

個人の好みに合わせた写真を提供するシステムの構築 Building a system that provides photos that suit individual tastes

田中敦大 戴瑩
岩手県立大学

1. はじめに

近年, 多数の画像解析に関する研究が行われているが, その多くが写真の美的要素が考慮されていないことや, 写真に対する普遍的な評価手順が存在しないこと, 個人によって美的嗜好が異なるため評価に一貫性を得ることが難しいという課題が挙げられている.

そこで本研究では, AlexNet を使用した転移学習²⁾等の技術を利用し, 個人の好みに合わせた写真を提供するシステムの構築を目指す. そのために, 本論文では既存のデータセットである xiheAA データセット¹⁾を用いて複数人の美的評価, 好みの評価を行うことで CNN に基づく複数の美的評価と好みの予測モデルを訓練する. さらに, 構築されたモデルを用い, 写真の美的特徴マップを抽出し分析を行う.

2. 提案方法

2.1 システム概要

図 1 にシステム構成図を示す. 本研究では, 個人の好みに合わせた写真を提供するシステムの作成を目指し, CNN を用いて, 美的評価モデルと好みモデルのそれぞれのモデルの訓練と検証を行う. 写真に対して知識のある人, 知識の無い人それぞれが既存の xiheAA データセットの写真に採点を行い, CNN を用いて写真の美的評価モデルの構築と写真の好みモデルを構築する. 構築されたモデルを利用し, 美的要素や個人の好みの要素の関連性の分析を行い, 個人の好みに合わせた写真の提供を行う.

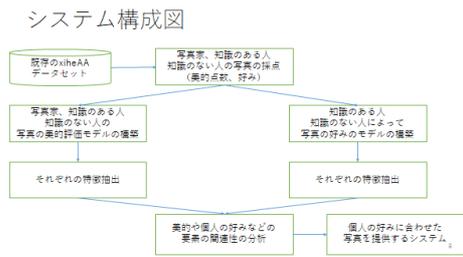


図 1 システム構成図

2.2 データセット

既存の xiheAA データセットを用い, 学習データセットとして活用した.

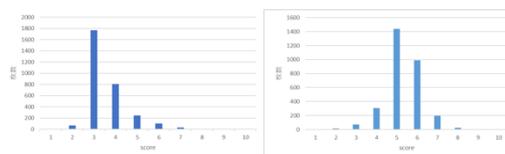


図 2 知識のある人の美的評価分布 (左)

図 3 知識の無い人の美的評価分布 (右)

美的評価モデルを訓練するために, 複数人から写真を 1 から 10 点の 10 種類でラベル付けを行った. 図 2 と図 3 により, 写真に対して知識のある一人は最頻値が 3 点となり, 知識の無い一人は 5 点となった. その後, 3040 枚の写真の内, ランダムに選んだ 4/5 の写真, 2432 枚を学習用データセットとし残りの 1/5 の写真, 608 枚をテストデータとして用いて知識のある人の美的評価モデル, 知識の無い人の美的評価モデルのモデル構築を行った.

好みモデルを訓練するために, 特に好きな写真と嫌いな写真の 2 種類でラベル付けを行い, 写真に対して知識のある人は好きな写真 94 枚, 嫌いな写真 24 枚の計 118 枚の内ランダムに選んだ 4/5 の写真 94 枚を学習用データセット, 残りの 1/5 の写真 24 枚をテスト用データセットとして好みモデル構築を行った. 同様に知識の無い人は好きな写真 23 枚, 嫌いな写真 17 枚の計 40 枚の内ランダムに選んだ 4/5 の写真 32 枚を学習用データセット, 残りの 1/5 の写真の 8 枚をテスト用データセットとして好みモデル構築を行った.

3. 実装環境と実現方法

CNN に基づくモデルの構築は, MATLAB(R2021a) を利用する. 実装は Alexnet に基づいた転移学習を用いて行う. 転移学習とは, 学習済みモデルを別の領域に転移し, 効率的に学習させる方法のことである²⁾. あとは学習された CNN モデルの第 15 層である relu5 層の特徴マップを抽出する上で美的特徴を分析する.

4. 評価実験

2.2 で述べた 4 つのモデルを用いて検証を行い、テストの精度を求める。精度を求める式を以下に示す。

$$A=(Y=T)/T$$

(A : 精度, Y=T:予測データとテストラベルの一致した数, T:テストラベルの数)

検証によって得られた知識のある人,知識の無い人の美的評価モデルの性能と好みモデルの性能を以下に示す。

表 1 知識のある人の美的評価モデルのテスト精度

回数	1	2	3	4	5
精度	56.4%	56.9%	55.3%	54.0%	53.8%

表 2 知識のない人の美的評価モデルのテスト精度

回数	1	2	3	4	5
精度	54.4%	52.1%	50.2%	52.3%	57.6%

表 3 知識のある人の好みモデルのテスト精度

回数	1	2	3	4	5
精度	83.3%	70.8%	79.2%	58.3%	70.8%

表 4 知識のない人の好みモデルのテスト精度

回数	1	2	3	4	5
精度	87.5%	50.0%	75.0%	75.0%	62.5%

美的評価モデルでは知識のある人,知識のない人のそれぞれが約 50%,好みモデルでは値のばらつきはあるが知識のある人,知識のない人のそれぞれが複数回 70%を上回るテスト精度を示しており,美的評価モデルより高い性能を持つと考えられた。

5. 美的特徴抽出と分析

作成した 4 つのモデルで,3.に述べた方法でサンプル写真の特徴マップの抽出を行った。



図 4 元画像
(知識のある人 9 点+特に好き
知識の無い人 7 点)



図 5 知識のある人の美的評価モデルから抽出された重ねる特徴マップ (左) (予測結果 3 点)

図 6 知識のない人の美的評価モデルから

抽出された重ねる特徴マップ (右) (予測結果 5 点)

知識のある人,知識の無い人の美的評価モデルから抽出された特徴マップは,実際の美的視点と同様に左の大きな建物から右の建物,さらにはその奥や星まで評価領域として抽出された。

しかし,予測結果は実際に付けた点数とは大きく異なり,人が点数を付ける際に重視する美的要素は正しく判断されなかったと考えられた。



図 7 知識のある人の好みモデルから抽出された評価領域 (左) (予測結果 好き)

図 8 知識の無い人の好みモデルから抽出された評価領域 (右) (予測結果 好き)

知識のある人,知識の無い人の好みモデルから抽出された特徴マップも,実際の美的視点と同様に広い範囲を評価領域として抽出した。

また,予測結果は知識のある人,知識の無い人のそれぞれが好きという結果になり,どちらでもないと答えた知識の無い人はこのモデルではどちらかといえば好きという結果になった。

6. まとめ

本稿では, AlexNet を使用した転移学習により,画像から人の美的評価と好みを予測するモデルの構築を行った。美的評価モデルより好みモデルの精度が高かったことから,モデル構築においてデータセットの数よりラベル付けの質は重要であると考えられた。

また,それぞれのモデルから抽出された特徴マップは実際の美的視点と一致していたが,美的評価モデルの予測結果は大きく異なっていたため,モデルの構築には改良の余地が見られた。

今後は,個人の好みに合わせた写真を提供するシステムの構築を目指す。

参考文献

- 1) Ying Dai, "Sample-specific repetitive learning for photo aesthetic auto-assessment and highlight elements analysis", Multimedia Tools and Applications, Vol. 80, pp. 1387-1402, DOI: 10.1007/s11042-020-09426-z, 2021
- 2) AlexNet を使用した転移学習
<https://jp.mathworks.com/help/deeplearning/ug/transfer-learning-using-alexnet.html>