

畳み込みニューラルネットワークを用いた風景画像を対象とした類似画像検索

矢森大貴 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

画像検索は大きくキーワードによる検索と画像をキーとした検索に分けられる。本研究で扱う類似画像検索は、「犬」や「猫」といったキーワードではなく画像そのものを検索キーとして、類似画像を検索する方法である。そのような手法の1つとして畳み込みニューラルネットワーク [1] を用いた類似画像検索が提案されている。この手法では、事前に画像認識の学習を行った畳み込みニューラルネットワーク [2] を利用する転移学習という手法がとられている。畳み込みニューラルネットワークを用いた類似画像検索では、事前に画像認識の学習を行った畳み込みニューラルネットワークに画像を入力したときの出力層の1つ手前の層の出力をその画像の特徴を表す特徴ベクトルとして利用して検索が実現される。画像と特徴ベクトルの組をデータベースに保存しておき、検索キーとなる画像を畳み込みニューラルネットワークに入力して生成した検索キーに対応する特徴ベクトルをデータベース内の特徴ベクトルと比較することで検索を行う。特徴ベクトルが類似している画像が検索されることになる。しかし、画像分類の学習を行った畳み込みニューラルネットワークを利用しているため、特定のオブジェクトが含まれるような画像を検索することは得意であるが、風景画像を対象として位置などを含めて類似した画像を検索することは困難であるといった問題がある。

それに対し、不応性を有する自己組織化特徴マップを用いて類似画像検索を行う方法 [3] も提案されている。これは、風景画像を対象とした類似画像検索の手法であり、画像内に含まれるオブジェクトの位置なども考慮して検索を行うことができる。入力として、色情報、大きさ・形状情報、テクスチャ情報などを使用し、これらの特徴を画像内での位置も含めて考慮することができるようになってきている。しかしながら、事前に特徴量を抽出する必要のない畳み込みニューラルネットワークを用いた類似画像検索の手法とは異なり、事前に特徴量の抽出を行う必要がある。また、検索の精

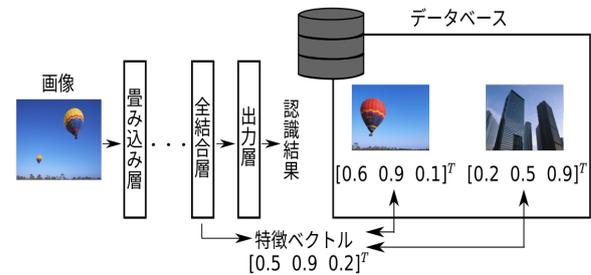


図 1: 畳み込みニューラルネットワークを用いた類似画像検索のイメージ

度も用いる特徴量に大きく依存している。

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を対象とした類似画像検索を提案する。提案手法では入力画像を複数の領域に分けて入力し、領域ごとに特徴ベクトルを生成することで画像内での大まかな位置情報を考慮した検索を目指す。

2 畳み込みニューラルネットワークを用いた類似画像検索

畳み込みニューラルネットワークを用いた類似画像検索では、検索用に新たに畳み込みニューラルネットワークの学習を行うのではなく、画像中に含まれるオブジェクトの認識を行うように学習を行った畳み込みニューラルネットワークを用いる転移学習が用いられる [4][5]。

畳み込みニューラルネットワークを用いた類似画像検索では、図 1 に示すように画像を入力したときの出力層の手前の全結合層の出力を特徴ベクトルとして用いる。

画像 p に対応する特徴ベクトル $\hat{c}^{(p)}$ は出力層の手前の全結合層の出力 $c^{(p)}$ を大きさが 1 になるように正規化したもので与えられる。生成した特徴ベクトルと画像を組み合わせたものはデータベースに保存しておく。検索を行う場合には、検索キーとして用いる画像を畳み込みニューラルネットワークに入力し、それに対する出力層の手前の全結合層の出力から特徴ベクトルを求める。求めた検索キーの特徴ベクトルとデータベースに記録してある特徴ベクトルを比較することで検索を行う。特徴ベクトルどうしのユークリッド距離

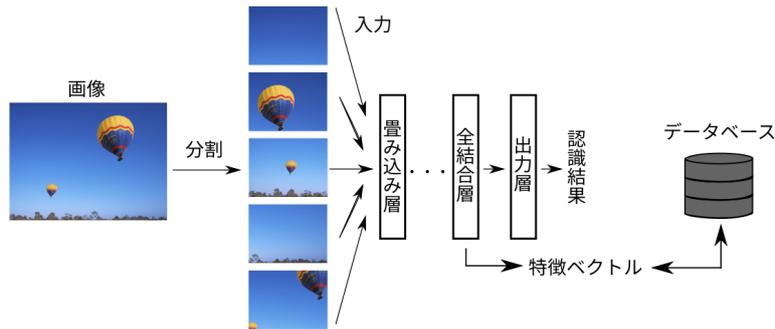


図 2: 提案手法による類似画像検索のイメージ

を求め、距離が近ければ類似していると判断し、検索結果として出力する。

画像 p の特徴ベクトル $\hat{c}^{(p)}$ と検索キーの特徴ベクトル \hat{c}^{key} とのユークリッド距離 $d(\hat{c}^{(p)}, \hat{c}^{key})$ は

$$d(\hat{c}^{(p)}, \hat{c}^{key}) = \sqrt{\sum_{n=1}^{N^{(L-1)}} (\hat{c}_n^{(p)} - \hat{c}_n^{key})^2} \quad (1)$$

で与えられる。ここで、 $N^{(L-1)}$ は出力層 (第 L 層) の直前の層のニューロン数である。

しかし、この手法で用いる畳み込みニューラルネットワークは画像認識用に学習されたものであるため、画像中のどこにそのオブジェクトが含まれていても同じように認識されるようになっている。そのため、検索に用いる場合にもオブジェクトの位置を考慮した検索を行うことはできないという問題がある。

3 畳み込みニューラルネットワークを用いた風景画像を対象とした類似画像検索

ここでは、提案する畳み込みニューラルネットワークを用いた風景画像を対象とした類似画像検索について説明する。提案手法では、従来の手法と同様に画像認識の学習を行った畳み込みニューラルネットワークを利用する。なお、画像認識の学習を行った畳み込みニューラルネットワークとしては VGG16[2] を用いることを考えている。VGG16 は Oxford 大学のチームが開発し、畳み込み層と全結合層が合わせて 16 層ある畳み込みニューラルネットワークである。VGG16 は ImageNet と呼ばれるデータセットで学習済みであり、画像認識でよく使われている。

従来の畳み込みニューラルネットワークを用いた類似画像検索手法とは違い、提案手法では図 2 に示すように画像を上下左右、中央の 5 つの領域に分けたものを入力として用いる。5 つの領域に分けて入力し、それぞれに対して出力層の手前の全結合層の出力から特徴ベクトルを求める。各領域に対する特徴ベクトルを

並べて 1 つにまとめたものを特徴ベクトルとして用いることで、大まかな位置情報を保持した特徴ベクトルの生成を可能としている。

検索は、従来手法と同様に特徴ベクトル間のユークリッド距離を求め、検索キーの特徴ベクトルと類似した特徴ベクトルと対応する画像を検索結果として出力することで行う。

4 計算機実験

550 枚の画像に対して実験を行い、 F 値 0.85 (再現率: 0.80, 適合率: 0.99) の精度で検索が行えることを確認した。

参考文献

- [1] 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.
- [2] K. Simonyan and A. Zisserman : “Very deep convolutional networks for large-scale image recognition,” Proceedings of International Conference on Learning Representation, 2015.
- [3] 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 – 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.
- [4] 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,” 電子情報通信学会音声研究会, 2015.
- [5] 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.