6R-02

ファインチューニングを用いた畳み込みニューラルネットワークによる人工物領域の特定精度の向上

山田創平　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

画像を検索キーとして検索を行う類似画像検索では、色や形状などの特徴の類似性に着目して検索を行うことで自然物のみで構成された風景画像に対しては高い精度で検索を行うことができる。しかしながら、色の特徴に着目した類似画像検索では、風景の中に建物などの人工物が存在している場合に検索精度が低下してしまうことがある。これは、自然物では色とオブジェクトが深く関わっているものが多いのに対し、人工的では同じ建物でも様々な色のものがあるように自然物と人工物では含まれる色から得られる情報に違いがあるためであると考えられる。それに対し、画像中で人工物が含まれる領域を特定し、その情報も用いて検索を行うことで検索精度が向上することも知られている [1]。

風景画像中で人工物が含まれる領域を特定する手法の1つとして、畳み込みニューラルネットワーク [2] を用いる手法 [3] が提案されている。この手法では、K 平均法 [4] を用いて領域分割を行い、各領域に人工物が含まれるかを畳み込みニューラルネットワークを用いて判定することで、約93%の精度で人工物領域の特定を実現している。この手法では、RBF ネットワークを用いて人工物領域の特定を行う手法 [1] で用いられているのと同じデータを用いて学習を行い、文献 [3] の手法よりも高い精度で人工物領域の特定を行うことができている。しかし、学習に用いたデータの数は十分とは言えない。学習に用いるデータを増やすことで特定精度を改善できる可能性が高いが、学習データを増やすためには領域ごとにクラスラベルを付与する必要があり、設計者にかかる負荷が大きい。

本研究では、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特定においてファインチューニング [5] を用いることで特定精度の向上を図る。

Improving Accuracy of Artifact Region Identification using Convolutional Neural Networks with Fine Tuning
Sohei Yamada and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

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

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 [1] では、画像を領域に分割し、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 [4] 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と座標の情報を用いる。画素 i に関するデータ $\boldsymbol{a}^i$ は

$$\boldsymbol{a}^i = \left(x_i, y_i, z_i, \frac{X_i}{W}, \frac{Y_i}{H}\right) \tag{1}$$

で与えられる。ここで、x_i, y_i, z_i は画素 i の表す色の HSV 色空間における座標である。また、X_i, Y_i は画素 i の座標、W, H は画像の横と縦の画素数を表している。

K 平均法で領域分割を行うと、画像は K 個の領域に分割されることになる。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 [3] では、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が起こらないように K の値は大きめに設定している。

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。この手法では、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、同じクラスタに分類されてい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る可能性があるため、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ようにしている。

Copyright ©2020 Information Processing Society of Japan.
All Rights Reserved.

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があるかの判定

畳み込みニューラルネットワークに認識対象となる領域以外を塗りつぶした図2の右のような画像を入力として与え、その領域がどのクラスに属するかを出力するように学習を行う。クラスとしては、建物、道路、その他の人工物、空、雲、木、草、花、水、岩、砂、その他の自然物の12のクラスを用いている。

3 ファインチューニング

他のデータにおいて学習されたネットワークの情報を利用して新しいモデルを構築する手法として、転移学習やファインチューニングなどの手法が提案されている。転移学習では、既存の学習済みのモデルの出力層以外の部分の重みを変更せずに特徴抽出器として利用し、出力層の部分のみを学習することで新しいモデルを構築する [6]。それに対し、ファインチューニングでは、学習済みのモデルの重みなどパラメータをコピーして、それを新しいニューラルネットワークのパラメータの初期値として使って学習を行う [5]。入力層に近い層の学習パラメータは学習に使用した画像の具体的な特徴が反映され、出力層に近い層のパラメータは分類を特定する特徴が反映されていると考えられる。また、中間付近のパラメータには抽象的な特徴が反映されていると考えられる。そのため、図3のように下位層の重みは固定しておき、上位層の重みのみを再学習させることが多い。

4 ファインチューニングを用い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

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 [3] において、ファインチューニングを用いることで人工物領域の特定精度の向上を目指す。学習済みのモデルとしては画像認識に用いられる ImageNet のデータ (VGG16 の学習済みモデル) を用いる。

5 計算機実験

提案手法を用いて人工物領域の判定を行い、従来の手法 [3] よりも高い精度で判定が行えることを確認した。

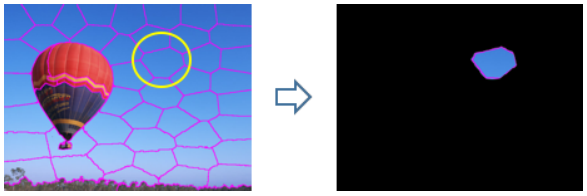

図 1: 領域分割画像と塗りつぶし画像

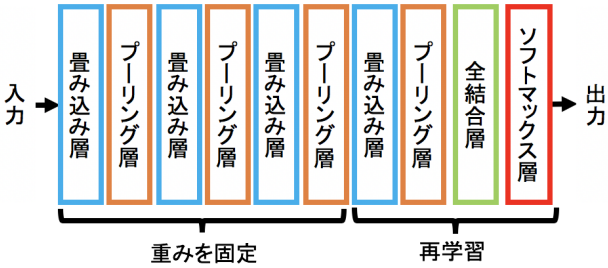

図 2: 重みの固定と再学習

参考文献

[1] 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.

[2] T. Lecun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and D. Jackel : “Backpropagation applied to hand-written zip code recognition,” Neural Computation, Vol.1, No.4, pp.541–551, 1989.

[3] 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.

[4] J. B. MacQueen : “Some methods for classification and analysis of multivariate observations,” Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol.1, pp.281–297, University of California Press, 1967.

[5] F. Chollet：“ファインチューニング, ” Python と Keras によるディープラーニング, マイナビブックス, pp.159–166, 2018.

[6] 中山英樹：“深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習,” 電子情報通信学会技術研究報告, Vol.115, No.146, pp.55–59, 2012.

Copyright ©2020 Information Processing Society of Japan.
All Rights Reserved.