

## 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

藤田賢之 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

我々が普段目にしているイラストには様々なタッチのものがある。資料などを作成する際にイラストを使用する場合にはタッチの統一されたイラストが使用したいという要望がある。既存のイラスト集などではイラストのタッチごとに分けられていることはほとんどなく、自分の欲しい内容であり、かつタッチの類似したイラストを探すのは難しい。キーワードのみで行う画像検索ではイラストのタッチを正しく指定することは困難であり、イラストによってはタッチを表すような名称が存在しない可能性もある。そのような場合には、イラストを入力することで類似したタッチのイラストを検索できるシステムが必要になる。

タッチの類似性を考慮したイラストの検索を行える手法として不応生を有する自己組織特徴マップを用いたシステム [1] が提案されている。このシステムでは、画像から彩度、明度、輪郭線などの特徴量を抽出し、それを不応性を有する自己組織化特徴マップ [2] に入力することでタッチの類似したイラストの検索を実現している。また、一方で、畳み込みニューラルネットワークを用いた画像認識に関する研究 [3] も盛んに行われている。畳み込みニューラルネットワーク [4] は Deep Learning の代表的な手法であり、画像認識や音声認識などにおいて従来の手法よりも優れた性能を示すとして注目されている。この手法では、従来の画像認識の手法とは違い、特徴量を事前に抽出する必要がなく、入力された学習データから色や形状などの特徴を学習により自動的に抽出することができる。

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を実現する。

### 2 畳み込みニューラルネットワークを用いた画像検索

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、はじめにパターン分類として畳み込みニューラルネットワークの学習を行う [5]。検索は、学習を行ったあとの畳み込みニューラルネットワークにおける出力層の直前の全結合層の出力を利用して行う。出力層の直前の全結合層の出力を利用して行う。出力層の直前の全結合層の出力をここでは特徴ベクトルと呼ぶものとする。この特徴ベクトルの値は、全結合層の出力をそのまま用いることも、2 値化したものを用いることもある。画像と対応づけられている特徴ベクトルをデータベースに蓄えておく。検索を行う際には、検索キーの特徴ベクトルとデータベース内の特徴ベクトルとを比較し、それらが類似していれば検索結果として特徴ベクトルに対応する画像が出力される。特徴ベクトルが類似しているかはユークリッド距離により判定する。

### 3 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

提案システムでは、タッチの類似性に基づいてカテゴリーに分類できるように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生成した特徴ベクトルを利用して検索を行う。

#### 3.1 構造

一般に畳み込みニューラルネットワークを用いて画像検索や画像認識を行う場合には、画像をそのままもしくはグレースケール画像に変換したものを入力し、色や形状の特徴を自動的に抽出する。それに対し、提案システムでは、形状や色そのものの特徴が類似している画像を検索したいのではなく、彩度や明度、輪郭線の太さなどの傾向が類似している画像を検索したい。そこで、提案システムでは、RGB 画像を HSV(色相・彩度・明度)に変換したものを特徴として用いる。ただし、色相が類似している必要はないため、彩度と明度のみを使用する。また、彩度と明度の情報をそのまま

Artwork Retrieval based on Similarity of Touch using Convolutional Neural Network  
Takayuki Fujita and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

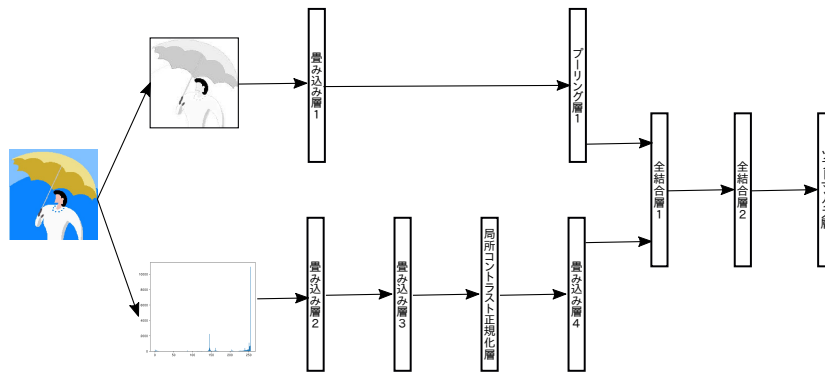


図 1: 提案システムで用いる畳み込みニューラルネットワークの構造

入力してしまうと、形状情報の特徴が抽出されてしまう可能性が高いため、提案システムでは、彩度や明度のヒストグラムを入力データとして用いる。また、輪郭線の太さなどの情報を得るためには形状情報を保った状態の入力も必要となるため、彩度や明度の値をそのまま保持したものも入力として用いる。

畳み込みニューラルネットワークも、一般的な構造のものではなく、図 1 のような構造のものを用いる。彩度や明度のヒストグラムを入力データとした部分に関しては従来の畳み込みニューラルネットワークと同様の構造で処理する。彩度や明度の値を入力とする部分では、畳み込み層の後ろに 1 つだけプーリング層を挿入する。この畳み込み層では、入力データから重みの表すフィルタによって輪郭線などの情報が抽出される。畳み込み層の出力に対し、フィルタに対応するチャンネルごとに、チャンネル全体を受容野としてプーリング層で平均プーリングを行うことで、どのような輪郭線がどのような割合で含まれているかといった特徴を抽出することが可能となる。

### 3.2 検索結果の出力

提案システムでは、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と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。閾値の値は、学習データに対する検索精度が高くなるように設定する。

## 4 計算機実験

500 枚の画像に対する提案システムと従来のシステム [1] における検索精度を表 1 示す。なお、従来のシステムでは 500 枚の画像をすべて学習させた状態で

表 1: 検索精度の比較

	再現率	適合率	F 値
従来システム [1]	0.443387	0.454866	0.449053
提案システム	0.702800	0.874346	0.779243
CNN (未学習)	0.551333	0.329679	0.412623
提案システム (未学習)	0.617111	0.577278	0.596530

の検索精度しか調べられていないが、表 1 ではテスト画像を除いたデータのみを用いて学習を行い、未学習のデータに対して検索を行ったときの精度も示してある。また、CNN は、通常の畳み込みニューラルネットワークに画像そのものを入力したシステムにおける検索精度である。この結果より、提案システムでは、従来のシステムや通常の畳み込みニューラルネットワークに画像そのものを入力したシステムよりも高い精度で検索が行えていることが分かる。

## 参考文献

- [1] 川端隼矢, 長名優子: “タッチの類似性を考慮したイラスト検索の精度向上 –特徴量と類似度の算出方法の変更–,” 情報処理学会 第 79 回全国大会, 2017.
- [2] H. Mogami, M. Otake, N. Kouno and Y. Osana: “Self-organizing map with refractoriness and its application to image retrieval,” Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Vancouver, 2006.
- [3] A. Krizhevsky, I. Sutsuker and G. E. Hinton: “ImageNet classification with deep convolutional neural networks,” In Advances in NIPS, pp.1097–1105, 2012.
- [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] C. H. Kuo, Y. H. Chou and P. C. Chang: “Using deep convolutional neural networks for image retrieval,” Proceedings of IS&T International Symposium on Electronic Imaging 2016 Visual Information Processing and Communication VII, No.6, pp.1–6, 2016.