

卒業論文概要書

Summary of Bachelor's Thesis

2018/01/30

学科名 Department	情報理工	氏名 Name	山下拓朗	指導 教員 Advisor	渡辺 裕 印 Seal
研究指導 Research guidance	オーディオビジュアル情報処理研究	学籍番号 Student ID number	1w120564-5 CD		
研究題目 Title	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 Deformation Influence Evaluation of Character Face Detection using CNN				

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[1]。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。また、深層学習アルゴリズムの発展に伴い、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[2]。しかし、深層学習により得られる特徴量は、人為的に定義されず、その内容は従来のものに比べ不明瞭である。

本研究では深層学習によるマンガキャラクターの顔検出における、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. 深層学習によるマンガキャラクターの顔検出

深層学習による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)[3]を用いる。SSDは検出性能とリアルタイム性に優れている。学習については、松井らにより公開されたマンガ画像データセット Manga109[4]から作成したアノテーションデータを利用し、マンガキャラクターの顔検出を行う。

3. 深層学習福笑い

日本の伝統的な遊戯である福笑いを参考に、マンガキャラクターの輪郭にマンガキャラクターの顔パーツを配置、深層学習によるマンガキャラクター顔検出をリアルタイムに行った。本実験を通して深層学習による物体検出における、マンガキャラクターの顔パーツの配置と深層学習によるマンガキャラクターの顔検出の関係性を直感的に確認した。

4. マンガキャラクターの検出における顔変形の影響評価

深層学習によるマンガキャラクターの顔検出について、キャラクターの顔を変形した際の検出状況の変化を、定量的に評価することを目的として実験を行った。キ

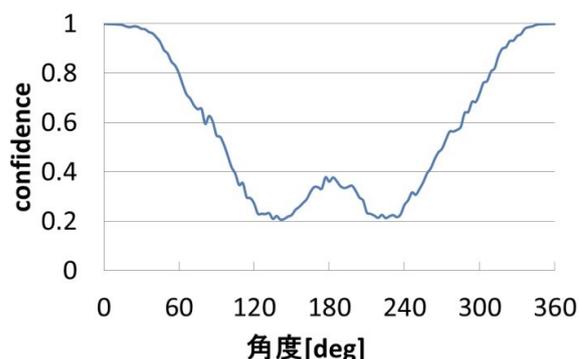


図 1: 回転した画像の検出結果

ャラクタの顔の変形として、キャラクターの顔の回転、目の領域の平行移動を行った。検出結果の例として回転の場合を図1に示す。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定量的に確認した。

5. まとめ

本研究では深層学習によるマンガキャラクターの顔検出の状況を直感的かつ定量的に評価した。課題としてキャラクターの顔変形項目を追加することや、深層学習における学習や検出に関するクラス分類の指標としての応用を検討している。

参考文献

- [1] 日本経済新聞, “電子書籍が変える読書 世界市場4年で急成長,” 日本経済新聞, <https://vdata.nikkei.com/datadiscovery/21book/>, 2018年1月19日参照。
- [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] 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.
- [4] 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.

2017 年度 卒業論文

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

Deformation Influence Evaluation of Character Face
Detection using CNN

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報理工学科

1W120564-5

山下 拓朗

目次

第 1 章 序論	1
1.1 研究の背景	1
1.2 本研究の目的	1
1.3 本論文の構成	2
第 2 章 関連研究	3
2.1 まえがき	3
2.2 MANGA109	3
2.3 物体検出における指標	3
2.4 SINGLE SHOT MULTIBOX DETECTOR (SSD)	4
2.5 むすび	5
第 3 章 予備実験	6
3.1 まえがき	6
3.2 深層学習福笑い	6
3.3 実験方法	6
3.4 実験環境	7
3.5 実験結果	7
3.6 考察	10
3.7 むすび	11
第 4 章 提案手法	12
4.1 まえがき	12
4.2 提案手法	12
4.3 実験方法	12
4.4 実験 1 の結果と考察	16
4.5 実験 2 の結果と考察	16

4.6 むすび.....	19
第5章 結論.....	20
謝辞.....	21
参考文献.....	22
図一覧.....	23
表一覧.....	24
研究業績.....	25

第1章

序論

1.1 研究の背景

近年、スマートフォンやタブレットといった電子端末の普及に伴い、電子書籍市場が拡大している。この傾向は世界的にも見られるものである。また、日本の電子書籍市場に占めるコミックの割合は 81% と非常に高く、これは他国にない特徴である[1]。クールジャパン政策において、マンガ・アニメを含む日本の魅力を展開し海外需要を獲得することが求められている[2]。電子コミックと従来のマンガにおける大きな違いはその媒体であり、従来は紙であったものが電子化されたデータとして流通することである。その中で、市場に流通する膨大なマンガデータをより効率的かつ効果的に活用することが求められている。このような背景から、マンガに対するさまざまな研究が行われてきた。

マンガに関する研究のひとつに、マンガオブジェクトの検出がある。それらの研究では、マンガ内のキャラクターやコマ、テキストといったオブジェクトを検出することで、メタデータの抽出を行っている。これら研究の応用に関しては、吹き出しからテキストを抜き出し翻訳、海外での展開に活用することや、マンガのタイトルや作者、掲載雑誌のみでなく、キャラクターの情報や台詞などによる検索の実現への活用などが考えられている。

1.2 本研究の目的

従来のマンガに関する画像認識の研究では、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[3]。しかし近年、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[4, 5]。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、その物体検出の様子は従来の特徴量を用いた手法に比べ不明瞭である。深層学習による検出の評価が可能になることによって、学習や検出を行うクラスのカテゴリの妥当性の検証や、行うデータオーグメンテーションの選択、メタデータの抽出を行う場合の根拠となり得ると考えられる。

本論文では、深層学習を用いた物体検出アルゴリズムの1つである **Single Shot MultiBox Detector (SSD)** を用いる。これをマンガについて用いることでメタデータを抽出する。このメタデータとは、テキストやコマ、吹き出し、キャラクターの顔とそれらのマンガ画像領域を対応付けたものである。本論文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、キャラクターの顔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。

1.3 本論文の構成

以下に本章以降の構成を示す。

- 第1章 本章であり、研究の背景及び目的について述べている。
- 第2章 関連研究について述べる。まず本研究で利用したマンガ画像データセットである **Manga109** について述べる。次に物体検出における指標について述べた後、一般物体検出アルゴリズムの一つである **SSD** について述べる。
- 第3章 予備実験として行った深層学習福笑いについて述べる。
- 第4章 提案手法である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について述べる。
- 第5章 本研究のまとめを述べる。

第2章

関連研究

2.1 まえがき

本章では関連研究について述べる。まず本研究で利用したマンガ画像データセットである Manga109 について述べる。次に物体検出における指標について触れた後、本研究で利用した一般物体検出アルゴリズムである SSD について述べる。

2.2 Manga109 [6]

Manga109 は松井らにより公開されたマンガ画像データセットで、日本のプロの漫画家によって描かれた 109 冊のマンガで構成されている。本研究はこのデータセットを利用した。

また、本研究では Manga109 について、手動でアノテーションを行ったものを利用した。アノテーションのラベルは、キャラクターの顔 (character) , 吹き出し (balloon) , テキスト (text) , コマ (panel) の 4 クラスであり、それぞれの画像における座標の情報を与えたものである。

2.3 物体検出における指標

FPS は Frame Per Second の略で 1 秒間に処理できるフレーム数を示し、リアルタイム性の指標となる。

検出性能の指標として再現率 R (Recall) と適合率 P (Precision) がある。検出されたデータは 4 種類に分類される。正解と検出された正解データを TP (True Positive) , 正解と検出された不正解データを FP (False Positive) , 不正解と検出された正解データを FN (False Negative) , 正解と検出された不正解データを TN (True Negative) とする。

再現率は,

$$R = \frac{TP}{TP + FN} \quad (2.1)$$

により得られる.

適合率は,

$$P = \frac{TP}{TP + FP} \quad (2.2)$$

により得られる.

平均適合率 AP (Average Precision) の定義について, N 個のデータ D_n ($n = 1, 2, \dots, N$) を入力としたとき, データに含まれる $TP + FN$ で表される正解データの数 C とする. このとき $I(r)$ をデータ D_r ($r = 1, 2, \dots, N$) が正解データである場合に 1 を, そうでない場合に 0 となる関数と定義し, 入力データが k 個のときの適合率 $P(k)$ は,

$$P(k) = \frac{1}{k} \sum_{i=1}^k I(i) \quad (2.3)$$

と表される. このとき, 平均適合率は,

$$AP = \frac{1}{C} \sum_{i=1}^N I(i) P(i) \quad (2.2)$$

により得られるものである.

更に平均適合率の平均を求めたものを mAP (mean Average Precision) といい, M 個の平均適合率 AP_m ($m = 1, 2, \dots, M$) としたとき, 以下の式

$$mAP = \frac{1}{M} \sum_{j=1}^M AP_j \quad (2.3)$$

で定義する.

物体検出においては検出するクラス毎の平均適合率 AP を算出, 全クラスについての平均をとり mAP とする.

物体検出における予測の重みの指標を確信度 **confidence** で表す. 確信度 **confidence** は 0 から 1 の値を取り, 1 に近いほど高い予測の重みを示す.

2.4 Single Shot MultiBox Detector (SSD) [7]

SSD とは Wei Liu らによって提案された深層学習を用いた一般物体検出アルゴリズムである. SSD は PASCAL VOC, COCO, ILSVRC データセットについて, その他の手法と同程度精度を得ながら非常に高速である. VOC2007 テストによる一般物体検出の精度の比較について, mAP が 74.3% で 59FPS という結果を発揮し, 当時非常に高い精度を示していた Faster R-CNN の mAP が 73.2% で 7FPS, YOLO の mAP が 63.4% で 45FPS という結

果を大きく上回った.

2.5 むすび

本章では, 本研究で用いるデータセットや指標, 物体検出手法に関して述べた.

第 3 章では, 本章で述べたものを用いた予備実験である深層学習福笑いについて述べる.

第3章

予備実験

3.1 まえがき

本章では，予備実験として行った深層学習福笑いについて述べる．まず概要を述べた後，実験についてその方法，環境，結果，考察について述べる．

3.2 深層学習福笑い

日本の伝統的な遊戯である「福笑い」を参考に，マンガキャラクターの輪郭にマンガキャラクターの顔パーツを配置，深層学習によるマンガキャラクター顔検出をリアルタイムに行う．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，その物体検出の様子は従来の特徴量を用いた手法に比べ不明瞭である．本実験を通し，深層学習による物体検出のマンガキャラクターの顔パーツの配置と深層学習によるマンガキャラクターの顔検出の関係性を直感的に確認する．

3.3 実験方法

本実験では，マンガキャラクターの顔検出のため深層学習を用いた物体検出アルゴリズムである SSD を用いる．SSD の学習については Manga109 のデータを使用し，キャラクターの顔 (character) ，吹き出し (balloon) ，テキスト (text) ，コマ (panel) の4クラスの学習を行ったものを用いる．学習画像の枚数は6作品から600枚の画像を使用，150枚を用いたテストにおける mAP は 0.799 であった．

「福笑い」を参考に，マンガキャラクターの顔の輪郭を描いた紙と，顔のパーツを書き切り取った紙を用意する．顔のパーツとしては目，口，鼻，髪を使用する．マンガキャラクターの顔の輪郭を描いた紙の例を図 3-1，顔のパーツを書き切り取った紙の例を図 3-2 に示す．

マンガキャラクターの顔の輪郭が描かれた紙に顔のパーツを配置し、その様子を WEB カメラで撮影、撮影した画像についてリアルタイムの画像処理を行う。この時顔のパーツの配置を変えた場合のマンガキャラクターの顔検出とその精度を調査する。



図 3-1: 顔の輪郭の例

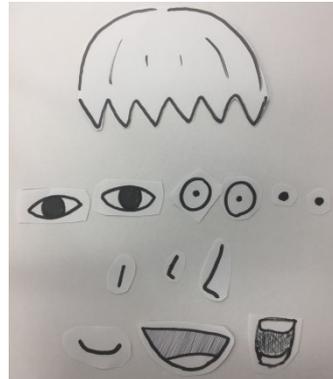


図 3-2: 顔のパーツの例

3.4 実験環境

実験を行った環境を以下の表 3-1 に示す。

表 3-1: 深層学習福笑いの実験環境

OS	Ubuntu 14.04
CPU	Intel(R) Core(TM) i7-6700 CPU @ 3.40GHz ×8
GPU	GeForce TITAN X
Deep learning framework	Caffe

3.5 実験結果

本実験における検出結果の例を保存した画像を図 3-3 から図 3-11 に示し、確信度 confidence を表 3-2 に示す。また、目を横に移動した場合の検出結果の画像を図 3-12、確信度 confidence を表 3-3 に示す。黄色の枠で囲まれた領域がマンガキャラクターの顔領域として検出された領域である。その他の色はマンガキャラクターの顔以外のクラスの検出領域を示すが本実験では考慮しない。リアルタイム性については 39 FPS 前後であった。

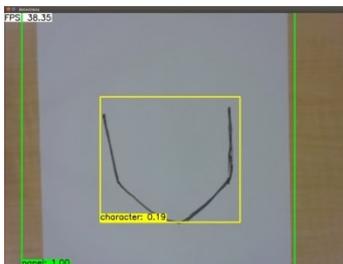


図 3-3: character A

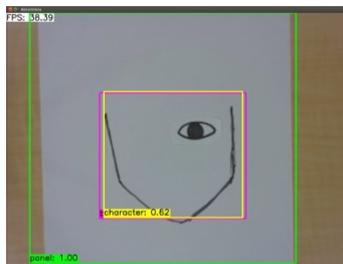


図 3-4: character B

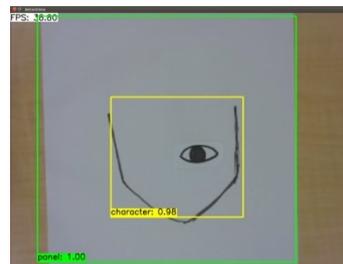


図 3-5: character C

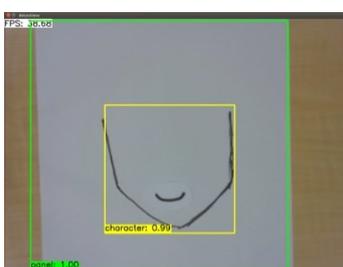


図 3-6: character D

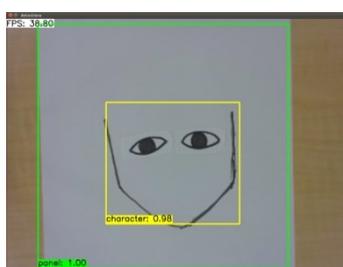


図 3-7: character E

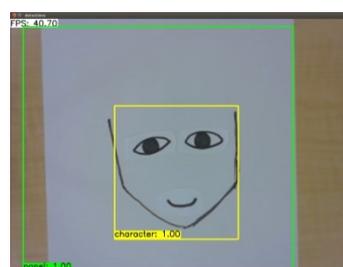


図 3-8: character F

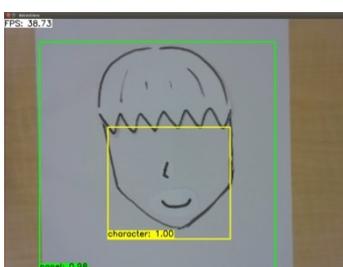


図 3-9: character G

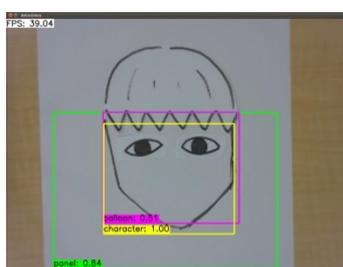


図 3-10: character H

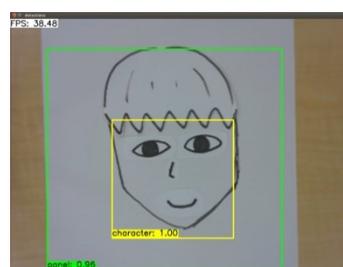


図 3-11: character I

表 3-2: 検出結果例の確信度 confidence

Character	A	B	C	D	E	F	G	H	I
Confidence	0.19	0.62	0.98	0.99	0.98	1.00	1.00	1.00	1.00

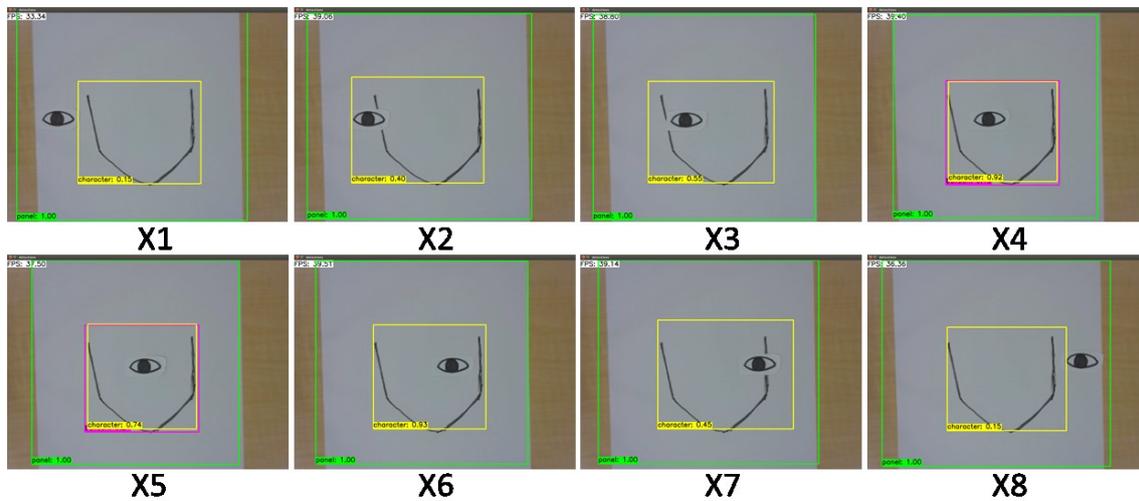


図 3-12: 目を横に移動した場合の検出結果

表 3-3: 目を横に移動した場合の確信度 confidence

Character	X1	X2	X3	X4	X5	X6	X7	X8
Confidence	0.15	0.40	0.55	0.92	0.74	0.93	0.45	0.15

3.6 考察

本実験の結果について、まず検出結果の例として示した **character A-I** について考察を行う。マンガキャラクターの顔のパーツと **confidence** の値の関係について、顔のパーツが置かれ実際のマンガキャラクターの顔に近くなるほど **confidence** の値が大きくなると考えられる。

character A についてはマンガキャラクターの輪郭のみが描かれている。この場合でも顔としては検出されるが、**confidence** の値は 0.19 と低い値を示している。

character B-I はマンガキャラクターの輪郭に顔のパーツを配置していった場合の検出結果は予想通り、輪郭のみの場合の **confidence** の値よりも高くなっている。

character B, C は同じ目が置かれているが位置の違いにより、**confidence** の値が **character B** が 0.62 に対し **character C** は 0.98 と 0.36 の差がある。このことから、顔のパーツの有無だけでなく位置も検出において重要であることがわかる。

character D では口のみが置かれているが、**confidence** の値は 0.99 と非常に高い。

character E-I では二つ以上の顔のパーツを、正面向きのマンガキャラクターの顔として不自然でないと考えられる位置に配置している。その結果として、**confidence** の値はどれも非常に高い値となっている。特に **character F-I** では **confidence** の値が 1.00 を示しており、どの **character** も三つ以上のパーツが配置されていることから、自然な位置により多くのパーツが配置されると正しく検出されやすいことがわかる。また、代表的な顔のパーツとして目、鼻、口、髪を配置してきたが、検出にすべてのパーツが揃っている必要はないことがわかる。

次に、目を横に移動した場合の検出結果について考察を行う。この検出では、顔の左側の輪郭の外側から右側の外側まで、おおよそ平行に移動させ 8 枚の画像の検出結果を示した。本実験における深層学習アルゴリズムの学習に使用したデータでは、マンガキャラクターの顔の向きについて、正面向きのものと横向きのを区別していない。また、横向きの顔には目が一つしか描かれないものが存在し、これは目が一つの場合でもよく検出が行われている原因となっていると考えられる。

X1 から X8 について **confidence** の値の変化を見ると、目が輪郭より外側にある場合である X1, X8 で 0.15 となっており、輪郭のみ描かれた場合よりも低くなっている。これはマンガキャラクターの顔の検出範囲外の要素も考慮していることを示すと同時に、顔パーツが配置次第で **confidence** の値を下げることもあることも示している。

次に輪郭と重なっている部分 X2, X3 と X7 を見ると **confidence** の値はそれぞれ 0.4, 0.55 と 0.45 とやや低い値となっている。これは目が輪郭と重なって描かれることが少ないためであると考えられる。

最後に X4, X5, X6 を見ると、X4, X6 は、輪郭のやや内側に目が配置され、横向きの場合に描かれる場合、正面向きの場合それぞれについて、目の描かれやすい位置に存在している。その結果 0.92, 0.93 といった高い **confidence** の値が得られていると考えられる。

X5 が隣り合う X4, X6 に比べ confidence の値が低い 0.74 を示している要因としては、輪郭から離れた顔の中心あたりではあまり目が描かれないことが考えられる。

以上の結果から、顔のパーツの有無だけでなく位置も検出において重要であること、配置次第で検出を妨げることもあること、自然な位置により多くのパーツが配置されるとよく検出されることがわかる。

3.7 むすび

本章では、予備実験として行った深層学習福笑いの実験方法、実験結果とその考察について述べた。

第 4 章では、提案手法である深層学習を用いたマンガキャラクタの検出における顔変形の影響評価について述べる。

第4章

提案手法

4.1 まえがき

本章では，提案手法である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について述べる．まず提案手法について述べ，次に実験方法について述べる．その後，実験について結果と考察を述べる．

4.2 提案手法

第3章で述べた深層学習福笑いでは，WEBカメラにより撮影すること，顔のパーツを手動で操作すること，実際のマンガの顔パーツではなく実験者により描かれた顔パーツを用いることなど，直感的な評価は可能であるが定量的な評価は困難であった．本実験では，実際のマンガ画像データを用い，深層学習によるマンガキャラクターの顔検出について，キャラクターの顔を変形した際の検出状況の変化を，より定量的に評価することを目的として実験を行う．

4.3 実験方法

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の実験方法について述べる．マンガキャラクターの顔検出は第3章で用いたSSDを用いる．

マンガ画像データから本実験の正例となるものを選ぶ．これは，Manga109の画像データに対しSSDによるマンガキャラクターの顔検出を行い，confidenceの値が1に非常に近いキャラクターの顔を含むものの中から，そのキャラクターの顔が正面を向いているものを5作品から142枚選び用意した．正例の画像の例を図4-1に示す．

これらの画像についてマンガキャラクターの顔の変形を行う．実験1としてマンガキャラ

クタの顔を中心とした回転を行った画像を作成した。回転は 360 度行い、3 度回転させる毎に画像を作成、マンガキャラクターの正例の画像 1 枚につき 120 枚の画像を作成した。作成した画像の例を図 4-2 に示す。

実験 2 としてマンガキャラクターの左目の領域の平行移動を行った画像を作成した。平行移動は、横軸及び縦軸方向に上下左右それぞれについて行った。150 ピクセルずつ動かし、2 ピクセル毎に画像を作成、マンガキャラクターの正例の画像 1 枚につき 150 枚の画像を作成した。平行移動の方向について、正例の画像の左目の位置を原点とし、横軸方向に関しては右側を正、左側を負とし、縦軸方向に関しては下側を正、上側を負とする。また、平行移動を行う前の左目の領域は白く塗りつぶした。作成した画像の例を図 4-3 に示す。

作成した画像について、SSD によるマンガキャラクターの顔検出を行い confidence の値を調査する。マンガキャラクターの顔の変形により、検出される領域の変形が生じる場合や、検出領域が重なりを持ち複数検出される場合がある。検出される領域の変形が生じる場合については、正例の顔領域を基準に検出領域の選択を行った。

検出領域が重なりを持ち複数検出される場合については、Intersection over Union (*IoU*) を用いて検出領域の選択を行った。*IoU* はバウンディングボックスの重なり具合を示す指標であり、ふたつのバウンディングボックスを結合した領域を U 、ふたつのバウンディングボックスの重なり合った領域を O とすると *IoU* は、

$$IoU = \frac{O}{U} \quad (4.1)$$

により得られる。



図 4-1: 正例の画像の例 (ありさ²©八神 健)

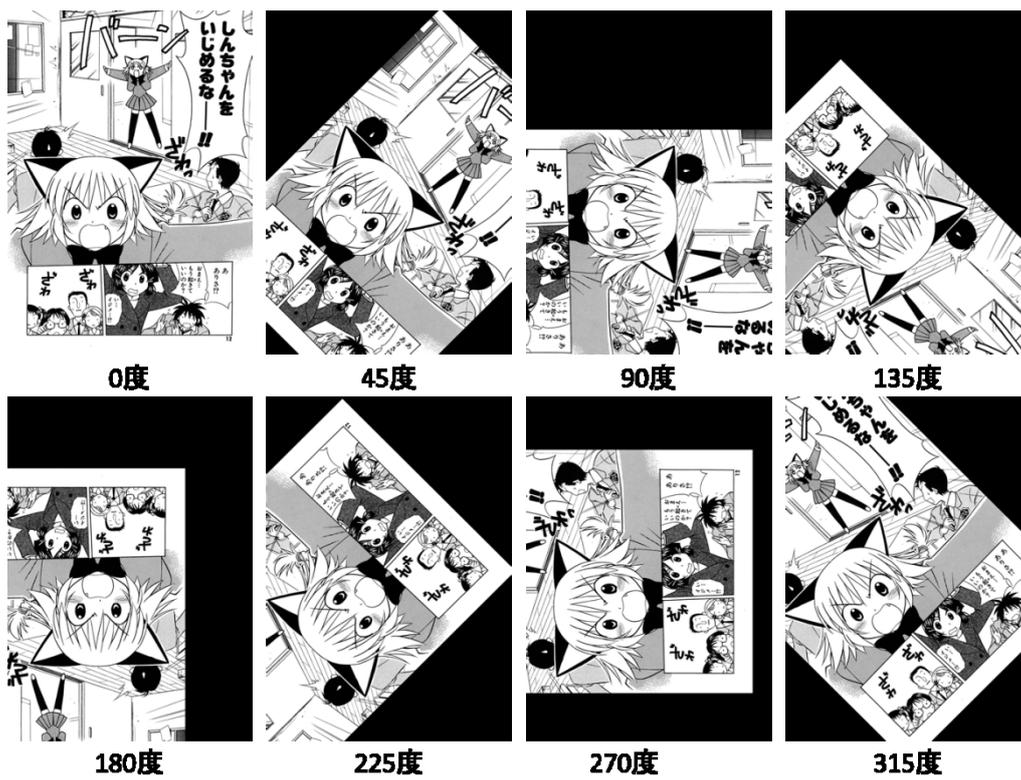


図 4-2: 回転した画像の例 (ありさ²©八神 健)

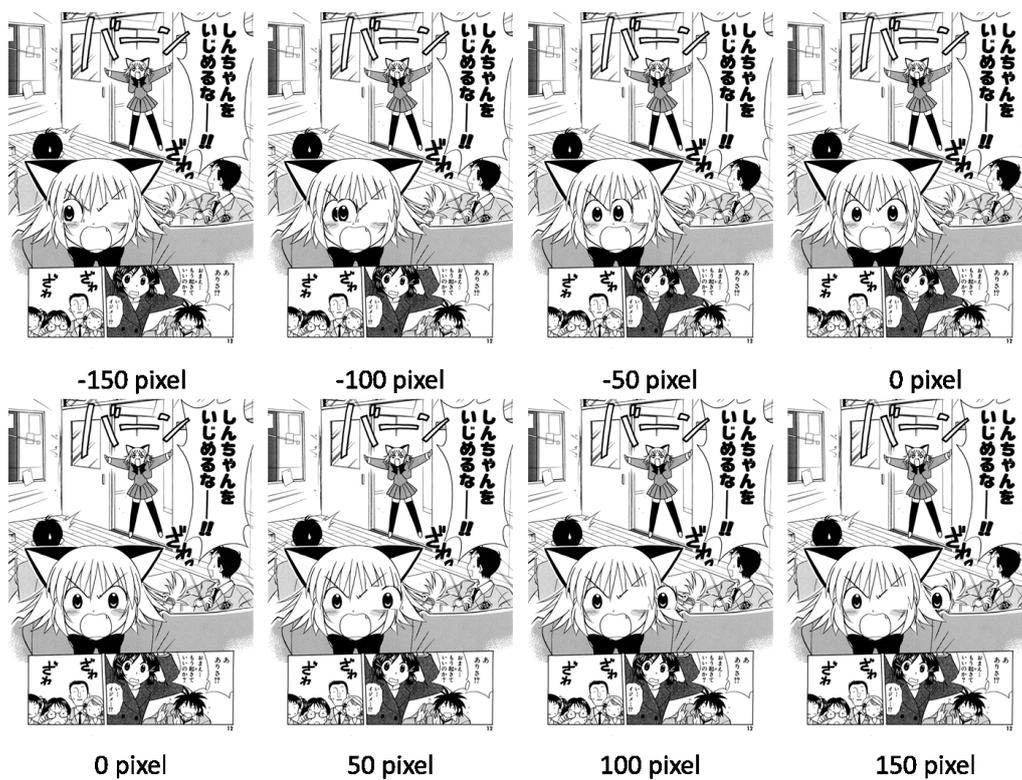


図 4-3: 横軸方向に平行移動した画像の例 (ありさ²©八神 健)

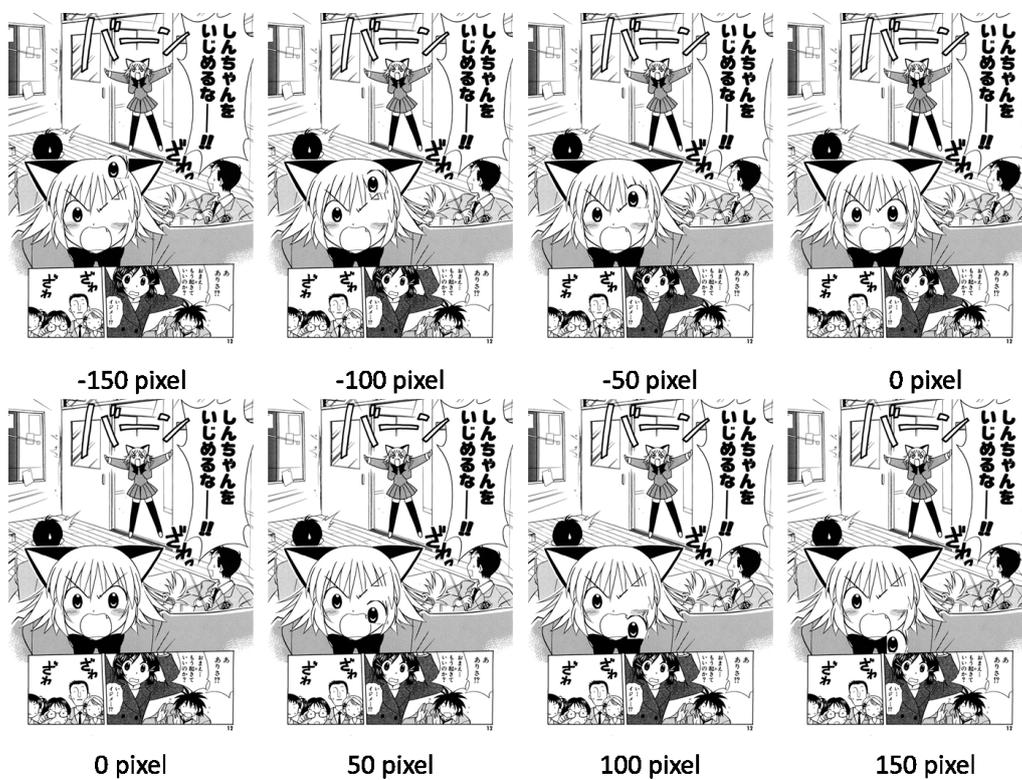


図 4-4: 縦軸方向に平行移動した画像の例 (ありさ²©八神 健)

4.4 実験1の結果と考察

マンガキャラクターの顔を中心とした回転を行った画像に対する SSD によるマンガキャラクターの検出結果について、回転角度と confidence の平均値の関係のグラフを図 4-5 に示す。

本実験で用いた SSD は学習において、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、正例の画像の顔の角度からの角度のずれが大きくなっていくに従って confidence の値が下がっていくことが予想される。実際にグラフを見ると、グラフでは時計回りに回転していった様子を示しているが、そのグラフは 180 度を中心に左右対称に非常に近い形となっている。このことから予想の通り、正例の画像の顔の角度からの角度のずれが大きくなっていくに従って、confidence の値が下がっていることが確認できる。

また、confidence の値は 140 度付近と 220 度付近で最も低い値となり、その値はおよそ 0.2 である。しかし、180 度付近では confidence の値が 0.4 付近まで上がっていることがわかる。その要因としては 180 度程度の回転をすると、目が上下反転してはいるが正例の画像に近い配置となること、目に対して自然な位置に口はないものの、マンガのキャラクターにおいては口や鼻が描かれないキャラクターもしばしばいることなどが考えられる。

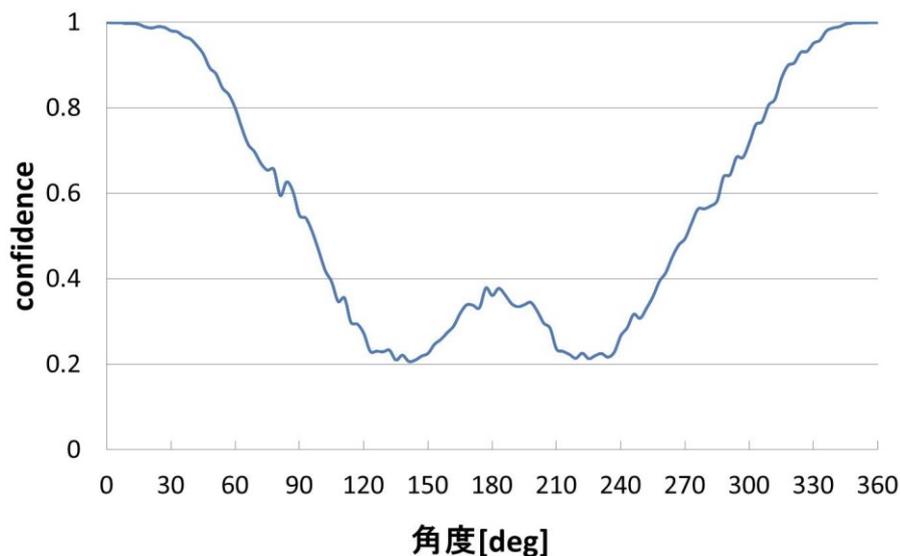


図 4-5: 回転した画像の検出結果

4.5 実験2の結果と考察

マンガキャラクターの左目の領域の平行移動を行った画像に対する SSD によるマンガキャ

ラクタの検出結果について、移動距離と **confidence** の平均値の関係のグラフを作成した。横軸方向の平行移動に関する検出結果のグラフを図 4-6 に、縦軸方向の平行移動に関する検出結果のグラフを図 4-7 に示す。

左目の領域の平行移動を行った画像に関して、移動距離の基準と考えられる正例の画像の顔領域の大きさの平均は幅 163 ピクセル、高さ 164 ピクセルであった。また、左目を移動し配置することで隠れる部分が生じる。移動距離が小さい 10 ピクセル以下の領域では、**confidence** の値の大きな変化は見られない。

横軸方向の平行移動を行った画像について考察する。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 ピクセルから 100 ピクセル付近で **confidence** の値が最大 0.97 付近まで下がっている。この要因としては、移動により右目を隠すことが考えられる。一方、100 ピクセル以降で **confidence** の値が大きくなっている。この要因としては、移動により隠されていた右目が隠れなくなることが考えられる。

縦軸方向の平行移動を行った画像について考察する。片目を縦軸方向に平行移動すると、2つの目の位置が上下にずれた関係となる。これはマンガキャラクターの顔に見られづらい状況であると考えられ、**confidence** の値を下げる要因になっていると考えられる。横軸の平行移動に比べ、縦軸の平行移動の方が低い値を示すのもこのためと考えられる。正の方向への移動における **confidence** の値は、移動距離 20 ピクセル付近から下がり始め、輪郭を隠す 100 ピクセル付近まで下がり最大 0.94 付近まで下がっている。その後 100 ピクセル以降で **confidence** の値が大きくなっている。この要因としては、移動により検出領域外に配置されるため、マンガキャラクターの検出領域内が隠されなくなることが考えられる。負の方向への移動における **confidence** の値は、髪やマンガキャラクターの検出領域を隠す、移動距離 10 ピクセルから 60 ピクセル付近まで **confidence** の値が最大 0.96 付近まで下がっている。移動距離 60 ピクセル付近から **confidence** の値が上がるのは、移動により検出領域外に配置されるため、マンガキャラクターの検出領域内が隠されなくなることが考えられる。

以上の結果から、単純に回転や平行移動が行われる場合には、正例の画像の顔の状態との違いが大きくなるにつれ、**confidence** の値が下がっていくこと、検出領域外の変化は **confidence** の値に大きく影響しないことがわかった。また、平行移動により隠れる部分が生じる場合についてはこの限りではなく、自然に描かれるマンガキャラクターの顔の状態と異なる状態になる場合、**confidence** に比較的大きな影響を与えることがわかった。

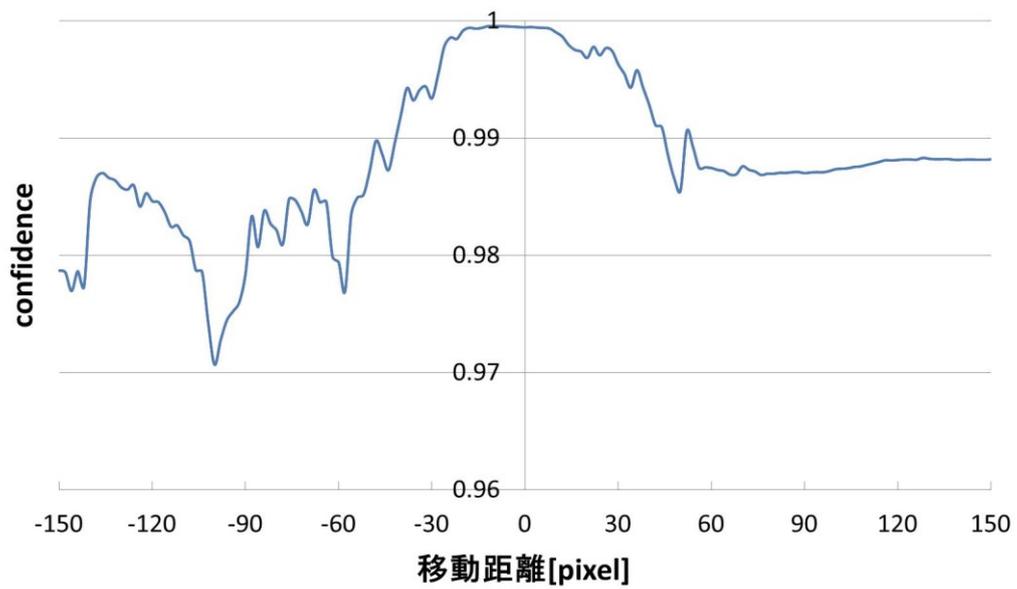


図 4-6: 横軸方向に平行移動した画像の検出結果

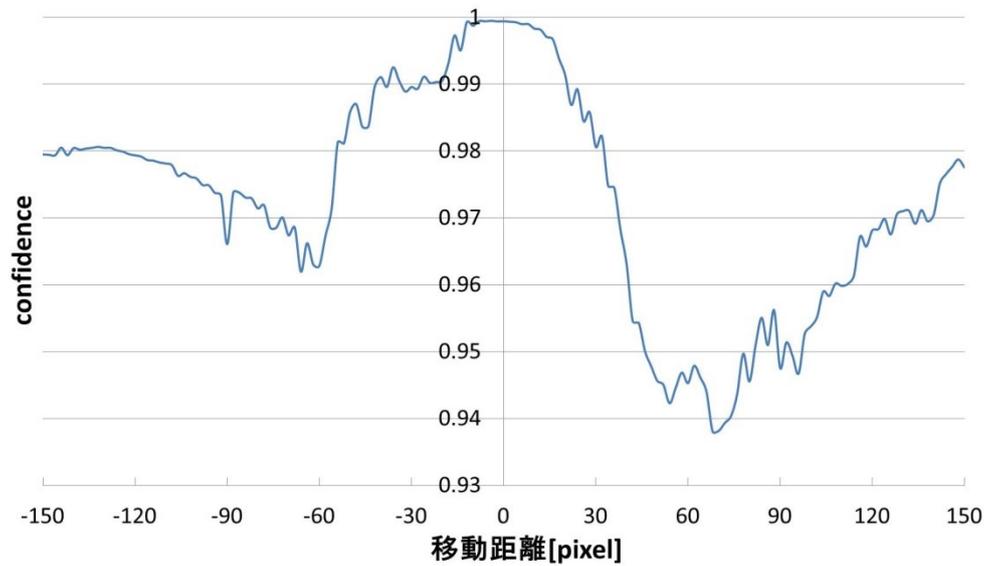


図 4-7: 縦軸方向に平行移動した画像の検出結果

4.6 むすび

本章では，提案手法である深層学習を用いたマンガキャラクタの検出における顔変形の影響評価について，実験方法，実験結果とその考察について述べた．

第 5 章では，本論文の結論について述べる．

第5章

結論

本論文では、深層学習によるマンガキャラクターの顔検出について、キャラクターの顔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した。まず、予備実験として行った深層学習福笑いでは、リアルタイムで直感的に深層学習による検出の様子を確認した。特に、顔のパーツの有無だけでなく位置も検出において重要であること、配置次第で検出を妨げることもあること、自然な位置により多くのパーツが配置されるとよく検出されることを確認した。

次に、マンガキャラクターの検出における顔変形の影響評価では、単純に回転や平行移動が行われる場合には、正例の画像の顔の状態との違いが大きくなるにつれ、**confidence** の値が下がっていくこと、検出領域外の変化は **confidence** の値に大きく影響しないことがわかった。また、平行移動により隠れる部分が生じる場合についてはこの限りではなく、自然に描かれるマンガキャラクターの顔の状態と異なる状態になる場合、**confidence** に比較的大きな影響を与えることがわかった。

今後の課題について、本論文ではマンガキャラクターの回転と目の平行移動に関する顔変形しか行っていない。顔領域の透過などのフィルタ処理、目以外のパーツに対する顔変形など、更なる実験により深層学習によるマンガキャラクターの検出状況の調査を行いたい。

また、本論文ではキャラクターの顔をひとつのクラスとし学習や検出を行い、正面向きと横向きの場合を区別していない。このことは本論文の実験結果に影響を与えていたと考えられる。これについて、正面向きと横向きの場合にそれぞれのクラスを分けた場合の検出結果の比較により、クラス分類に対する指標としての応用を検討したい。

謝辞

本研究は JSPS 科研費 17K0511 の助成を受けたものである。

本研究に際して充実した研究環境を整えて頂き、熱心なご指導を頂きました渡辺裕教授に心より感謝いたします。

日頃から研究に関する助言や相談に乗って下さいました渡辺研究室の皆様に感謝いたします。

最後に、私をここまで育ててくださり、生活を支えてくださっている家族に感謝いたします。本当にありがとうございました。

参考文献

- [1] 日本経済新聞, “電子書籍が変える読書 世界市場 4年で急成長,” 日本経済新聞, <https://vdata.nikkei.com/datadiscovery/21book/>, 2018年1月19日参照.
- [2] 経済産業省, “クールジャパン政策について,” 経済産業省商務情報政策局生活文化創造産業課, http://www.meti.go.jp/policy/mono_info_service/mono/creative/file/170705CooljapanseisakuJuly.pdf, 2018年1月19日参照.
- [3] H. Yanagisawa, D. Ishii, and H. Watanabe: “Face Detection for Comic Images with Deformable Part Model,” The 4th International Workshop on Image Electronics and Visual Computing (IEVC2014), 4A-1, Oct. 2014.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.

図一覧

図 3-1: 顔の輪郭の例	7
図 3-2: 顔のパーツの例.....	7
図 3-3: character A	8
図 3-4: character B	8
図 3-5: character C	8
図 3-6: character D	8
図 3-7: character E	8
図 3-8: character F	8
図 3-9: character G	8
図 3-10: character H.....	8
図 3-11: character I.....	8
図 3-12: 目を横に移動した場合の検出結果	9
図 4-1: 正例の画像の例（ありさ ² ©八神 健）	14
図 4-2: 回転した画像の例（ありさ ² ©八神 健）	14
図 4-3: 横軸方向に平行移動した画像の例（ありさ ² ©八神 健）	15
図 4-4: 縦軸方向に平行移動した画像の例（ありさ ² ©八神 健）	15
図 4-5: 回転した画像の検出結果.....	16
図 4-6: 横軸方向に平行移動した画像の検出結果	18
図 4-7: 縦軸方向に平行移動した画像の検出結果	18

表一覧

表 3-1: 深層学習福笑いの実験環境	7
表 3-2: 検出結果例の確信度 confidence.....	8
表 3-3: 目を横に移動した場合の確信度 confidence.....	9

研究業績

- [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] 山下, 柳澤, 渡辺: “深層学習福笑い,” 2017年画像符号化シンポジウム・2017年映像メディア処理シンポジウム (PCSJ/IMPS2017), P-5-2, Nov. 20-17. 2017.
- [3] 柳澤, 山下, 渡辺: “マンガキャラクター顔画像クラスタリングの改良のおける一検討,” 映像情報メディア学会冬季大会, 22B-8, Dec. 2017
- [4] H. Yanagisawa, T. Yamashita and H. Watanabe: “A Study on Object Detection Method from Manga Images Using CNN,” International Workshop on Advanced Image Technology (IWAIT2018), No.16, pp.1-4, Jan. 2018
- [5] 山下, 柳澤, 渡辺: “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,” 電子情報通信学会総合大会, March. 13-15. 2018(発表予定) .

