

畳み込みニューラルネットワークを用いたタッチを考慮したキャラクターイラストの検索

松井大樹 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

我々が普段利用している画像検索機能を使用してキャラクターイラストの検索を行うと、検索に利用したキャラクターイラストと同じ色が使用されているイラストや、イラスト全体の構図が似ているイラストが検索されることになる。しかしながら、キャラクターイラストを検索するような場合には、色や構図などが似ているイラストを探したいのではなく、輪郭線の描き方や色の塗り方などのイラストのタッチが似ているイラストや同じ作者によって描かれたイラストを検索したいことが多い。同じ作者によって描かれたキャラクターイラストでは、目の描き方や塗り方、形状に共通した特徴があることが多い。

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、畳み込みニューラルネットワークを用いてタッチの類似性を考慮してイラスト検索を行う手法 [1] が提案されている。この手法では、彩度や明度、輪郭線の太さなどの全体の傾向が類似している画像の検索を行うことができる。キャラクターイラストの検索では、目などの特徴のある部分に着目してイラストの検索を行いたいことがある。しかし、文献 [1] の手法ではキャラクターの目などのイラストの一部に着目した検索を行うことはできない。

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を考慮したキャラクターイラストの検索を提案する。提案システムは、畳み込みニューラルネットワークを用いたタッチを考慮したイラスト検索 [1] に基づいたシステムである。提案システムでは、画像内の特定のオブジェクトを抽出する手法である You Only Look Once (YOLO)[2] を用いてキャラクターイラストから目の領域を抽出し、それも入力として利用する。

2 畳み込みニューラルネットワークを用いたタッチを考慮したキャラクターイラストの検索

提案システムでは、検索キーとして用いる画像の彩度・明度、彩度と明度のヒストグラム、YOLO[2] によって抽出した目の領域の RGB 画像を畳み込みニューラルネットワークに入力し、出力層の手前の全結合層の出力を特徴ベクトルとして、タッチを考慮したキャラクターイラストの検索を行う。

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。提案システムは、文献 [1] のシステムに基づいたシステムであり、画像の彩度・明度、彩度と明度のヒストグラムを入力とする部分は、文献 [1] のシステムと同様の構造を持つ畳み込みニューラルネットワークを用いる。

彩度や明度の値を与える部分では、畳み込み層の重みを表すフィルタによって輪郭線などの情報が抽出されることを想定している。提案システムでは、輪郭線の太さなどの情報は検索の際に利用したいが、画像のどこにどのような輪郭線があるかというような情報は必要としない。畳み込み層の出力は各フィルタに対応するチャンネルを持っているが、プーリング層においてそれぞれのチャンネル全体を受容野として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になる。

一般的な畳み込みニューラルネットワークでは、事前に抽出した特徴量を入力として用いることはないが、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として検索を行った場合、未学習のデータに対して十分な検索精度を得られない。そこで、彩度と明度のヒストグラムを事前に抽出し、それを特徴量として用いるものとしている。彩度と明度のヒストグラムの値を入力する部分では、畳み込み層を重ねた構造になっている。

目の領域を与える部分では、YOLO[2] を用いて抽

Character Illustration Retrieval considering Touch with Convolutional Neural Network
Hiroki Matsui and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

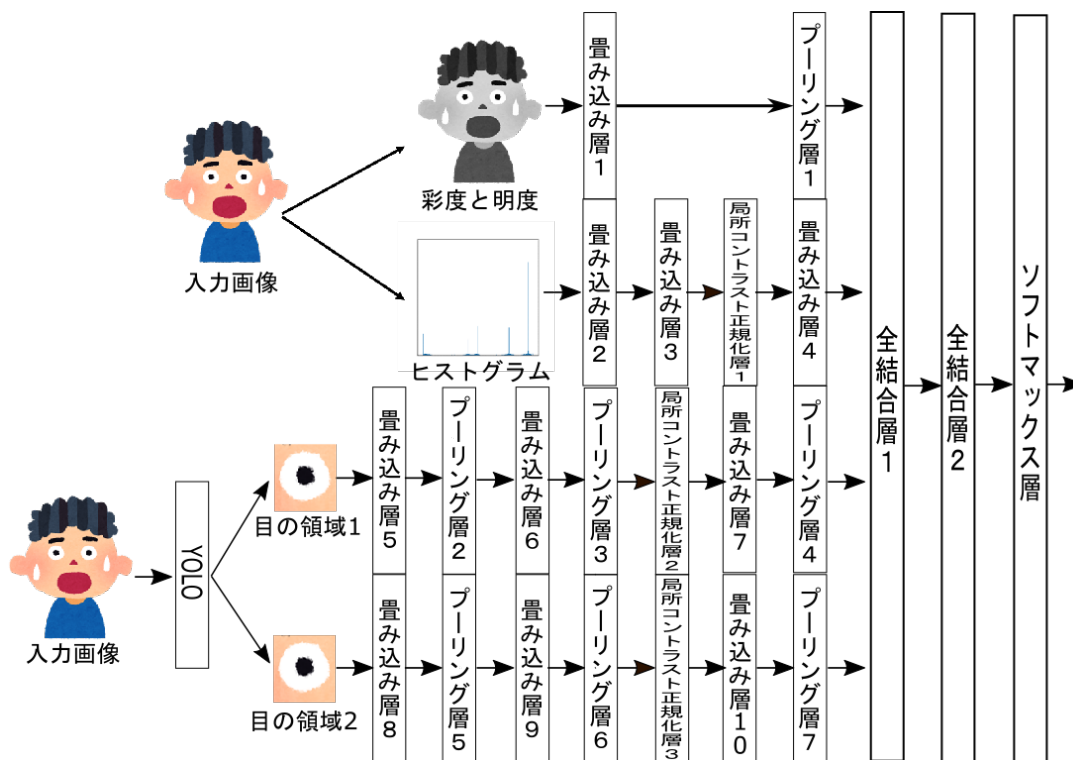


図 1: 提案システムで用いる畳み込みニューラルネットワークの構造

出した目の領域の RGB 画像を入力として用いる。目に関しては形状や色などが類似しているかどうかを判断したいので、この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に処理を行う。

最後に 2 つの全結合層を通して 3 つの入力に関する情報を統合し、最終的にはソフトマックス層において入力が各クラスに属する確率が出力される。提案システムでは、同じ作者のイラストを 1 つのクラスとして扱うものとする。

2.2 検索

提案システムでは、タッチの類似性を考慮して作者ごとの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。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。

3 計算機実験

500 枚 (20 枚 × 25 グループ) のキャラクタイラストの画像を学習させ、実験を行った。提案システムにお

いてすべてのデータを学習データとして用いて学習を行ったときの F 値は 0.690 であった。それに対し、目の領域のデータを用いない従来システム [1] において同様の実験を行ったところ F 値は 0.245 であった。また、未学習のデータに対して同様の実験を行ったところ、提案システムでの F 値は 0.447、従来システム [1] での F 値は 0.141 であった。いずれの場合にも提案システムにおいて従来システムよりも高い精度で検索が行えることが分かった。

参考文献

- [1] 藤田賢之, 長名優子 : “タッチの類似性を考慮したイラスト検索,” 情報処理学会第 80 回全国大会, 2018.
- [2] J. Redmon, S. Divvala, R. Girshick and A. Farhadi : “You Only Look Once: unified, real-time object detection,” <https://arxiv.org/abs/1506.02640>
- [3] 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.