

畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定 — ファインチューニングの利用による精度の向上 —

小川京太 長名優子

東京工科大学 コンピュータサイエンス学部

1 はじめに

近年、画像をキーとして検索を行う類似画像検索に関する研究が盛んに行われているが、類似画像検索において風景画像中に建物などの人工物が含まれていると検索精度が低下してしまうことが知られている。また、それに対し、画像中の人工物が含まれる領域を特定し、その領域を考慮することで検索精度が向上することもわかっている [1]。それに対し、画像において人工物が含まれる領域を特定する手法がいくつか提案されている。ニューラルネットワークの一種である RBF ネットワーク [2] やサポートベクトルマシン [3] を用いて色情報と直線情報、テクスチャ情報を考慮して風景画像中の人工物領域の特定を行う手法では、評価指標である F 値が 0.778 程度の精度で人工物領域の特定が行えることが分かっている [4][5]。また、畳み込みニューラルネットワーク [6] を用いた人工物領域の特定 [7] では、 F 値が 0.806 程度の精度で人工物領域の特定が実現できることが分かっている。さらにこの手法では RBF ネットワークやサポートベクトルマシンを用いた手法に比べ、特定物の認識に関する精度も向上することが確認されている。畳み込みニューラルネットワークを用いた手法では、RBF ネットワークやサポートベクトルマシンを用いた手法よりも精度が若干向上してはいるが、大幅に向上してはいない。これは、訓練データとして従来手法 [4][5] と同じデータを用いたため、訓練データ数が少なかったことが原因として考えられる。また、従来手法と同じ条件で実験を行うために特定できる対象が 12 種類と少ないことも問題としてあげられる。いずれの問題に対してもデータ数や特定できるカテゴリ数を増やすことで解決できると考えられるが、データを増やすのは非常に大変であり、コストもかかる。

それに対し、ファインチューニング [8] という手法が提案されている。ファインチューニングは別の課題を学習済みのネットワークモデルを初期値として再学

習を行う手法で、ファインチューニングを利用すると十分な訓練データを用意することができない場合でも高い精度を実現できることが知られている。

本研究では、畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定においてファインチューニングの利用する手法を提案する。

2 畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定

ここでは、従来の畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定 [7] について説明する。この手法では、 K 平均法を用いて画像を領域分割し、領域ごとに人工物が含まれているかの判定を行う。

2.1 K 平均法による領域分割

画像の領域分割は、 K 平均法を用いて行う。 K 平均法を用いて画像の領域分割を行う際は、データとして画素ごとの色 (HSV 色空間における座標) と座標の情報を用いる。 K 平均法で領域分割を行うと、分割される領域の数は K の値に依存して決まる。この手法では、領域ごとに人工物が含まれているかどうかの判定を行うため、1つの領域内に複数のオブジェクトが含まれているような状況は望ましくない。適切な領域の数は画像によって異なるが、文献 [7] では、1つの領域内に複数のオブジェクトが含まれないように K の値は大きめに設定されている。また、 K 平均法を用いて領域分割を行うと、画素の座標を考慮して分割を行ったとしても不連続な領域が生成されてしまうことがある。不連続な領域はそれぞれ別の領域として扱う必要があるため、領域分割後に領域のラベリングを行い、同じクラスに分類されても不連続な領域は別の領域として扱うものとしている。また、領域のサイズが小さすぎると人工物が含まれているかどうかの判定が正しく行えない可能性があるため、面積が一定以下の領域 (微小領域) に対して、隣接した領域と色が類似していれば統合する処理を行う。

Identification of artificial areas in landscape images using convolutional neural networks — Improving accuracy by using fine tuning —
Kyota Ogawa and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

2.2 各領域に対する畳み込みニューラルネットワークを用いた人工物領域の判定

画像に対して K 平均法によって領域分割を行い、認識を行いたい1つの領域以外の領域を黒で塗りつぶした画像を畳み込みニューラルネットワークへ入力することで人工物領域であるかの判定を行う。文献 [7] では、出力層が建物、道路、その他の人工物、空、曇、木、草、花、水、岩、砂、その他の自然物の12クラスに対応するニューロンから構成される畳み込みネットワークを用いており、建物、道路、その他の人工物のいずれかの分類された場合には人工物領域であると判断されることになる。

3 ファインチューニングを利用した畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定

提案するファインチューニングを利用した畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定では、従来の畳み込みニューラルネットワークを用いた風景画像中の人工物領域の特定 [7] と基本的には同様の方法で人工物領域の判定を行う。なお、畳み込みニューラルネットワークとしては画像分類を学習済みのモデルを用い、ファインチューニングを利用することで少ない学習データでの精度の向上を目指す。

本研究では学習済みのネットワークモデルとして VGG16[9] という畳み込みニューラルネットワークを用いることを考えている。VGG16 は画像データベース ImageNet を利用して100万を超える画像を学習された16層の畳み込みニューラルネットワークであり、1000個のオブジェクトカテゴリに分類することができる。

また、畳み込みニューラルネットワークの出力としては以下のような3つを考える。

(1) 自然物か人工物かを判別

入力された領域が自然物か人工物かの2クラス分類を行う。

(2) 12のオブジェクトカテゴリに分類した上で判別

文献 [7] と同様の12のオブジェクトカテゴリに分類し、その結果に基づいて人工物か自然物かを判定する。

(3) VGG16で分類されている1000個のオブジェクトカテゴリを参考により詳細に分類した上で人工物と自然物に判別

ファインチューニングに利用する VGG16 で分類されている1000個のオブジェクトカテゴリを参考により詳細な分類をした上で人工物か自然物を判定する。

4 計算機実験

計算機実験を行い、(1)~(3)の場合に対し、それぞれ0.883, 0.982, 0.361の F 値が得られることを確認した。(3)に対しては精度が低くなっているが、従来手法で行っていた(2)に関しては高い精度で判定が行えることが分かった。

参考文献

- [1] 一戸竜司, 長名優子: “不応性を有する自己組織化特徴マップによる類似画像検索 一色の特徴量の次元の調整による検索精度の向上 —,” 情報処理学会第80回全国大会, 2018.
- [2] J. Moody and C. J. Darken: “Fast learning in networks of locally turned processing units,” *Neural Computation*, Vol.1, pp.281–294, 1989.
- [3] 阿部重夫: *パターン認識のためのサポートベクトルマシン入門*, 森北出版, 2011.
- [4] 藤崎豪人, 長名優子: “RBF ネットワークを用いた風景画像中の人工物領域の判定における — SIFT, LBP の利用による精度の向上 —,” 情報処理学会第75回全国大会, 2014.
- [5] 福田孝真, 長名優子: “サポートベクトルマシンを用いた風景画像中の人工物領域の判定,” 情報処理学会第77回全国大会, 2016.
- [6] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, Vol.86, No.11, pp.2278–2324, 1998.
- [7] 砂川皓, 長名優子: “ディープラーニングを用いた風景画像中の人工物領域の特定,” 情報処理学会第79回全国大会, 2017.
- [8] 内田祐介, 山下隆義 “畳み込みニューラルネットワークの研究動向,” *パターン認識・メディア理解研究所*, 2017.
- [9] K. Simonyan and A. Zisserman: “Very deep convolutional networks for large-scale image recognition,” *Proceedings of International Conference on Learning Representation*, 2015.