

画風の類似性に着目したイラスト検索における輪郭線情報の利用による精度の向上

金井陸 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

イラスト検索において特徴ごとの重み付けを考慮した畳み込みニューラルネットワークを用いたイラスト検索 [1] が提案されており、色の出現頻度や特徴ごとの重み付けを考慮した精度向上などが検討されている。この手法では彩度・明度・色相のヒストグラム、スタイル行列を入力とし、画風の類似性に基づいたグループに分類する問題を学習した畳み込みニューラルネットワークを利用してイラスト検索を行っている。検索には出力層の手前の全結合層の出力を特徴量として利用し、特徴ベクトル間の距離を計算することによって類似したイラスト検索を行うが、このとき、重視される特徴ごとの重み付けを考慮した検索を行うことができるようにしている。しかしながら、この手法では輪郭線の情報が適切に扱えていないため、検索において画風の類似性が十分に考慮できていないという問題点がある。それに対し、Anime2Sketch[2] というイラスト画像をスケッチ画像に変換する手法が提案されている。この手法は、輪郭線抽出を行う手法ではないが、ファインチューニングを行うことで、輪郭線抽出に利用できる可能性がある。

本研究では、Anime2Sketch を輪郭線抽出が行えるようにファインチューニングを行い、画風の類似性に着目したイラスト検索における輪郭線情報の利用による精度の向上を目指す。

2 Anime2Sketch

Anime2Sketch はイラスト画像をスケッチ画像に変換するという手法であり、敵対的オープンドメイン適応 (Adversarial Open Domain Adaptation: AODA)[3] というスケッチ画像からの写真の生成において使われている技術の一部を利用してスケッチ画像への変換を実現している。

敵対的オープンドメイン適応では、写真からスケッチへと変換する生成器 G_s 、スケッチから写真へと変換する生成器 G_p の2つの生成器と2つの識別器 D_s 、 D_p

を使用している。また、クラスラベルを予測する分類器 R により G_p が同じクラスの実際の写真と区別できない写真を生成するようにする。この手法では非対応の写真 p と実際のスケッチ s を使って学習を行っており、写真からスケッチ、スケッチから写真への変換とともに学習し、写真 p から再構成された画像と、実際のスケッチ s から生成された写真の間のギャップを狭めることで、訓練データにないフリーハンドスケッチからでもリアルな写真を合成することができるようにしている。Anime2Sketch は、この手法の写真からスケッチへと変換する部分を利用したものになっている。

3 画風の類似性に着目したイラスト検索

従来のシステム [1] では、彩度・明度・色相のヒストグラム・スタイル行列を入力として用いていたが、提案システムでは Anime2Sketch で生成した画像も入力として利用し、輪郭線の特徴も考慮した検索が行えるようにする。画風の類似性に基づいたグループに分類する問題として畳み込みニューラルネットワークを学習し、出力層の手前の全結合層の出力を特徴量として、重みづけも考慮して検索を行う。

3.1 畳み込みニューラルネットワークの構造

一般的な畳み込みニューラルネットワーク [4] は入力から出力まで一連の処理で行うが、提案システムでは、従来のシステムと同様に、輪郭線、彩度・明度のヒストグラム、色相のヒストグラム、スタイル行列のような入力ごとに出力層まで独立した形になっている (図1)。これは、検索を行う際に特徴ごとの内部状態をもとに特徴ごとの重みを求めるためであり、なお、スタイル行列に関しては、それ自身が画風情報を持つため畳み込み層などには通さず、全結合層に直接入力されるような構造にしている。

3.2 入力データ

Anime2Sketch の出力は、図2に示すように輪郭線の太い部分などは黒く塗りつぶされない形で抽出されてしまうことがある。そこで、提案システムでは、

Improving Accuracy by using Contour Information in Illustration Search focusing on Similarity in Style
Riku Kanai and Osana Yuko(Tokyo University of Technology, osana@stf.teu.ac.jp)

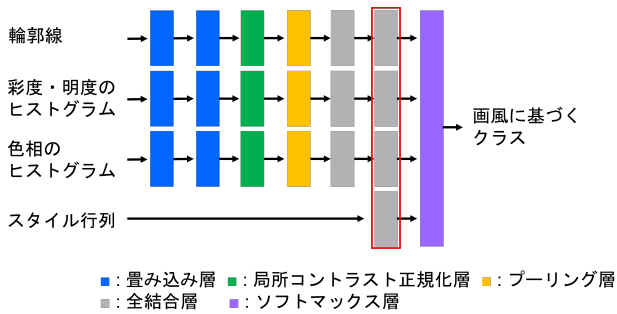


図 1: 提案システムで用いる畳み込みニューラルネットワークの構造

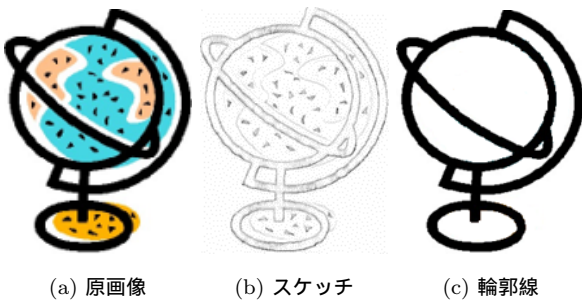


図 2: 輪郭線抽出

Anime2Sketch をファインチューニングすることで図 2(c) のような輪郭線抽出が行えるようなネットワークを作成する。輪郭線抽出ネットワークに学習させる輪郭線データは Anime2Sketch の出力結果を加工することで作成し、輪郭線抽出ネットワークの出力で生成したデータを入力データとして使用する。

3.3 スタイル行列

スタイル行列とは、画風変換 [5] において画風を表すのに用いられる特徴量である。画風変換とは、その画像が何を表すかを示すコンテンツ画像と画風情報として使用するスタイル画像の 2 つを用い、スタイル画像風のコンテンツ画像を生成する手法のことである。スタイル行列は、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 [4] である VGG16[6] の中間層の出力のチャンネル間の内積を利用したもので、第 l 層のチャンネルの特徴マップ i とチャンネルの特徴マップ j の内積を成分にもつグラム行列をスタイル行列として用いる。つまり、今回の i, j はその中間層のチャンネルを表し、同じ層内でのチャンネル間の内積を求めたものがスタイル行列となる。

3.4 画像蓄積過程

図 1 のような畳み込みニューラルネットワークの学習を行う。学習後のネットワークに学習に用いたデータを 1 つずつ入力し、全結合層の出力を正規化した値

を特徴ベクトルとし、これを画像と関連付けてデータベースに保存しておく。

3.5 画像検索過程

画像の検索は、検索キーとデータベース中の画像の特徴ベクトルを比較し、距離がしきい値以下の画像を検索結果として出力することで行う。なお、提案システムでは、特徴ごとの重み付け係数を特徴ベクトル間の距離を計算する際に利用することで、特徴ごとの重み付けを考慮した検索を行う。

4 計算機実験

計算機実験を行い、提案システムにおいて画風の類似性を考慮したイラスト検索が行えることを確認した。

参考文献

- [1] 神田夏, 長名優子: “畳み込みニューラルネットワークを用いた画風の類似性に着目したイラスト検索,” 情報処理学会第 85 回全国大会, 2023.
- [2] Anime2Sketch: A Sketch Extractor for Anime Arts with Deep Networks, <https://github.com/Mukosame/Anime2Sketch>, (2024/01/07 参照).
- [3] X. Xiang, D. Liu, X. Yang, Y. Zhu, X. Shen and J. P. Allebach: “Adversarial open domain adaptation for sketch-to-photo synthesis,” <https://arxiv.org/pdf/2104.05703.pdf>, 2021 (2024/01/07 参照).
- [4] 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.
- [5] L. A. Gatys, A. S. Ecker and M. Bethge: “Image style transfer using convolutional neural networks,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.2414–2423, 2016.
- [6] K. Simonyan and A. Zisserman: “Very deep convolutional networks for large-scale image recognition,” <https://arxiv.org/pdf/1409.1556.pdf>, 2014 (2024/01/07 参照).