

マンガページ画像からのキャラクタ顔検出に関する 検討

添田, 碧人
九州大学工学部電気情報工学科

伊東, 栄典
九州大学情報基盤研究開発センター

<https://hdl.handle.net/2324/7329572>

出版情報 : Record of Joint Conference of Electrical and Electronics Engineers in Kyushu. 77 (09-1P-09), pp.186-187, 2024-09-19. Committee of Joint Conference of Electrical, Electronics and Information Engineers in Kyushu

バージョン :

権利関係 : © 2024 Committee of Joint Conference of Electrical, Electronics and Information Engineers in Kyushu



マンガページ画像からのキャラクタ顔検出に関する検討

添田 碧人* 伊東 栄典**

(九州大学 *工学部電気情報工学科 **情報基盤研究開発センター)

* soeda.aoto.467@s.kyushu-u.ac.jp, ** ito.eisuke.523@m.kyushu-u.ac.jp

1 はじめに

近年、著作権を無視してマンガ画像を提供する海賊版マンガサイトや、著作権を無視した最新話マンガのスライド動画の投稿が問題になっている。我々は、マンガの画像からマンガのタイトル・巻数・話数を推定する判別手法を研究している [1]。判別精度が十分で、かつ処理速度が高速ならば、迷惑メールフィルタのような海賊版マンガフィルタが実現でき、海賊版利用の抑止となる。マンガのスライド動画公開に対しても、著作権者への利益還元が実現可能であろう。

本稿では、マンガのページ画像から登場キャラクタの顔部分を検出する手法について検討する。マンガの原作を多数作成した小池一夫氏が述べる [2] ように、マンガにおいて登場キャラクタは非常に重要である。画像からキャラクタを判別できれば、引用された画像がどのマンガを利用したものかを高精度で判別できる。与えられたマンガページ画像に含まれる登場キャラクタ抽出のために、マンガページ画像からのキャラクタ顔部分抽出を検討した。本稿では、用いたデータと、適用した手法、および検出精度について報告する。

2 従来研究

マンガ画像からキャラクタの顔部分を検出する研究は現在様々な手法を下に行われている。マンガ描写はその著者により大きく異なる。同じ作品内のキャラクターでもデフォルメされてその見た目が異なる場合もあるので、そのようなキャラクターの描写の変化に柔軟に対応できる顔検出器の作成が望まれる。過去研究では陳らが幾何学的変化に強く対応できる HOG(Histograms of Oriented Gradients) 特徴量と AdaBoost を利用したマンガキャラクター検出を提案した [8]。柳沢らは DPM(Deformable Part Model), R-CNN での顔検出や、Faster R-CNN を用いたキャラクターおよびコマ、フキダシ、文字列の検出の有効性を示した [9][10]。

本研究では、Faster R-CNN と公開されている 2 つの 2 次元顔検出分類器を用いて、マンガページからキャラクターの顔の位置情報を検出するモデルの作成を目指す。

3 用いたデータ

本稿の実験では、Manga109 データセットと筆者らが独自取得した電子マンガのページ画像を用いた。Manga109 データは東京大学の相澤ら [3, 4] により作成されたデータセットである。109 冊のマンガのページ画像と共に、XML 形式の注釈 (annotation) が付与されている。各ページに登場するキャラクタも、顔部分・体部分・キャラクタ名が注釈で提供されている。

また、自分が所有する電子マンガのページ画像もキャプチャで取得した。こちらは次節で述べるキャラクタ顔部分特定手法が、Manga109 データ以外でも適切に機能するのかの確認にのみ用いた。

4 キャラクタの顔部分特定手法

ここでは、今回適用した 3 種類の顔検出手法について述べる。2 種類は既に公開されている手法で、マンガやアニメなどの二次元画像からの顔画像検出に特化した分類器である。3 つ目の手法は、既存手法を用いて我々が独自に学

習させた分類モデルである。

4.1 animeface (Viola-Jones)

このモデルは OpenCV の Viola-Jones 物体検出フレームワークに基づいたカスケード分類器の一種で、nagadomi 氏により作成・提供されている [6]。Viola-Jones 物体検出フレームワークは、リアルタイムでの物体検出に広く利用されている手法である。単純な特徴と効率的な学習アルゴリズムを組み合わせることで、高速かつ高精度な物体検出を実現している。

4.2 animeface (CascadeClassifier)

animeface は、シンプルで高速なアニメ顔検出に適したモジュールである。このモデルは OpenCV の CascadeClassifier をベースに開発されており、アニメ顔の特徴を学習したカスケード分類器を用いて、画像からアニメ顔を検出する [7]。

4.3 Faster R-CNN (Detectron2)

Faster R-CNN は、画像内の物体を検出する真相学習モデルの一種で、特にリアルタイムな物体検出に置いて高い性能を発揮する。このモデルは、深層学習の中でも特に畳み込みニューラルネットワーク (CNN) を活用した代表的な手法である。CNN の高い特徴抽出能力と、Region Proposal Network による効率的な物体候補領域の生成により高精度が実現されている。

Detectron2 は、Facebook AI Research が開発した、PyTorch ベースの物体検出ライブラリである。

5 実験

3 種類の検出器について、マンガキャラクターに対する検出率を比較する。Manga109 データセットの images を入力として用いて、出力として得られたキャラクターの顔の位置情報と、annotation に記載されたキャラクターの顔の位置情報を比較し、その重なり具合 (IoU: Intersection over Union) が指定値を超えていたものを検出成功例とする。入力の画像例と検出の様子を以下に示す。なお、表では、animeface (Viola-Jones) を af-VJ, animeface (CascadeClassifier) を af-CC, Faster R-CNN を F.R-CNN と表す。

表 1: 判別精度 (IoU=0.5)

Method	Prec.	Rec.	F-s.	AP.
af-VJ	0.9813	0.1502	0.2648	0.1865
af-CC	0.9927	0.3387	0.5051	0.4937
F.R-CNN	0.7661	0.8306	0.7971	0.8112

公開されていた検出器はどちらも Precision は高い。しかし、Recall が低い傾向が見られた。これは検出したものはほぼ全て顔であったが検出できた顔の割合が低い。また、Faster R-CNN では Recall, Precision とともに高い値を出している。これらの数値は図 1 の検出結果の様子からも見て取れる。

af-VJ モデルに関しては、全くページ内の顔画像を検出できていない。af-CC モデルに関しては、真正面の顔かつ顔のほとんどの部分が描写されているもののみ検出してお



(a) Manga109 データセットにおける各検出器の結果の比較 (左:af-VJ, 中央:af-CC, 右:F_R-CNN)[11]



(b) 電子マンガページにおける各検出器の結果の比較 (左:af-VJ, 中央:af-CC, 右:F_R-CNN)[11]

り、横顔やフキダシ等で顔がほとんど隠れてしまっている場合は検出できない傾向がある。F_R-CNN モデルに関しては、真正面の顔だけでなく、横顔や顔の半分がフキダシで隠れているものや顔が小さく描写されているものまで検出に成功している。

また、このモデルは顔を検出したときにその検出結果がどれくらい正確を確信度という指標で示している。今回は確信度 50 % 以上のものを使用して判別精度を算出した。この確信度を上げて性能を評価することで、検出結果例のような顔でない部分に誤反応している検出候補を消去できる。それによりさらに精度が上がると考えられる。また、Faster R-CNN には確信度以外にもパラメータが多数あるので、より強度モデルを作成するために、パラメータを変化させたとき、評価指標がどう変化するか今後検討したい。

6 おわりに

本論文では、マンガのページ画像からの登場キャラクターの顔部分特定を試みた。キャラクターの顔部分特定は、公開されている分類器よりも Faster R-CNN で学習させた分類器のほうが顔検出に適しているということが示された。

本研究の主たる目的は海賊版マンガを抑止するための、マンガ画像の自動特定である。この目的のために今後は、Faster R - CNN 以外の Detectron2 に格納された他のモデル、例えば YOLOv2 なども検討し、モデルの違いによる検出精度の変化についても調査を進める必要がある。また、キャラクター判別を用いたページ・話・巻の特定も試したい。したがって、様々な類似度や距離を利用した他の手法も検討し、より汎用性の高い検出手法の構築を目指すべきである。最終的には、自然画像やアニメ画像との比較分析を行い、より高精度で多様な環境に対応できるシステムを構築することを目指す。

参考文献

- [1] 山仲一颯, 伊東栄典: 大規模マンガ画像判別手法の検討, 情報処理学会 火の国情報シンポジウム 2024, A4-1, Mar.14, 2024.
- [2] 小池一夫: 小池一夫のキャラクター新論 ソーシャルメディアが動かすキャラクターの力, ゴマブックス, 2017.
- [3] Manga109, <http://www.manga109.org/ja/index.html> (accessed on 05-Aug-2024).
- [4] K. Aizawa, et.al.: Building a Manga Dataset “Manga109” with Annotations for Multimedia Applications, IEEE MultiMedia, vol.27, no.2, pp.8-18, 2020.
- [5] Wei-Ta Chu, Wei-Wei Li: Manga FaceNet: Face Detection in Manga based on Deep Neural Network, ICMR’17, pp.412-415, 2017.
- [6] nagadomi: lbpcascade_animeface, https://github.com/nagadomi/lbpcascade_animeface (accessed on 05-Aug-2024).
- [7] <https://github.com/opencv/opencv/tree/master/data/haarcascades> (accessed on 05-Aug-2024).
- [8] 陳 明, 柳澤 秀彰, 張 傑, 石井 大祐, 渡辺 裕: マンガにおける HOG+AdaBoost による顔検出の性能評価, 映像情報メディア学会年次大会, 17-4, Aug.31, 2014.
- [9] 柳澤 秀彰: マンガキャラクターを対象とした多視点顔検出の研究, 早稲田大学修士論文, Feb.1, 2016.
- [10] 柳澤 秀彰: ストーリー理解を目的とした漫画オブジェクトの抽出, 早稲田大学, 博士論文, 甲第 5715 号, 2019.
- [11] 赤坂アカ, 横槍メンゴ, “推しの子”, 集英社.