

独立成分分析と深層学習による セマンティックセグメンテーション手法の検討

渡辺 圭祐[†] 鈴木 幸司[†]

[†]室蘭工業大学

1 はじめに

画像をピクセルごとにクラス分類するセマンティックセグメンテーションは、自動運転をはじめ、医療用画像での異常領域検出や農業への応用など、幅広い分野での活躍が見込まれている。セマンティックセグメンテーションに限らず、深層学習の分野全体は大きな発展を遂げているが、深層学習を使用するための前提として、大規模なデータセットが存在していることが必要となる。データセットが整っていない場合、十分な精度を得ることが難しくなるため、データの希少性などの理由で小規模なデータセットしか用意できない対象への深層学習の適用が困難になる。

本研究では独立成分分析 (ICA) によって画像の基底を抽出を行い、この基底と線形結合を用いることで任意の画像を生成する。ここで生成する画像を人工物、もしくは自然物とすることで線形結合の値を画像の特徴量とすることができる。この特徴量をニューラルネットワークに入力することで、小規模データセットでセマンティックセグメンテーションを行う分類器を作成した。また DeepLab v3+ を同様のデータを用いて学習を行い、分類器と提案手法で精度の比較を行った。その結果、提案手法での分類精度が DeepLab v3+ 精度を上回ることを確認した。

2 提案手法 [2]

2.1 ICA による基底作成

ICA とは複数の信号が重なって生成される混合信号から、個々の信号とその大きさを推定する技術である。この ICA を画像に適用することで、画像を構成する基底を無相関の状態で抽出することができる。この時、画像を人工物だけが写っているものと、自然物のみが写っているものの2つに分別することで、それぞれのパターンのみが反映された基底を抽出することができる。本研究では 255 個の基底を作成し、その中から固有値が高い順に 64 個の基底を使用する。

2.2 基底と結合係数による画像の復元

64 個ある基底を用いて画像ブロックを復元する場合、復元の式は以下ようになる。

$$x = a_1 s_1 + a_2 s_2 + \dots + a_{64} s_{64} \quad (1)$$

Semantic segmentation using ICA and deep learning

Keisuke Watanabe [†], Yukinori Suzuki [†]

[†]Muroran Institute of Technology

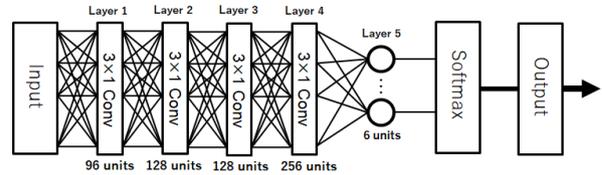


図 1: ニューラルネットワークの概略図

画像ブロックを x , n 番目の基底を s_n , また、基底にかかる結合係数を a_n とした。結合係数の値は復元対象とする画像ブロックごとに異なる値を持つため、画像ブロックが持つ特徴量として扱うことができる。

カラー画像を対象に復元を行う場合、複数チャンネル分の結合係数が作成される。今回は色空間を HSV に変更してから復元を行っているため、色相 (H)、彩度 (S)、明度 (V) の特徴量を持った結合係数がそれぞれ作成される。

2.3 ニューラルネットワーク

ICA 基底により画像を復元したときに得られる結合係数は 1×64 のベクトルデータである。これが人工物の基底から作成されたものと、自然物の基底から作成されたものとで 2 つ存在する。この 2 つのベクトルデータを水平方向に結合し、 1×128 のベクトルデータにしたのち、HSV の 3 チャンネル分の結合係数を順に垂直方向に重ねていき、 9×128 のデータを作成する。この 9×128 の結合係数のデータをニューラルネットワークに入力し、学習及び分類を行う。

ニューラルネットワークの構造は図 1 となっており、ソフトマックス層から「建物」、「道路」、「車両」、「植物」、「空」、「水面」の 6 つのカテゴリの確率を出力する。

3 分類器の作成

MATLAB を用いて、セマンティックセグメンテーションを行う分類器の作成を行う。まず画像を復元するための基底の作成は、人工物画像 15 枚と自然物画像 15 枚の計 30 枚の画像に対してオーバーラップありのブロック分割を行い、ICA を適用することで作成した。次に図 1 に示したニューラルネットワークの学習を行う。学習用の結合係数を作成するための画像は、人工物が 9 枚、自然物が 10 枚の計 19 枚である。基底作成と同様にオーバーラップありのブロック分割を行い、切り出した画像を対象にした復元を行って結合係数を得た。

今回のセマンティックセグメンテーションのタスクを人工物と自然物の2値分類に設定している。そのため、ニューラルネットワークから出力される6つの確率の中で最も高い確率を参照し、それが「建物」、「道路」、「車両」のいずれかである場合人工物、「植物」、「空」、「水面」のいずれかであった場合は自然物とした。

3.1 モルフォロジー

ニューラルネットワークによる分類後に発生する誤認識をモルフォロジーによって補完する。分類後は、人工物は1、自然物は0としたバイナリデータが出力される。このバイナリデータに対して、注目画素を中心とした 3×3 の近傍を読み取り、1が5つ以上あった場合は注目画素を1に、そうでない場合は注目画素を0にするという処理を行う。これにより、微小な誤認識を排除することができる。

4 実験内容

提案手法により作成した分類器と、既存のセマンティックセグメンテーション手法であるDeepLab v3+を用いて同じ画像に対して分類を行い、精度の比較を行う。

分類結果の精度検証には正解データとの一致率を表す「正解率」、「適合率」、「再現率」、二つの調和平均である「F値」を用いた。

4.1 DeepLab v3+[4]

DeepLab v3+はエンコーダ-デコーダ構造をとっており、エンコーダ部分では atrous 畳み込みによる特徴抽出と、低レベル特徴抽出の2種類の特徴の抽出を行う。本研究では、低レベル特徴作成部分には ResNet18[5]を用いた。

デコーダ部分ではエンコーダ部分で抽出した2種類の特徴を連結し、小さいフィルタでの畳み込みによって特徴の修正を行った後に、アップサンプリング処理を施してセマンティックセグメンテーションを行う。

DeepLab v3+のネットワーク構成はMATLABを用いて行った。また、学習データは、提案手法の分類器の学習に用いた19枚の画像を使用した。分類方法も提案手法に合わせて6つのカテゴリの確率を出力したのち、人工物と自然物に分類する。

5 実験結果と考察

分類を行った画像を図2に示す。図2の(c)と(d)が分類後の画像であり、画像中で赤色に着色された領域が人工物と予測されたピクセルであり、緑色に着色された領域が自然物と予測されたピクセルである。図2(c)を見てみると正解率78%、F値は75%ほどの分類結果となった。対して図2(d)では自然物領域が大部分を覆ってしまっており、正解率58%、F値46%となっている。分類精度をまとめた表1でも、提案手法がDeepLab v3+の精度を上回っていることが確認できる。表2に20枚の画像に対して分類を行った時の平均

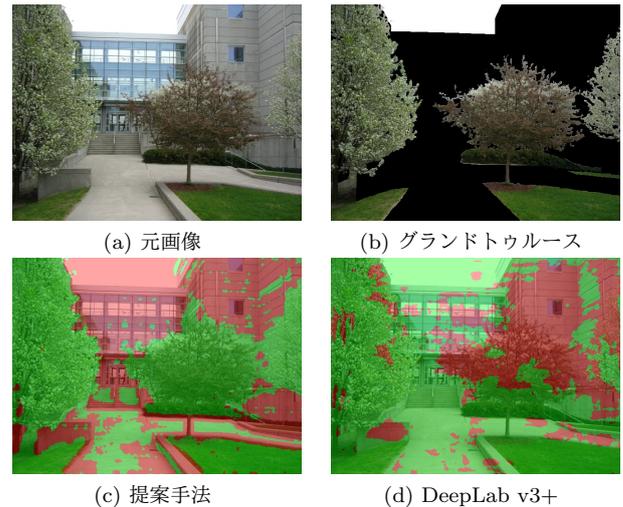


図2: 画像のセマンティックセグメンテーション

表1: 図2の分類精度の評価

分類方法	正解率	適合率	再現率	F値
提案手法	78.18%	79.97%	72.01%	75.78%
DeepLab v3+	58.16%	55.69%	39.85%	46.46%

結果を示す。提案手法において、主に分類精度を下げている要因となっていたのが、道路の部分に発生している自然物領域である。道路部分のニューラルネットワークからの出力を見てみると、「水」と判定されることが大半であった。この改善案としては学習用の画像の彩度、明度を変更することによるデータ拡張が挙げられる。これによりロバスト性の向上を行う。

表2: 20枚の画像での評価の平均

分類方法	正解率	適合率	再現率	F値
提案手法	75.10%	74.58%	64.62%	66.56%
DeepLab v3+	69.91%	70.39%	56.36%	58.77%

参考文献

- [1] 渡辺 圭祐, 鈴木 幸司, "独立成分分析と深層学習を用いた画像の人工物と自然物の分類", 令和元年度電気・電子関係学会北海道支部連合大会, pp.167, 2019.
- [2] A. Hyvarinen, J. Karhunen, E. Oja, "詳解 独立成分分析 信号解析の新しい世界", 根本幾, 川勝真喜訳, 東京電機大学出版局, 2005.
- [3] Olmos A, Kingdom F, "A biologically inspired algorithm for the recovery of shading and reflectance images", Perception, 33, pp.1463-1473, 2004.
- [4] Chen, Liang-Chieh et al. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.", The European Conference on Computer Vision, pp. 801-818, 2018.
- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition", The IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.