5ZB-06

機械学習を用いたランドマーク識別

庄司　將孝†　　秋岡　明香‡
明治大学総合数理学部　　明治大学

1. はじめに

ランドマーク画像は撮った際の時間や天気、季節など様々な要因により姿が変わる。近年、SNSの普及により、個人が写真を投稿することはより身近なものとなった。それに伴い、サンプル画像のような同一の姿のものだけでなく、様々な場面のランドマークを見る機会が増えた。しかし、Google や Yahoo!といった検索エンジンの類似画像検索で調べた際、それらの要素が一つに限定されているものしか出てこない。本研究では類似画像検索において同一とみなされない時間や環境などにより変化するランドマークについて正しい識別をすることを目指す。

2. 関連研究

市川らは鉄筋コンクリート工事における鉄筋識別の手段として畳み込みニューラルネットワークを用いた自動判別を試みている。[1] しかし、入力画像として使用している画像は連続した鉄筋を連写して撮っているものであり、背景の変化が少ないと考えられる。本研究では、背景などの環境の変化時に、正確な識別ができるかという部分に重きを置いているため、差別化できていると考える。

3. 提案手法

3.1 提案手法の概要

本研究ではランドマーク画像に対し、畳み込みニューラルネットワーク[2]を用いた学習モデルを作成し、ランドマークの識別を行う。そして、Grad-Cam[3]を用いて畳み込みニューラルネットワークの学習を可視化し、学習モデルが妥当なものであるか判断、考察を行う。

Landmark recognition using machine learning.
†Masataka Shoji, Meiji University
‡Sayaka Akioka, Meiji University

3.2 Grad-Cam

Grad-Cam は畳み込みニューラルネットワークで構築したモデルに使用すると、そのモデルが画像のどこを特徴としているか可視化する技術である。Grad-Cam は入力画像に対し、学習モデルが学習した特徴をヒートマップとして出力する。入力画像とヒートマップを合わせることで、学習モデルの特徴量を視覚情報で見ることができる。図 1 は Grad-Cam の出力例である。ヒートマップには、特徴の重みによって色分けがなされており、赤、黄、緑、青の順に重みが下がっていく。本研究では、図 1 のようにヒートマップ上でランドマークに該当する部分に反応するかどうかで学習モデルの妥当性を判断する

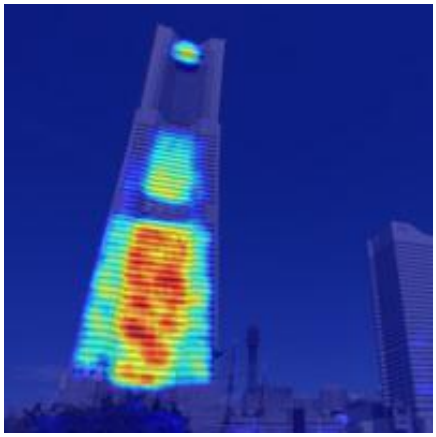

図 1 Grad-Cam 出力例

4. 実験と評価

学習画像として、東京タワー、ランドマークタワー、東京スカイツリー、首里城を用いる。学習画像については近傍画像を google と Yahoo!の二つの検索エンジンから集めたものである。それぞれのランドマークの画像を 400 枚ずつ用意し、学習用に全体の 7 割である 1120 枚、検証用に 2 割、240 枚、テスト用に 1 割、120 枚の配分で割り振った。これを学習にかけ正答率の検証を行う。その後、Grad-Cam を使う。それぞれの

Copyright ©2021 Information Processing Society of Japan.
All Rights Reserved.

ランドマーク対して建物自体を特徴として取れているかの確認を行う。

実験結果は、検証データに対しての正答率は96%、テストデータに対しての正答率は95%をそれぞれ記録した。図２、図３、図４はランドマークにGrad-Camを通した時の入力画像とヒートマップを重ねた画像である。図２、図３については建物に反応が出ていることがわかる。また、図４に関しては、建物ではなく背景の雲に反応してしまっている。図５は首里城の画像だが、東京タワーと出力されたものである。そのため、首里城とは違い、建物自体に反応があることがわかる。

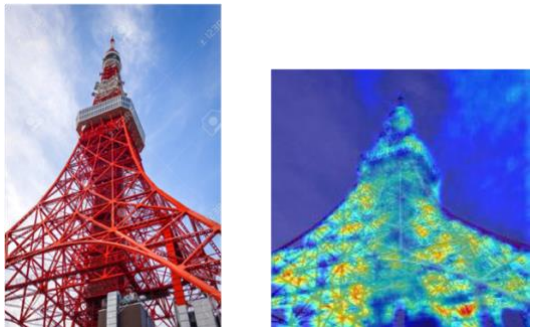

図 2 東京タワー 入出力画像

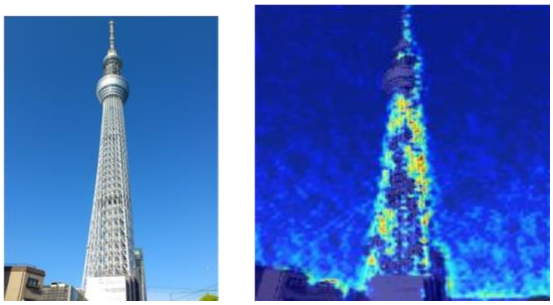

図 3 東京スカイツリー　入出力画像

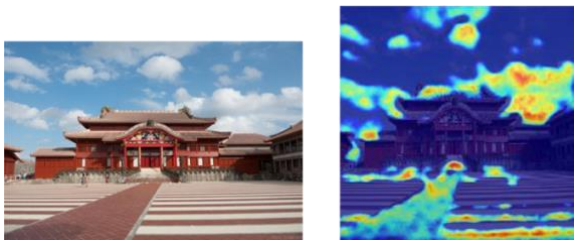

図 4　首里城　入出力画像

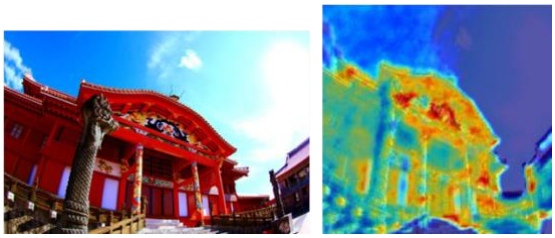

図 5 東京タワーと識別された首里城

5. 考察

首里城だけ建物を特徴として取らない理由は色にあると考えられる。これは図２、図３、図５の結果からも見て取れる。図５の首里城は図４のものと比べると建物の彩度が高く、図２の東京タワーに近い色合いをしていることがわかる。ことからも、この学習モデルでは赤い特徴は東京タワーのものであり、首里城は別の部分を特徴として学習していると考えられる。これを解決するためには、畳み込みニューラルネットワークが持つ色に釣られやすいという欠点を解決する必要がある。そのために、入力画像に対する前処理を加え、入力画像のラベルを増やすことで釣られにくくなるのではないかと考えている。

6. おわりに

本研究では畳み込みニューラルネットワークを用いたランドマークの識別を行なった。前項でも挙げたように、今回作成したモデルでは色や出現位置に依存する点について解決していない。その問題を解決できるモデルに改善する必要がある。また、今回は４種類のランドマークを用いて実験を行ったが、今後はランドマーク数の増加と精度の維持を行ないたいと考える。

7. 参考文献

[1]. 市川 享祐, 石畑 宏明；“画像認識技術の鉄筋自動識別への試適用”, 人工知能学会全国大会論文集, 2018

[2]. Francois Chollet；“Deep Learning with Python”, Manning Publications, 2017

[3]. Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam；“Grad-CAM: VisualExplanationsfromDeepNetworksviaGradient-basedLocalization”

Copyright ©2021 Information Processing Society of Japan.
All Rights Reserved.