目や衛星写真といったRGBカメラでは見えない現象を可視化することが可能となる。なお、最尤法はある地点の分光反射特性がどの分類クラスの特性と近似しているか定量的に比較する土地被覆分類である。

"Investigation of a recursive land cover classification with time series remote sensing imagery"

[†] Yu Ohya, Tokyo University of Science

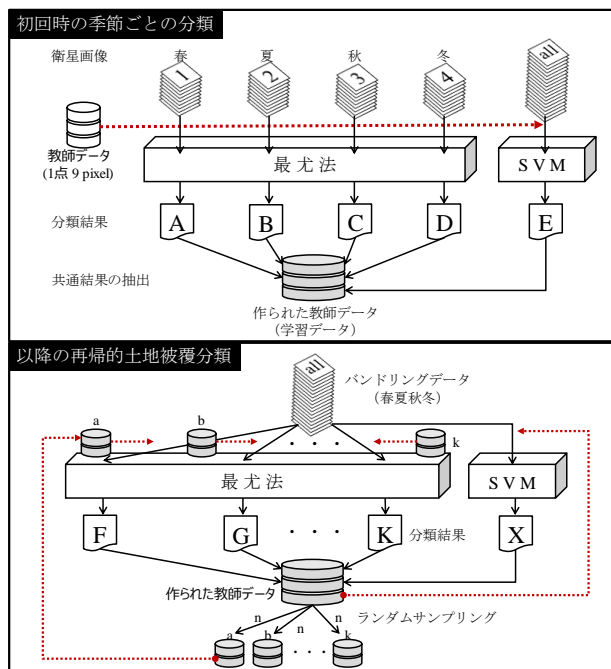


図1 提案手法の流れ

本研究では最尤法によって得られる4つの分類結果に加え、バンドリングデータをSVMで分類した結果も用いて、5つ全ての結果に共通した地点を探し出す。そして、その地点は正しい分類であった可能性が高いと想定し、それを次の分類の教師データとみなす。

2.2 バンドリングデータの再帰的分類

続けて、先ほど正しいと仮定し教師とみなしたものの一部を用いて、バンドリングデータを分類する。本手法は学習する教師データの取り方次第で何回でも分類が行えるため、本研究では、教師データをN個ずつランダムに抽出し、その抽出できた回数だけ分類を行う。ここで最尤法はSVMに比べ計算量が少ないため、この試行に適している。ただし計算の都合上、教師データの数Nは特徴の数より多くなるよう設定する必要がある。一方、計算時間を要するが分類精度の高いSVMには全ての教師データを与えて1回だけ分類を行う。この最尤法とSVMによる複数の分類から共通した判別結果を持つ場所の探索を行い、新たな教師データを作成する。

本手法の利点として、この教師データの作成と分類の作業を再帰的に行うことができることがあげられる。本研究では作成した教師データの個数が収束するまでに行い、最後にSVMで各教師データがその分類結果で正しいかどうか判断し、精査する。

3. 実験結果

本提案手法のもと実験を行った。なお、本研究では対象地域として自然豊かである鎌倉市を選択した。衛星画像には Landsat 8 を用いた。また詳細な植生状況を知る必要があると想定し、分類クラスをスギ、ヒノキなど 12 種類と細かく設定した。ここで、教師データは各クラスにつき 1 点である一方、最尤法は教師データ数が衛星画像の特徴数より多い制約を有するため、初回では教師データの周囲も同一クラスとみなし、9 pixel 分の衛星画像データから学習を行った。なお特徴数も 9 つが限度であるため Landsat 8 のバンド 1~5, 7, 10, 11 を選択し利用した。

以上の条件のもと検証を行った。その結果として表 1 は初回時の季節ごとの分類精度と、それら分類にて共通した結果を持った地点だけに着目した時の精度を表している。表 2 ではそれ以降の分類で再帰するごとに正解と推定された結果（“作られた教師データ”）に対する精度を示している。さらに、最終的に正しいと推薦した結果と先行研究の比較を行い、表 3 に記した。なお、精度には Kappa 係数を用いた。

4. 考察

表 1 より、季節ごとの分類結果では Kappa 係数が 0.38 から 0.51 と幅がある上、全体的に正解率が低いことがわかる。しかし、4 つの分類結果から正解の可能性のある場所を探したところ、21% のエリア（判別結果）が推薦され、そのうちの 80% が実際に正解だった。そこに SVM の結果も参照することで推薦されたエリアの正解率は大幅に上昇した。

表 2 より、以降の再帰的分类では Kappa 係数 0.65 から 0.71 の揺れ幅の中で徐々に正解データ数を増やしている。本研究では、3 回目と早い段階で収束してしまっただけ、3 回目で正解の可能性のある結果に対してのみ SVM で再検討してみたところ、対象エリアの 36.9% に対して分類結果は正解であると本手法は推薦した。実際にその推薦エリアの分類結果は正解率 90% となり、正しく分類できた場所の推薦を行えたといえる。本研究では全体の分類結果でなく、分類したものの中から正解の可能性のある結果のみを取り上げたため正確な比較はできないが、鎌形ら[09] と比べて空間分解能 (1 pixel あたりに被覆する土地面積) が粗く分類クラス数も多いが、本研究の方が高精度であることが表 3 からわかる。

すでに、ハイパースペクトルデータと呼ばれる高解像度 (空間分解能、波長分解能が比べ物にならないほど細か) な衛星画像に対して高精度な機械学習 (SVM など) で分類する研究は多くなされている。しかしながら、そのほとんどが対象地の半数近くを教師データとして学習しているため、詳細の不明な未利用植生に対して実用的ではなかった。そこで、本研究のように 1 箇所という僅かな教師データからでも時系列データを活用することで、分類を何回も

表1 初回時の分類結果

	κ	Accuracy	CR*	
最尤法	春	0.380	0.464	---
	夏	0.514	0.608	---
	秋	0.486	0.582	---
	冬	0.393	0.482	---
Common**	0.679	0.813	21.1%	
Common & SVM	0.756	0.896	16.6%	

* CR は “作られた教師データ” の対象地域に対する被覆割合 (Cover Ratio)

** Common は最尤法による4つの分類から結果の共通した部分だけに着目した時の精度

表2 再帰ごとの作られた教師データの精度

	1st	2nd	3rd	Last*
κ	0.712	0.647	0.665	0.709
Accuracy	0.863	0.812	0.813	0.896
CR**	33.6%	45.5%	45.5%	36.9%

* Last は3回目で作成された教師データで学習したSVMでその教師データが正解と推定した結果のみの精度

表3 先行研究との比較

	鎌形ら [09]	本研究
対象地	鎌倉市	鎌倉市
地域	山間部	平野部
面積	25km ²	36.9% of 102,724km ²
衛星画像	IKONOS	Landsat 8
空間分解能	4m	30m
分類クラス数	7	12
Kappa係数	0.526	0.709

試行し、正解地点の推薦をし、この分類と推薦を再帰的に行うことで良質な教師データを作成した。今後、既存手法に対してこの作成したデータを学習し、未利用森林への土地被覆分類の前処理として有効であることを示す必要がある。

5. 結論

本研究では、時系列データに注目することで1つの教師データから複数回分類を行い、それら分類結果から正解の可能性のあるデータを探し、それを新たな教師データと見なして再度複数回分類を行う手法を提案した。その結果、分類結果の中から正解と思われるものだけを抽出した。

現在、自然再生エネルギーの利用などで未利用植生が注目されている。本研究で提案した手法は詳細の不明な森林の土地被覆分類に対して高いポテンシャルを有している。

参考文献

- [1] Pal M. and Mather P.M. "Support vector machines for classification in remote sensing", *International Journal of Remote Sensing*, Vol.26, No.5, pp.1007-1011, 2005
- [2] Yu Ohya, Katsutoshi Kanamori and Hayato Ohwada, "Estimating Resource Distribution Using Satellite Images to Utilize Woody Biomass", *Lecture Notes on Information Theory*, Vol.2, No.2, pp.158-162, 2014