

キャラクターの顔領域の情報を用いたイラスト作者推定

笠松慎之介† 櫻井義尚‡

† 明治大学大学院先端数理科学研究科 ‡ 明治大学総合数理学部

1. はじめに

今日のインターネット上では様々な情報やデータが公開されており、その中の一つにイラストがある。多くのイラストは Pixiv¹などのイラスト投稿サイトを通じ公開されるが、近年では Twitter などの SNS に投稿されることも珍しくない。近年ではイラストの作者以外が投稿を行うケースもあり、イラストの作者情報が欠落した発信も存在している。イラストの作者について手がかりがない場合、作者への到達は困難である。

しかし、人間は訓練を行うことで作者の推定が可能である。この時判断の基準の一つに作者の作風という概念がある。作風とは作者の作品に共通する雰囲気や指向である。作風が強く表れるとされる領域は主に顔であり、同キャラクターについて異なる作者が描いた場合、その作風によって描かれたキャラクターから受ける印象は大きく異なることもある。

そこで本研究ではイラスト群から顔の領域の情報を抽出したデータから、作者情報のあるイラストを機械学習によって学習し、作者情報のないイラストの作者の推定を行うことを目的とする分類モデルを作成し評価を行った。

2. 関連研究

藤澤らはキャラクターの目のパーツに注目したイラスト作者の推定の手法を提案している[1]。イラスト内にほぼ必ず存在し、作品の違いの影響を受けにくい箇所を使用することで作者を推定する。識別実験では再現率平均 35%だが、作者によっては 90%の再現率があった。この研究は目のみを使用している点と、色情報について触れられていない点から本研究とは異なるが、目的は同じである。

また、藤澤らは色の出現頻度から作品ジャンルを分類する手法も提案している[2]。色の出現頻度を用いたヒストグラムに対し閾値以下の頻度の色から特徴を抽出し、Support Vector Machine によって学習した学習器を用いて対照的な作風の分類を行う。この研究では色情報のみを取り扱っている点と、作品ジャンルまでの分類にと

どまっており、作者の同定までには至っていない点で異なる。

松井らは畳み込みニューラルネットワーク (Convolutional Neural Network, 以下 CNN と記載) を用いたキャラクターイラストの検索手法について提案している[3]。単純に CNN を用いるだけでは精度が出ないという問題を解決すべく、イラスト内の目の部分とイラストの彩度と明度、RGB ヒストグラムを別々に入力することで精度の向上に成功している。目の領域を採用しなかった従来のモデルが未学習データに対し F 値 0.141 に対しこの研究のモデルでは F 値 0.447 だった。この研究はイラスト検索精度の向上をさせるため、作者ごとにラベル分けをして特徴ベクトルを作成している。この過程が本研究の目的と一致しているものの、学習過程で目以外の領域について形質的情報を使用していない点が異なっている。

3. 実験手法

3.1. 実験手法の概要

本研究では作者が既知であるイラスト群から顔領域を抽出し、そのデータから学習器を作成することで未知のイラストの作者の推定をすることを目的とする。

3.2. 手順

判定したい作者のイラストを、作者情報付きで収集するため Pixiv で公開されているイラストを用いてデータセットを作成した。本実験で使用したイラストは、フルカラーのイラストであること、漫画形式でないこと、両目の位置がわかる構図であること、デフォルメされていないこと、以上の四点を満たした各イラストに対し、顔認識を行い得られた画像群から顔全体が入っていないなどのノイズを除去してデータセットを作成した。各作者につき訓練用の枚数 180 枚、検証用画像 20 枚、テスト用の枚数 15 枚の計 215 枚を 23 作者分用意し計 4945 枚を使用した。訓練用データには Augmentation として水平反転と左右 45° の回転を用いて学習用のバッチデータを作成した。このデータセットを次項で説明するモデルへ入力し、評価を行った。

Illustration author estimation using information on the character's face area

Shinnosuke Kasamatsu† Yositaka Sakurai‡

† Graduate School of Advanced Mathematical Sciences, Meiji University

‡ School of Interdisciplinary Mathematical Sciences, Meiji University

3.3. 使用したモデル

本研究では VGG16 を fine-tuning して学習器を作成した。モデルに追加した層について図 1 に示す。

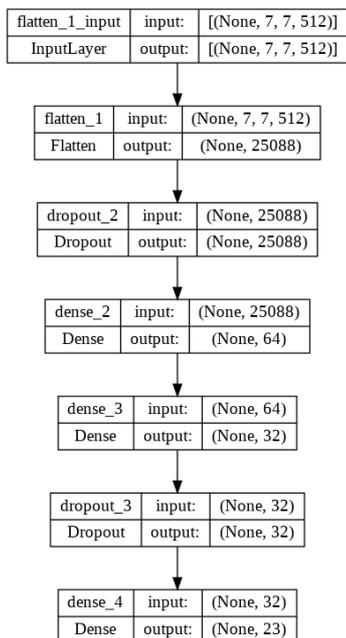


図 1 VGG16 に追加した層

4. 実験

epoch 数 20 で学習を行った。学習結果を図 2、図 4 で示す。

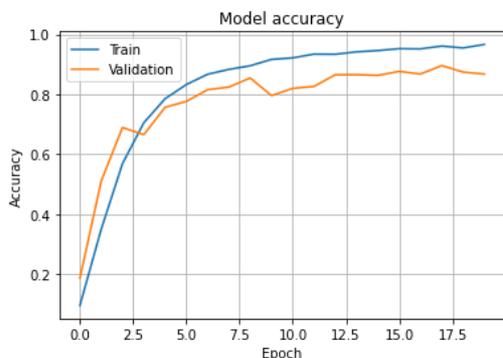


図 2 VGG16 で学習したモデルの Accuracy

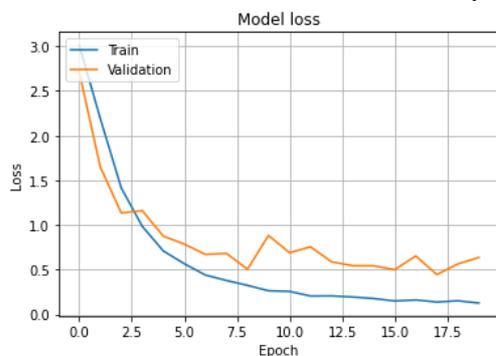


図 3 VGG16 で学習したモデルの Loss

テスト結果は Loss196.052、Accuracy が 0.557 であった。

5. 考察

4 章の結果から、顔領域を用いた作者の推定は一定程度効果がある事が判明した。今後の課題として Data Augmentation 手法の検討や、各モデルの特徴を踏まえた fine tuning 時の層の追加やパラメータ指定などがあげられる。

6. 参考文献

[1] 藤澤日明, 松本和幸, 吉田稔, 北研二, “キャラクターの顔を用いたイラスト画像の作者識別”, HCG シンポジウム 2013, vol.HCG2013-I-2-1, pp.523-530, 2013 年 12 月.
 [2] Akira Fujisawa, Kazuyuki Matsumoto, Minoru Yoshida, Kenji Kita, “An Illustration Image Classification Focusing on Infrequent Colors” International Journal of Advanced Intelligence Volume 8, Number 1, pp.72-83, May, 2016.
 [3] 松井大樹, 長名優子, “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” 情報処理学会第 82 回全国大会, 2-543 - 2-544 2020 年.

¹ イラストコミュニケーションサービス pixiv : <https://www.pixiv.net/>