

投稿写真の撮影内容に基づく 地域の特徴を表すピクトリアルマップの自動生成

周 文彤[†] 牛尼 剛聡[‡]

[†]九州大学大学院芸術工学府 〒815-8540 福岡県福岡市南区塩原 4-9-1

[‡]九州大学大学院芸術工学研究院 〒815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: [†]2DS18093P@s.kyushu-u.ac.jp, [‡]ushiana@design.kyushu-u.ac.jp

あらまし 近年、スマートフォンや SNS の普及によって、膨大な写真が SNS 上に投稿されている。風景の美しさが表れるコンテンツである風景写真には、それぞれの地域の異なる魅力が反映される。本研究では、ジオタグ付きの風景写真を用いて、地域の特徴を表す代表的な写真を抽出するシステムを開発することを目標とする。本システムを利用して、本論文では、Flickr 上に投稿された風景写真データから、地域の魅力を分かりやすく伝えるピクトリアルマップを自動的に生成する手法を提案する。これにより、地理情報検索や観光等様々な応用をサポートすることができると考えられる。

キーワード 風景写真、深層学習、美しさ、SNS、可視化

1. はじめに

1.1. 背景

近年、スマートフォンと SNS の普及によって目の前の美しい風景を見たとき、その瞬間に感じた美しさを写真に撮影して、SNS に投稿することが一般的に行われるようになった。世界中の SNS ユーザは、SNS に投稿された写真を見ることで、様々な地域の魅力や面白さを感じることができる。そのため、SNS を用いることで、これまでは知ることができなかった知らない地域の美しさを発見できるようになった。

一方、近年の人工知能(AI)の技術の大幅な向上により、画像の処理技術はめざましい発展を遂げた。機械学習の手法の一つであるディープラーニング (Deep Learning) を通して、写真の美しさといった高度な特徴も自動的に抽出できる可能性が高まってきた。そのため、投稿された膨大なジオタグ付き写真データを収集・分析することで、それぞれの地域の異なる魅力が発見されるのではないかと考えられる。

SNS に投稿された写真が地域の特徴をどのように反映するのかについて、従来の研究は、写真の一部の特徴だけ (例えば、写真の色情報) を考慮している^{[1][2][7]}。しかし、本研究は写真の撮影内容に基づいて、学習済みディープラーニングモデルを利用して、画像の全般的な特徴を抽出して利用する。

さらに、従来の研究で使用されたデータは一定期間内に投稿した全ての風景写真である^[2]。しかし、本研究は地域の魅力をより正確に伝えるために、まず全ての風景写真から美しい写真を抽出し、そして代表的な写真を決定してピクトリアルマップを生成する。

本研究の目標とは、収集したジオタグ付き風景写真を利用して、地域の特徴を自動的に認識・分析するこ

とで、地域の代表的な写真を選別する。そして、各地域の特徴の違いと地域全体的な特徴を把握するために、代表的な写真を地図にマッピングして、ピクトリアルマップを自動的に生成することを目標とする。

2. 関連研究

本章では、様々な画像データを利用した研究を 2 種類の分野に分けて紹介する。一つ目は SNS に投稿された写真に反映されている地域の文化に関する研究、もう一つは風景を撮影した画像から地理的特徴を認識することに関する研究である。

2.1. SNS に投稿された写真に反映されている地域の文化に関する研究

SNS に投稿された写真には投稿者の価値観や投稿者が所属する文化圏の特徴が反映されていると考えられる。近年、SNS に投稿された写真からの特徴抽出に関する研究が活発に行われている。

Redi^[3]らは Instagram に投稿された写真の内容とスタイルを分析して可視化することで、全世界有名な五つの大都市間の文化的類似点と相違点を明らかにしている。

Manovich^[4]らは Instagram に投稿する自撮り写真の内容を分析して可視化することで、世界中の 5 都市間のユーザの特性を明らかにしている。

以上に挙げた研究と同様に、本研究は SNS に投稿された風景写真を利用して、体系的に分析することで、写真に反映されている地域の魅力を発見することを目指す。

2.2. 風景および地理的特徴の認識

これまでにも、位置情報付きデータを利用した風景および地理的特徴の認識に関する研究が多数存在する。久保田^[1]らは、代表的な写真 SNS である Instagram

を利用して、特定の地域で投稿された膨大な写真データの視覚的な特徴を利用して、地域の「見どころ」と「見ごろ」を分かりやすく伝える可視化手法を提案した。高ら^[2]は、位置情報付きの写真に対してクラスタリングにより代表的な地域を抽出し、写真の色情報を利用して、その地域の代表的な写真を抽出する手法を提案している。

Zhou^{[5][6]}らは、風景（屋外&屋内）のシーン画像データセットを基に畳み込みニューラルネットワーク(CNN)を用いた転移学習により、シーン画像の識別器を構築している。

Doersch^[7]らは、膨大なジオタグ付きデータから地理的位置を特徴付ける視覚的特徴を自動的に発見する方法を設計した。地域の視覚的特徴の二つの尺度：出現頻度と差別（Frequent and Discriminative）この2つを組み合わせることで特徴の重みを計算して、地域の代表的な特徴を選別した。

以上に挙げた研究の中で、Doersch らの研究では、本研究と似たように、写真による地域の代表的な特徴の抽出を目的としている。しかし、彼らの研究ではHOG 特徴と色情報を組み合わせた識別機を利用して物体の局所特徴量を検出している。しかし、近年、畳み込みニューラルネットワーク(CNN)の特徴抽出器を利用して行う研究は驚異的な性能を達成し、注目を集めている。そこで、本研究では、Zhou らが構築したシーン画像の識別器を用い、風景写真の属性を自動的に識別する。そして、Doersch らと同様に、風景写真の属性の二つの尺度である出現頻度と差別度を組み合わせることで特徴の重みを計算して、地域の代表的な写真を選別するシステムを構築することを目指す。

3. アプローチ

本研究では、まず、学習済みディープラーニングモデルを利用して、収集したジオタグ付き風景写真から美しい写真を抽出する。次に、写真に付けられたジオタグ情報に基づいて、クラスタリングを行って、魅力的な地点を発見する。そして、もう一つの学習済みディープラーニングモデルを使って、風景の属性と特徴を自動的に認識する。また、特徴間の類似度に基づく代表的な写真を抽出する。最後に、写真の地理情報に基づいて、代表的な写真を地図にマッピングする。

図1に提案手法の流れを示す。

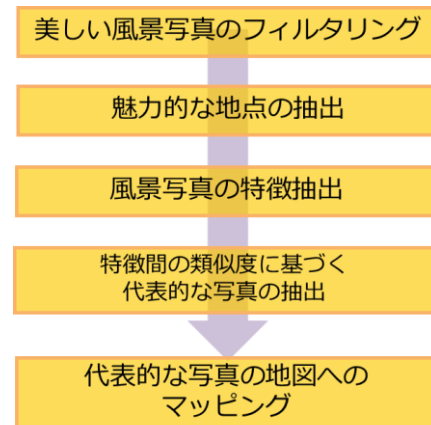


図 1 提案手法の流れ

4. 提案手法

本章では、地域の特徴を表す代表的な写真を選別する手法について述べる。本手法を5段階に分けて説明する。

4.1. データセットの構築

写真共有 SNS から、ジオタグ付き写真データを収集できる。例えば、人気がある写真共有 SNS には Instagram や Flickr などがある。

本研究では Flickr API を利用して、日本国福岡県の4120 枚ジオタグ付き風景写真を取得して、風景写真のデータセットを構築した。

取得した風景写真のデータは、写真の url、ジオタグ情報（緯度と経度）、写真のタイトル、によって構成されている。この風景写真のデータから url を抽出し、各ソースの Web ページ上の写真にアクセスし、写真を取得する。

しかし、地域の美しさや魅力を反映したピクトリアルマップを生成するためには、本研究は美しい写真だけを使い、単なる記録のために撮影された写真や美しくないと考えられる写真を除外する必要がある。

4.2. 美しい写真を自動的に抽出する

4.1.で収集した風景写真から美しい写真を自動的に抽出する手法について述べる。

従来から、コンピュータで写真の美しさを判別するため、手動で特徴を設計して教師あり学習を用いたパターン認識モデルを利用して写真の美しさを識別する研究が行われてきた。しかし、2014 年以降、画像の美しさに関する研究にはディープラーニング技術を採用されることが多くなった。これはディープラーニング技術を通じて自動的に特徴を識別できるようになったからである。それにより、写真の良さ悪さの判別の確率を大幅に上がっている。AVA データセット（a Large-scale database for aesthetic visual analysis）^[8]において今まで分類確率が一番高いモデルは

ILGNet-Inc.V4^[9]である。

本研究は学習済み ILGNet-Inc.V4 を利用して全ての風景写真から美しい風景写真を識別する．スコア値 5 以上は良い写真として識別して抽出する，同様にスコア値 5 以下は良くない写真と判断して棄却する(図 2)．



図 2 美しい写真抽出の流れ

上記の処理により，4120 枚の写真から 566 枚の美しい風景写真を得る．図 3 と図 4 に ILGNet-Inc.V4 による写真の識別結果の例を示す．図 3 は美しいと判断された写真であり，図 4 は美しくないと判断された写真である．



図 3 ILGNet-Inc.V4 によって分類された美しい風景写真

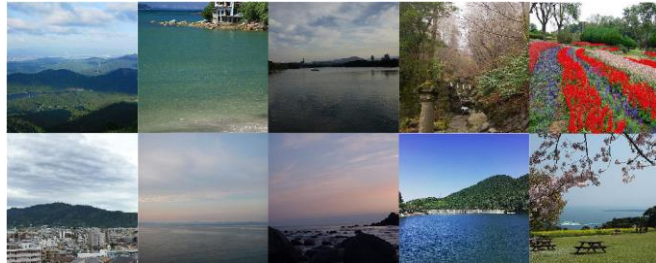


図 4 ILGNet-Inc.V4 によって分類された美しくない風景写真

4.3. 魅力的な地点の抽出

本節では，写真のジオタグ情報に基づいて魅力的な地点を抽出する手法について述べる．

本研究において，美しい写真の投稿数が密集する場所を魅力的な地点と定義する．そのため，4.2 節で抽出した美しい風景写真に付与されている緯度と経度を特徴量として，密度ベースのクラスタリング手法の 1 つ DBSCAN^[10] (Density-based spatial clustering of applications with noise)を用い，魅力的な地点を抽出する．

DBSCAN 手法は，クラスタ間の距離の閾値 Eps とクラスタを構成する最小データ数の閾値 $MinPts$ との 2 つの閾値を持つ．ある点 x から，距離 Eps 内にある点集合

を近傍 $N_{Eps}(x)$ と定義し，以下の接続関係を満たす時，同じクラスタに分類する．

$$(1) y \in N_{Eps}(x)$$

$$(2) |N_{Eps}(x)| \geq MinPts(\text{コア点})$$

つまり，ある座標から半径 Eps 以内に $MinPts$ 以上の座標集合が存在するならば，同じクラスタに分類する．図 5 に DBSCAN によるクラスタ結果を示す．

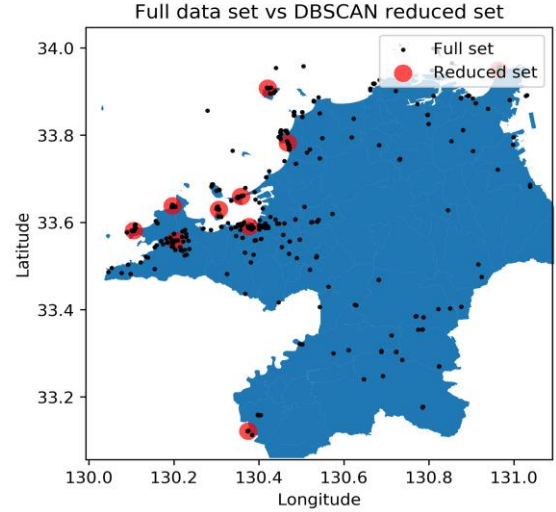


図 5 DBSCAN によるクラスタ結果 (福岡県)

また，生成されたクラスタの中心点を求める．クラスタ内のスポットに属するそれぞれの写真の撮影位置の緯度と経度を (ϕ_i, λ_i) ，クラスタの重心を (ϕ_o, λ_o) ，また $\Delta\phi$ ， $\Delta\lambda$ はその差の絶対値とした際に，2 点間の中心角 $\Delta\sigma$ は球面余弦定理より以下の式で求める．

$$\Delta\sigma = \arccos(\sin \phi_i \cdot \sin \phi_o + \cos \phi_i \cdot \cos \phi_o \cdot \cos(\Delta\lambda))$$

大円距離 d ，すなわち円弧長は，球の半径 r ，弧度で表された $\Delta\sigma$ を用いて以下の式で求める．

$$d = r\Delta\sigma$$

クラスタ C の中心点 (x_c, y_c) を以下の式で求める．

$$(x_c, y_c = \text{Min}(d_1, d_2, \dots, d_i, \dots, d_n))$$

つまり，スポットと重心の大円距離をそれぞれに計算して，距離は最小となるスポットをクラススの中心点として判断する．この中心点は魅力的な地点とする．図 6 に DBSCAN による抽出した魅力的な地点の結果を示す．

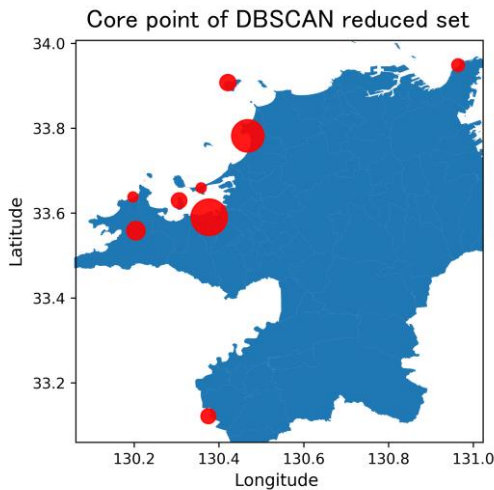


図 6 DBSCAN による抽出した魅力的な地点（福岡県）

4.4. 風景写真の属性抽出

本節では、風景写真の属性を自動的に抽出する手法について述べる。

Hwang^[11]らは、画像の視覚属性 (visual attribute) を、物体カテゴリ間で共有される人間が理解可能な属性であると定義している。本研究において、風景写真の属性はシーンカテゴリ間で共有される人間が理解可能な属性とする。一方、画像認識分野において、畳み込みニューラルネットワーク (CNN) は多くのタスクで優れた性能を発揮している。本研究は学習済み Places365-CNNs^[12]モデルを用いる。Places365-CNN では、シーン認識データセットを用いた転移学習したモデルである。シーン認識データセットには 400 以上のシーンカテゴリを含む 1000 万以上の画像が含まれている。本研究は学習済み Places365-CNNs を利用して、4.2.節で抽出した美しい風景写真のシーンの属性 (Scene Attributes) を取得できる (図 7)。



図 7 風景写真のシーンの属性抽出の流れ

そこで、4.3.で抽出された魅力的な地点に投稿された、それぞれの風景写真のシーンの属性を追加することで、各地点のシーンの属性を取得する。

4.5. 代表的な写真の抽出

本節では、4.4.で抽出した各地点のシーンの属性をモデル化し、比較することにより、地域代表的な写真の抽出方法について述べる。

まず、それぞれのシーンの属性の出現回数をベクトル化して、各地点の属性のベクトルの生成を行う。ま

た、あるシーンの属性 a の重要度は以下の 2 つの要因によって決定されると仮定する：

(1)出現頻度 (Attribute Frequency)：ある地点 A の属性 a の出現数

(2)差別度 (Discriminative Frequency)：属性 a が地点 A 以外に出現しないことが多い

この考え方は、文書中に含まれる単語の重要度を評価する手法である TF-IDF と同じ概念である。 af はある地点 A に出現する属性 a の頻度 (Attribute Frequency), isf はある範囲に全ての地点数/属性 a が出現する地点の総数の対数 (Inverted Spot Frequency) であり, $afisf$ はその積となる。地点集合 S に存在する地点 s_j に含まれる属性 a_i のスコア $afisf_{i,j}$ は、出現頻度 $af_{i,j}$ 及び差別度 $isf_{i,j}$ の 2 つの指標を用い、以下のように計算される。

$$afisf = af_{i,j} \cdot isf_{i,j}$$

$$af_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

$$isf_{i,j} = \log \frac{|S|}{sf_i} + 1, sf_i = |\{s: s \ni a_i\}|$$

ここで、 $n_{i,j}$ は地点 s_j における属性 a_i の出現回数を表す。

この $afisf$ の計算式により得られた風景写真の属性のベクトル a と地点の属性のベクトル b 間の類似度を以下に示すコサイン相関値で計算する。

$$\cos(a, b) = \frac{a \cdot b}{|a||b|}$$

各地点の風景属性のベクトルと各地点に対する風景写真の属性のベクトルの類似度はコサイン類似度を用いて計算して、類似度の高い順に 2 枚を代表的な写真として扱う。

図 8 に本節のアルゴリズムにより抽出した代表的な写真の例を示す (スペース制限のため、一部の写真だけを示す)。左側は各地点におけるランダムな風景写真である。右側は各地点の代表的な写真である。



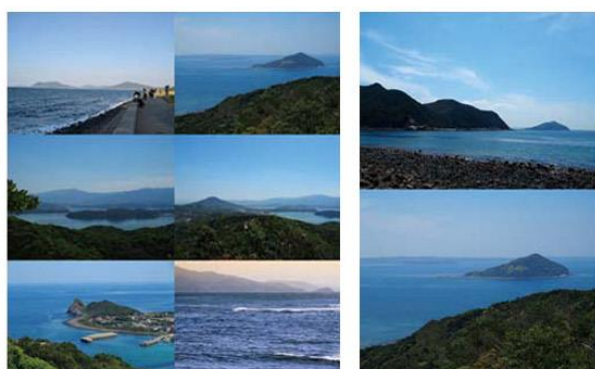
福津市の海岸におけるランダムな風景写真

抽出した代表的な写真



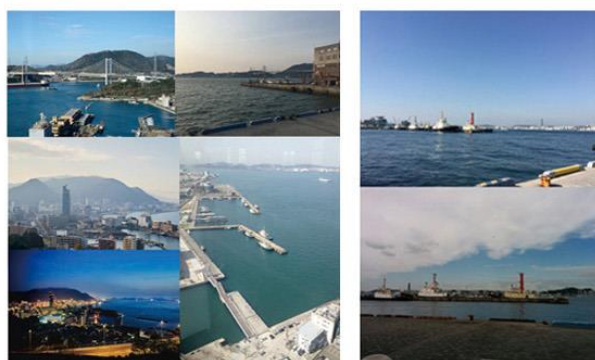
宗像大島におけるランダムな風景写真

抽出した代表的な写真



能古島におけるランダムな風景写真

抽出した代表的な写真



門司港レトロにおけるランダムな風景写真

抽出した代表的な写真

図 8 抽出した代表的な写真

5. 実験と評価

本章では、抽出された代表的な写真の適合性を検証するために行った実験内容と結果について述べる。

5.1. 地域代表的な写真に関する評価

5.1.1. 実験内容

本実験では、提案手法に基づいて抽出した代表的な写真とインターネットで公開されている二つの記事「福岡県の絶景スポット」^[13]、「福岡県の写真撮影スポット」^[14]に添付されている各スポットの写真の内容と比較し、提案手法の有効性を検証した（図 9 と図 10）。

5.1.2. 考察

図 9 と図 10 を見ると、提案手法による抽出した写真と人間による選択された写真と比べると、両方とも各地点の風景のユニークな魅力を反映されることが分かる。また、記事「福岡県の絶景スポット」と「福岡県の写真撮影スポット」に載っていないスポットは「×」にする。そこで、提案手法から、知られていないスポットを発見したことがわかる。

5.2. 発見された新しいスポットに関する評価

5.2.1. 実験内容

発見された新しいスポットが価値のあるスポットであるかどうか検証するために、被験者調査を行った。

被験者は福岡県福岡市に在住する大学生 4 人、写真家 1 人である。まず、発見された新しいスポットは被験者全員よく知っていないことが確認した。そして、抽出された写真を被験者に提示し、被験者に A、B の質問をした。

A 写真が表示されている地点は魅力的な地点だと思いますか？

B 都合、お金、時間などといった要素を考慮せず、写真が表示されている地点に行きたいのか？

表 1 と表 2 に質問への回答を示す。

表 1 質問 A への答え

全くそう 思わない	あまりそう 思わない	どちらとも いえない	ややそう 思う	とても 思う
0 人	0 人	0 人	4 人	1 人

表 2 質問 B への答え

はい	いいえ
5 人	0 人

5.2.2. 考察

表 1 は「写真が表示されている地点は魅力的な地点だと思いますか？」の結果を示している。表 1 の結果から、「ややそう思う」を選ぶ被験者の割合が最も高かった。表 2 は「都合、お金、時間などといった要素を考慮せず、写真が表示されている地点に行きたいのか？」の結果を示している。表 2 の結果から、被験者全員「はい」を選びました。

実験結果で示した、提案手法による発見された新し

いスポットは「隠れた魅力的な風景スポット」であると言えるだろう。ただし、今回の実験では、被験者数が少ないため、今後、より大規模な実験を行う必要がある。

5.3. ピクトリアルマップの例

本節では、提案手法に基づいて生成されたピクトリアルマップの例を示し、その特徴について述べる。

図 10 で提案手法に基づいて生成されたピクトリアルマップの例を示す。福岡県のピクトリアルマップから公園風景、海岸風景、島の風景、花の海、港・船の見える風景の写真が見える。福岡県にある多くの美しい海の風景がピクトリアルマップに直観的に反映されているのがわかる。

6. おわりに

本研究では、ジオタグ情報付き風景写真を用いて、地域の特徴を表す代表的な写真を抽出するシステムを開発する。本システムを利用して、Flick 上に投稿された福岡県の風景写真データから、県内各地の魅力を分かりやすく伝えるピクトリアルマップを自動的に生成した。ピクトリアルマップを自動的に生成することによって、地理情報検索や観光等様々な応用をサポートすることを期待している。

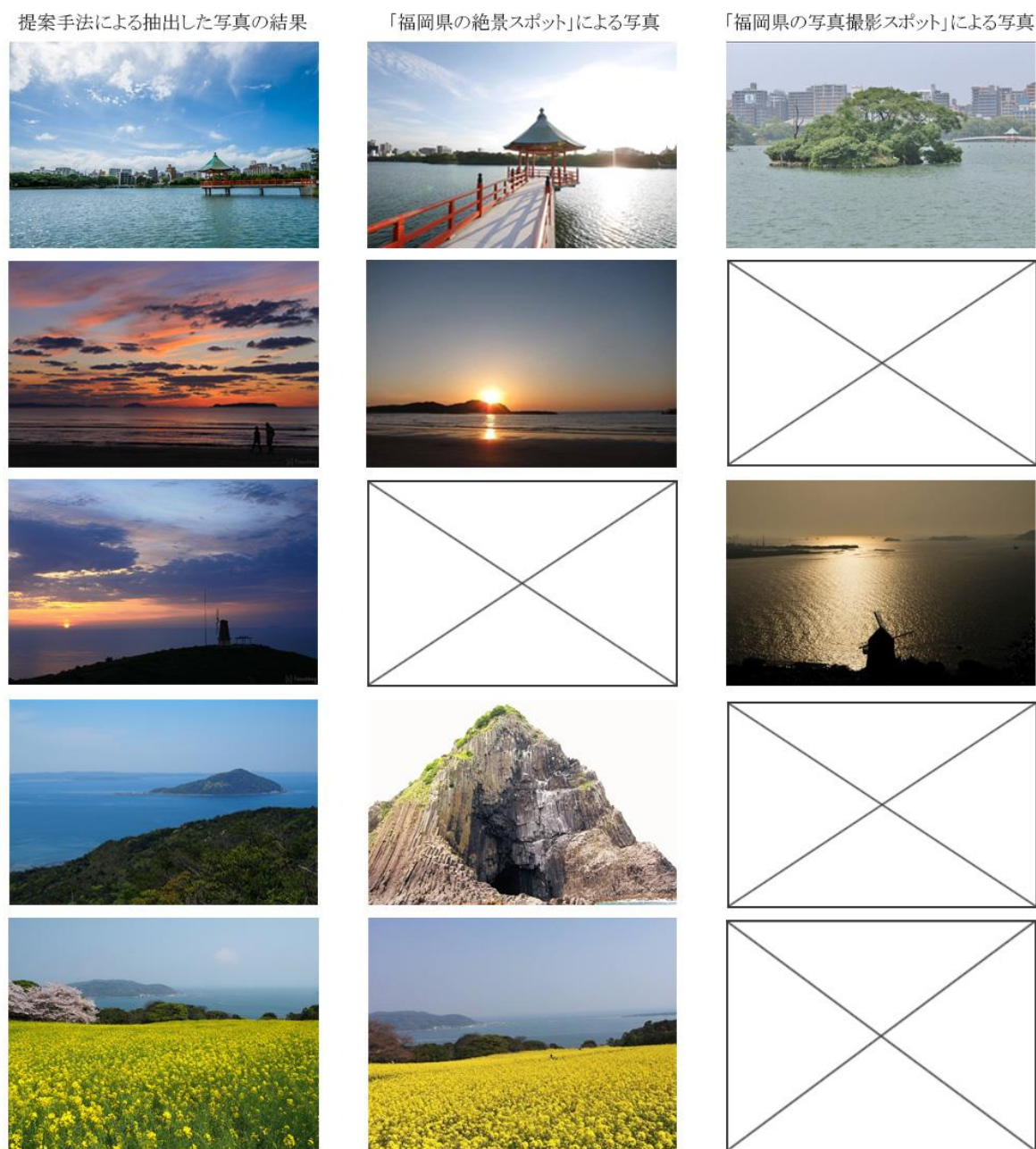


図 9 提案手法による抽出した写真と Web サイトで公開されている写真 (1)

提案手法による抽出した写真の結果



「福岡県の絶景スポット」による写真



「福岡県の写真撮影スポット」による写真

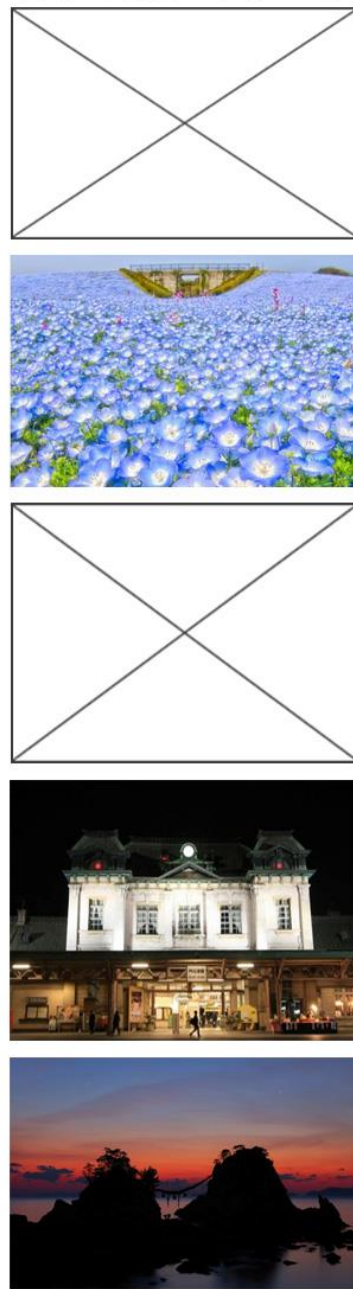


図 10 提案手法による抽出した写真と Web サイトで公開されている写真（2）

今後の課題として、まず、美しい風景写真を識別するために、2 分類問題にするのではなく回帰分析を使ってより適切な写真の美しさの識別方法を検討する必要がある。また、今回使用したデータは福岡県に半年間投稿された写真であったが、今後、全日本一年間の写真データを収集し、時間とともに変化する特徴を考慮して実験する予定である。

参 考 文 献

- [1] 久保田 麻美, 牛尼 剛聡, “SNS による文化と文士の可視化”, DEIM Forum 2015, G7-1, 2015.
- [2] 高 尚暉, 牛尼 剛聡, “SNS を利用したピクトリアルマップの自動生成”, DEIM Forum 2017, D7-4, 2017.
- [3] Miriam Redi, Damon Crockett, Lev Manovich, Simon Osindero, “What Makes Photo Cultures Different?”, Proceedings of the 24th ACM international conference on Multimedia, pp. 287-291, 2016.
- [4] “Selfiecity”, <http://selfiecity.net/#intro>.
- [5] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, “Learning deep features for scene recognition using places database”, in Advances in neural information processing systems, pp. 487-495, 2014.
- [6] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba, “Places: A 10 million Image Database for Scene Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.1-14, 2017.
- [7] Carl Doersch, Saurabh Singh, Abhinav Gupta, Josef Sivic, and Alexei A. Efros, “What Makes Paris Look like Paris?”, ACM Transactions on Graphics (SIGGRAPH 2012), August 2012, vol. 31, No. 3.
- [8] Murray N, Marchesotti L, Perronnin F, “AVA: A large- scale database for aesthetic visual analysis”, Proceedings of the 25th IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.2408-2415, 2012.
- [9] Xin Jin and Jingying Chi and Siwei Peng and Yulu Tian and Chaochen Ye and Xiaodong Li, “Deep image aesthetics classification using inception modules and fine-tuning connected layer”, 8th International Conference on Wireless Communications & Signal Processing, pp.1-6, 2016.
- [10] Martin Ester, Hans-Peter Kriegel, Jorg Sander, Xiaowei Xu, “A density-Based Algorithm for Discovering Clusters”, KDD-96, pp.226-231, 1996.
- [11] S. J. Hwang, F. Sha, K. Grauman, “Sharing Features Between Objects and Their Attributes”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.1761-1768, 2011.
- [12] M. C. Science and A. I. Laboratory. Places CNN. [Online]. Available: <http://places.csail.mit.edu/downloadCNN.html>
- [13] 福岡県の絶景スポット, <https://zekkei-project.com/areas/countries/153/prefectures/44/spot>
- [14] 福岡県の写真撮影スポット, <https://ganref.jp/spot/photo/jpn/fukuoka.html>



図 11 ピクトリアルマップの例