

畳み込みニューラルネットワークを用いたタッチの類似性を考慮したイラスト検索 － 色相の出現頻度に着目した精度向上 －

新関智也 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

タッチの類似性を考慮した検索が行えるシステムとしてニューラルネットワークを利用したシステムがいくつか提案されている [1]-[3]。文献 [2], [3] のシステムでは、畳み込みニューラルネットワーク [4] を用いてタッチの類似性を考慮した検索を実現している。畳み込みニューラルネットワークを用いて画像検索を行う場合には、一般的に画像に含まれる色や形状の特徴に基づいて検索が行われることになるが、タッチの類似性を考慮した検索を行う場合には、形状ではなく、使われている色の傾向や輪郭線の太さなどに着目して検索を行う必要があるため、単に畳み込みニューラルネットワークを用いただけでは検索を行うことができない。文献 [3] のシステムでは彩度・明度の値、彩度・明度のヒストグラム、RGB のヒストグラムを入力とし、イラストのタッチに応じてどのグループに属するのかを出力する分類問題として学習を行った畳み込みニューラルネットワークを用いて、画像の検索を実現している。このシステムでは、彩度・明度の値を入力として用いている部分では、輪郭線の太さの情報を抽出しようとしているが、畳み込み層で行われている処理によっては意図したとおりに動作していない可能性がある。また、RGB のヒストグラムを入力として用いているが、R, G, B のそれぞれのヒストグラムを用いているため、画像中に含まれる色が異なってもヒストグラムが一致してしまうことがあるという問題がある。

本研究では、畳み込みニューラルネットワークを用いたタッチの類似性を考慮したイラスト検索 [3] において入力データやネットワークの構造の一部を変更することで精度の向上を目指す。

2 畳み込みニューラルネットワークを用いたタッチの類似性を考慮したイラスト検索システム

2.1 構造

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を図 1 に示す。提案システムでは、(1) 2 値化画像、(2) 彩度・明度のヒストグラム、(3) 色相のヒストグラムを畳み込みニューラルネットワークへの入力として用いる。この畳み込みニューラルネットワークの下位の部分はそれぞれの入力に対して独立した構造になっており、上位の全結合層で 3 つの情報を統合し、出力層のソフトマックス層で入力の属するイラストのタッチに応じたグループに対応するニューロンが発火するように学習が行われる。(1) の 2 値化画像を入力とする部分は、プーリング層を 2 つ重ねた構造になっている。1 つ目のプーリング層では局所的な受容野に対して平均プーリングを行い、受容野内に含まれる輪郭線の要素の割合を抽出する。2 つ目のプーリング層は前のプーリング層全体を受容野として持つようにし、最大プーリングを行う。(2) の彩度・明度のヒストグラムを入力とする部分は、畳み込み層などを重ねた一般的な構造を持つ。ヒストグラムを入力として用いることで、画像内に含まれる形状の情報ではなく、含まれている色の彩度や明度の傾向を抽出する。(3) の色相のヒストグラムを入力とする部分も畳み込み層などを重ねた一般的な構造を持つ。従来システム [3] では、これに対応する部分に RGB のヒストグラムを用いている。しかしながら R, G, B それぞれのヒストグラムを用いているため、画像中に含まれる色が異なってもヒストグラムが一致してしまうことがあるという問題がある。ここで知りたい色の使われ方の傾向は、青っぽい色が使われていないといった大まかなものであり、厳密に青 (RGB で (0, 255, 0)) という色が使われているかどうかではないため、提案システムでは RGB ではなく、色相のヒストグラムを用いることとしている。なお、色相だけでは色を特定することができないが (2) の部分で彩度と明度の値を用いているため、(3) の部分では色相のみを用いる。

Artwork Retrieval based on Similarity of Touch using Convolutional Neural Network – Improvement of Accuracy Focusing on Appearance Frequency of Hue –
Tomoya Niizeki and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

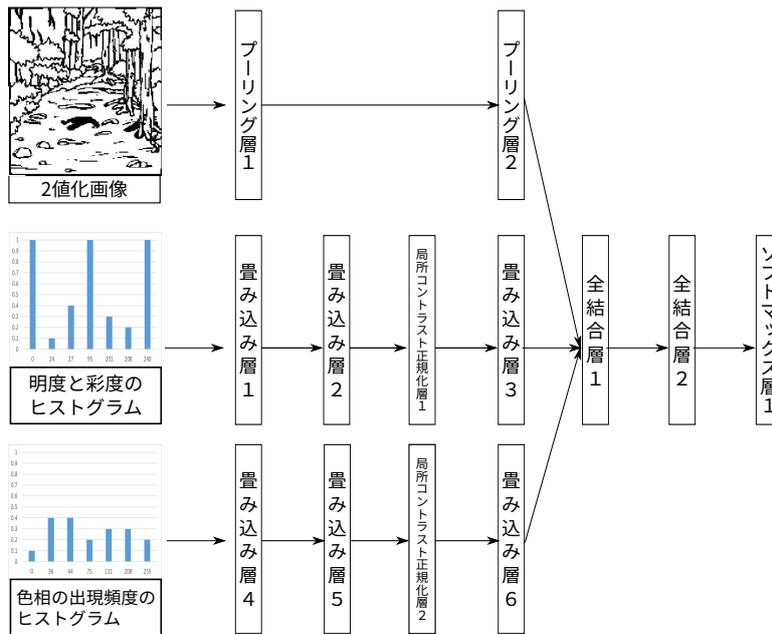


図 1: 提案システムで用いる畳み込みニューラルネットワークの構造

2.2 学習

提案システムでは、学習データの画像の彩度・明度の値、彩度・明度のヒストグラム、RGBのヒストグラムを入力とし、その画像が各クラスに属する確率を出力するような分類問題として学習を行う。ここではタッチの類似した画像を1つのグループとして扱い、1枚の画像は1つのグループにのみ属するものとする。誤差関数としては交差エントロピー誤差関数を用い、確率的勾配降下法により学習を行う。また、学習済みの畳み込みニューラルネットワークを用いて学習データの画像に対して特徴ベクトルを作成し、データベースに蓄積する。このシステムでは、出力層の手前の全結合層のニューロンの出力ベクトルを大きさが1になるように正規化したものを特徴ベクトルとして用いる。

2.3 検索

提案システムにおいて検索を行う場合には、学習済みの畳み込みニューラルネットワークに検索キーとして用いる画像のデータを入力し、出力層の手前の全結合層の出力から特徴ベクトルを生成する。検索キーの特徴ベクトルとのユークリッド距離が閾値以下になるようなデータベース中のすべての画像を検索結果として出力する。

3 計算機実験

計算機実験を行い、提案システムにおいて従来システム [2][3] よりも高い精度で検索が行えることを確認

した。

参考文献

- [1] H. Kawai and Y. Osana : “Search accuracy improvement in artwork retrieval based on similarity of touch,” Proceedings of 18th IEEE Workshop on Nonlinear Dynamics of Electronic Systems, Como, 2015.
- [2] T. Fujita and Y. Osana : “Artwork retrieval based on similarity of touch using convolutional neural network,” Proceedings of International Conference on Artificial Neural Networks, Rhodes, 2018.
- [3] D. Mizuno and Y. Osana : “Artwork retrieval based on similarity of touch using convolutional neural network – improving accuracy focusing on the frequency of color appearance –,” Proceedings of International Symposium on Nonlinear Theory and its Applications, Kuala Lumpur, 2019.
- [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.