

イラスト検索のためのキャラクタ属性に基づく イラストベクトル表現生成

小山 雄大[†] 福原 知宏^{†‡} 山田 剛一[†] 阿倍 博信[†] 増田 英孝[†]
東京電機大学[†] マルティスープ株式会社[‡]

1. はじめに

イラスト投稿系の SNS においてユーザが投稿したイラストを検索するための機能としてタグによる検索機能がある。

しかし、これにはまずユーザが手動でイラストに対して適切なタグを付与してあるという前提が必要となる。アニメやマンガのキャラクタのイラストに着目してみると人手でタグ付けを行うという仕様上、描かれているキャラクタの名前やそのキャラクタが登場する作品名等の具体的な情報のタグ付けは行われるが、髪型・服装等の自明な属性や表情・姿勢等の抽象的な細かい属性についてのタグ付けは行われない傾向がある。そのためユーザは特定の属性を持つキャラクタのイラストのみを閲覧したい場合でもタグが付与されていないため検索することができない。

また、従来のタグ付け手法ではタグが付いているかないかという情報でしか属性を表現できないためキャラクタの特徴を十分に表せているとは言えない。よってタグを用いたキーワード検索では本当にユーザが求めているイラストを絞り込むことが難しいという問題がある。

そこで、本研究では畳み込みニューラルネットワーク(CNN)を用いてイラストに描かれたキャラクタの特徴を抽出することで属性の情報をもとにイラストをベクトル化する方法を提案する。例えば髪型であればロングヘア、ショートヘア、ツインテール等のイラストの特徴を学習させることで、キャラクタの髪型におけるイラストの特徴量を得ることができる。これを服装等の他の普遍的な属性についても行うことでキャラクタ全体がどのような見えているかを属性に基づくイラストの特徴量を成分とする n 次元のベクトルで表現することができる。これによりタグ付けを行うよりも多くの情報をイラストに付与することが可能になる。

2. 関連研究

イラストのベクトル表現の生成を目指した研究として、Saito らによる *Illustration2Vec: A Semantic Vector Representation of Illustrations*[1]がある。この研究ではキャラクタ名、作品名、キャラクタ属性、背景等のイラストに付与されるあらゆるタグ情報を用いてイラスト全体に対するベクトル表現を生成し、類似画像検索やセマンティックモーフィングを実現している。

本研究ではベクトル表現を生成するために抽出する特徴をイラストに描かれているキャラクタの一般的な属性に絞りキャラクタ属性に基づくベクトル表現を生成する。これにより従来のタグ付け手法では付与されることが少なかったキャラクタ属性の情報をベクトル表現という形で自動的に付与することが可能になりキャラクタの属性を用いた検索を実現する。

3. 提案手法

本研究はイラストにおけるキャラクタ属性の特徴を抽出しベクトル表現を生成することでキャラクタ属性を用いたイラスト検索を行うことを目的とする。イラスト検索に用いるベクトル表現を自動的に生成するための CNN による特徴抽出器を提案する。イラストのベクトル表現の生成の流れを以下に示す(図 1)。

1. 予め作成したキャラクタ属性のイラストデータセットを用いて CNN を訓練する。
2. ベクトル表現を生成したいイラスト I を訓練された CNN に入力する。
3. 入力されたイラスト I に対して各属性を持つかどうかの予測値 p を出力する。
4. 予測値を出力する前の中間層の出力をイラスト I のベクトル表現とする。

得られたイラスト I のベクトル表現は学習したキャラクタ属性に基づく特徴量を成分とする。この特徴量を用いてベクトル間の距離をもとに検索結果をランキングする等の従来のキーワード検索では行えなかったベクトルを用いた検索が期待できる。

Generation Method of Vector Representation of Illustration Based on Character Attributes for Illustration Search

[†]Koyama Yuta, ^{†‡}Fukuhara Tomohiro,

[†]Yamada Koichi, [†]Abe Hironobu, [†]Masuda Hidetaka

Tokyo Denki University ([†]), MULTISOUP CO., LTD. ([‡])

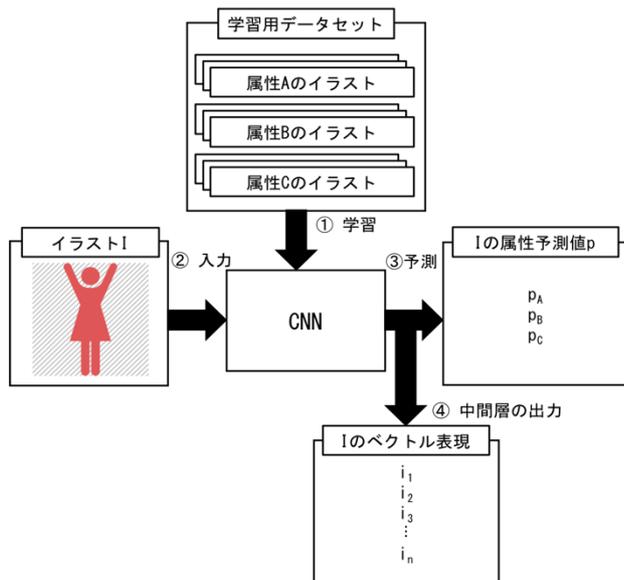


図 1. ベクトル表現生成の流れ

4. ベクトル表現の生成

4.1 CNN の構築

キャラクタ属性をラベルとするイラストのデータセットを用いて各キャラクタ属性の特徴を抽出可能な CNN を訓練する. CNN は事前学習済み VGG16 モデルをベースに全結合層を各ラベルがそれぞれ独立に付与される確率を出力するものに置き換え転移学習を行う. その後入力されたイラストに対して確率の出力を行う前の中間層の出力を取得しそれをイラストのベクトル表現とする.

イラストデータセットは他のイラスト投稿サイト等から引用されたイラストに対してユーザがキャラクタ属性や背景等の情報まで細かくタグ付けを行うことのできる Danbooru[2]から約 290 万のイラストとタグのデータを取得し, その中から白背景で女性キャラクターのみが描かれたものを利用する. 学習に用いるキャラクタ属性は「髪型」についての属性のみとする. それらの属性に関するタグの中から 1000 個以上のイラストに付与されているタグをラベルとし, 各ラベルに対し 1000~10000 個のイラストを抽出したデータセットを作成する.

4.2 評価

作成したデータセットで訓練を行った CNN によるタグ分類器に対してテストデータのイラストを入力し正しくタグを判定できるかを評価する. 各タグのイラストに対する予測値トップ 3 に正解タグが含まれている場合を正解としたときの精度を表 1 に示す.

表 1. タグと精度

タグ	精度	タグ	精度
one_side_up	0.234	ahoge	0.715
twintails	0.790	wavy_hair	0.218
short_hair	0.777	antenna_hair	0.333
sidelocks	0.586	half_updo	0.000
hair_intakes	0.023	long_hair	0.745
two_side_up	0.520	hair_flaps	0.356
drill_hair	0.285	hair_bun	0.008
tied_long_hair	0.000	very_long_hair	0.678
ponytail	0.593	braid	0.487

4.3 データセットの改善

評価の結果キャラクタ属性によって精度に大きく差が出ることが確認できた. 「twintails」等は十分な精度が得られたが「half_updo」等は不十分であった. その理由として髪結び等が関係する複雑な髪型であることや, 同じ髪型でも多様性の高いものがあるためだと考えられる.

また, 今回は Danbooru データセットの中からタグの有無を用いて機械的に作成したが今後はこの結果をもとに人手によるラベル付けを行ったデータセットを作成していく必要がある. 現在マンガキャラの髪型資料集[3]によるヘアスタイルの分類をもとに「ロングヘア」「ミディアムヘア」「ショートヘア」の大分類から「ストレート」「パーマ」等に細かく分類する体系づけられたラベル付けを行っている.

5. まとめ

本研究ではイラスト検索を行うためのキャラクタ属性の特徴を CNN によって抽出し, イラストのキャラクタ属性に基づくベクトル表現を生成した.

実験の結果, 髪型に関するキャラクタ属性では同じ属性である場合でも髪型自体のバリエーションやキャラクタのポーズ等によってイラスト全体としての特徴が複雑化してしまうため上手く特徴の抽出ができないことが分かった. そのため髪型・ポーズ等についても条件を設け精査されたデータセットを作成する必要がある.

参考文献

- [1] Masaki Saito, Yusuke Matsui, “Illustration2Vec: A Semantic Vector Representation of Illustrations,” SA '15 SIGGRAPH Asia 2015 Technical Briefs Article No. 5(2015).
- [2] Danbooru, <https://danbooru.donmai.us/>
- [3] アミューズメントメディア総合学院, マンガキャラの髪型資料集, 廣済堂出版(2017).