

深層学習を用いたマンガキャラクターの検出における 顔変形の影響評価

山下拓朗[†] 柳澤秀彰[‡] 渡辺裕^{†‡}

早稲田大学基幹理工学部情報理工学科[†] 早稲田大学大学院基幹理工学研究科[‡]

1. はじめに

近年、電子端末の普及に伴い電子書籍市場が拡大している。また、日本の電子書籍市場に占めるコミックの割合は 81%と非常に高い。このような背景からマンガデータのより効率的かつ効果的な利用について様々な研究が行われてきた。

従来のマンガ画像に関する研究では、特徴量として HOG 特徴量や SURF 特徴量などが用いられてきた。一方で近年深層学習アルゴリズムの発展による一般物体検出の精度向上に伴い、マンガに対する画像認識においても、深層学習を用いた手法が高い性能を示している[1, 2]。しかし、深層学習により得られる特徴量は、人為的に定義されないものであり、その内容は従来のものに比べ不明瞭である。

本稿では深層学習によるマンガキャラクターの顔検出について、顔パーツを変形した際の検出状況の変化について調査を行った。これにより、深層学習による検出がどのように行われるかを直感的かつ定量的に評価することを目的とする。

2. 実験

松井らにより公開されたマンガ画像データセット Manga109[3]のアノテーションデータを利用し学習した。Manga109 のマンガ画像データは学術目的の利用に限り、コマの加工を含めた学術論文への掲載利用が認められている。深層学習による一般物体検出アルゴリズムとして、Wei Liu らによって提案された Single Shot MultiBox Detector (SSD)[4]を用いる。

本実験では学習したキャラクターの顔、コマ、吹き出し、テキストの 4 クラスのうちキャラクターの顔のみを評価の対象とする。マンガ画像データから SSD による confidence の値が 1 に非常に近いキャラクターの顔画像のうち、正面を向いているものを 5 作品から 142 枚用意しこれを正

例とする。

これらの画像について、キャラクターの顔を中心にした回転と、キャラクターの目の領域を横軸および縦軸方向に平行移動を行った。

平行移動は、横軸および縦軸方向に上下左右それぞれ 2 ピクセル刻みで 150 ピクセルずつ動かし画像を作成した。回転は 3°刻みで、360°分回転し画像を作成した。変形していない元画像を図 1 に、作成した画像の例を図 2-4 に示す。図 1-4 は Manga109 のマンガ画像データを利用した。

作成した画像について confidence の値を SSD により調査した。画像の変形により、SSD による領域の検出が重なりを持ち複数もたらされる場合がある。Intersection over Union (IoU)は領域の重なり具合を示す指標であり、これを用いて、重なった複数の領域の選択を行った。

3. 実験結果

回転した画像について、回転角度と confidence の関係を図 5 に示す。本実験では SSD の学習において、回転に関するデータオーグメンテーションを行っていないため、元画像の位置から角度のずれが大きくなっていくにつれ、confidence の値は下がっていると考えられる。180°の付近で confidence の値が上がっている原因としては、目が上下反転しているものの元の画像に近い配置であること、マンガにおいて口が描かれないキャラクターもしばしばいることが考えられる。

左目を平行移動した画像について、移動距離と confidence の関係を図 6, 7 に示す。画像の平行移動に関しては、異動後の目の配置により、元の画像で隠れる部分が生じる。また、移動距離の基準と考えられる元の画像の顔領域の大きさの平均は幅 163, 高さ 164 であった。

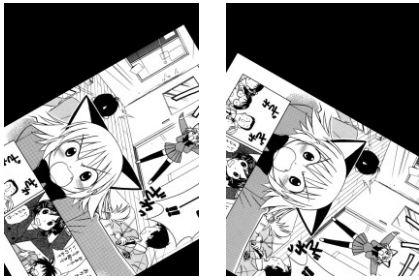


図 1. 元の画像 (ありさ²©八神 健)

Deformation Influence Evaluation of Character Face
Detection using CNN

[†]TAKURO YAMASHITA, School of Computer Science
and Engineering, Waseda University

[‡]HIDEAKI YANAGISAWA, HIROSHI WATANABE,
Graduate School of Fundamental Science and
Engineering, Waseda University

図 2. 回転した画像 (ありさ²©八神 健)図 3. 横軸平行移動した画像 (ありさ²©八神 健)図 4. 縦軸平行移動した画像 (ありさ²©八神 健)

左目を横軸方向に平行移動した画像について、正の方向への移動では、顔領域の境界と考えられる 60 ピクセル付近までは徐々に confidence の値が下がり、その後は大きな変動は見られない。負の方向への移動では、移動距離 60 から 100 ピクセル付近で confidence の値が下がっている。これは右目を隠すことが原因と考えられる。その後隠された右目が隠れない部分では confidence の値が大きくなっていることがわかる。

左目を縦軸方向に平行移動した画像について、正の方向への移動では、移動距離 20 ピクセルあたりから下がりはじめ、口付近の領域を隠す 100 ピクセルあたりまで confidence の値が下がっている。また、片目を縦軸方向に動かすと、2 つの目の位置が上下にずれた関係となることも confidence を下げる原因と考えられる。負の方向への移動では、移動距離 10 ピクセルあたりから下がりはじめ、髪の領域を隠す 60 ピクセルまで confidence は下がっている。移動距離 60 付近から confidence スコアが上がるのは、隠れていた髪の領域が表れてくること、目が顔領域の上部にあることが多く顔領域から外れた位置で多く移

動するためと考えられる。

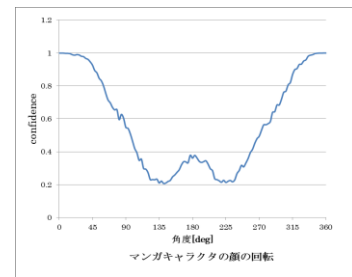


図 5. キャラクタの顔の回転

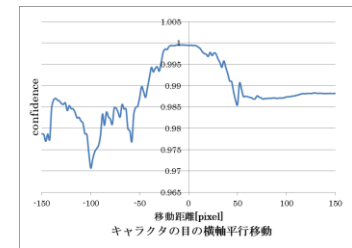


図 6. キャラクタの横軸平行移動

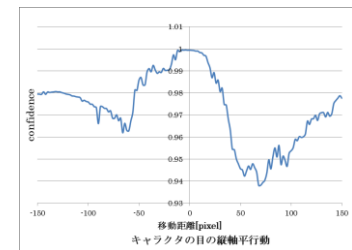


図 7. キャラクタの縦軸平行移動

4. まとめ

本稿では、深層学習を用いたマンガキャラクターの顔検出について、画像の回転と顔パーツの平行移動の影響を調査した。実験を通し、回転や移動によるパーツの位置関係の変化、隠れる顔パーツにより confidence の値が変化していく様子を確認した。

謝辞

本研究は JSPS 科研費 17K0511 の助成を受けたものである。

参考文献

- [1] 柳澤, 山下, 渡辺: “畳込みニューラルネットによるマンガオブジェクト認識メカニズムの一検討”, 電子情報通信学会, パターン認識・メディア理解研究会, PRMU2017-79, Oct. 2017
- [2] 小川, 山崎, 相澤: “漫画物体検出に向けた検出器の並列化”, FIT2017, pp 33–40, Sep. 2017
- [3] Y. Matsui, K. Ito, Y. Aramaki, A. Fujimoto, T. Ogawa, T. Yamasaki, and K. Aizawa: “Sketch-based Manga Retrieval using Manga109 Dataset,” Multimedia Tools and Applications, Springer, Vol.66, Issue 20, pp.21811–21838, Oct. 2016.
- [4] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and A. Berg: “SSD: Single Shot MultiBox Detector,” European conference on computer vision (ECCV’16), O-1A-02, Oct. 2016.