

# 投稿イラストのトレース回避のための キャラクターの姿勢類似度判定

進藤 光太<sup>†</sup> 中島 克人<sup>‡</sup>

東京電機大学未来科学部情報メディア学科<sup>†‡</sup>

## 1 はじめに

近年、イラストをネットや SNS に投稿する人達が増加している。しかし投稿の増加に伴い、他人のイラスト作品を作者の許可なく模倣する“トレース”を行って公開してしまう事やトレースを意図していないにも関わらず、疑いを掛けられて、作者が被害を受ける事が増加している。

そこで、トレースの疑いをかけられる被害の防止や故意にトレースを行う人への抑止に繋がる事を目的として、我々はトレースかどうかの自動判別を目指している。

イラスト画像では(人間の)キャラクターの姿勢がトレース対象となることが多いため、今回我々は、深層学習手法に基づく姿勢推定技術をイラストのキャラクターに適用し、姿勢類似度の判定を行う手法を提案し、評価を行った。

## 2 関連研究

姿勢推定技術は、画像から人物の関節点を学習しておき、推定した関節点同士を結ぶことで人物の姿勢を検出する技術である。代表的なものとして OpenPose[1]がある。実画像の人物だけでなく、漫画やアニメのキャラクターに対する姿勢推定技術[2]も研究されている。しかし、イラストのキャラクターに対する姿勢推定の研究は見当たらない。

## 3 提案手法

今回の対象としているイラストはキャラクターが描かれているものであり、事前実験では、背景が無地であれば人用に学習された OpenPose によりキャラクターの関節情報も高い精度で出力できることが判明した。しかしながら、投稿される多くのイラスト画像では複雑な背景が描かれる事が多く、正確な姿勢推定を阻害する。そこで、深層学習手法に基づく領域分割技術を用いて、イラスト画像の背景の影響を低減するための画像処理を行った後に姿勢推定を行う事とした(図 1)。

背景領域の分割には、代表的領域分割手法の 1 つである U-Net[3]を用いる。学習データにはキャラクターの描かれたイラスト画像の背景に手で

アノテーションを付与したものを用いる。

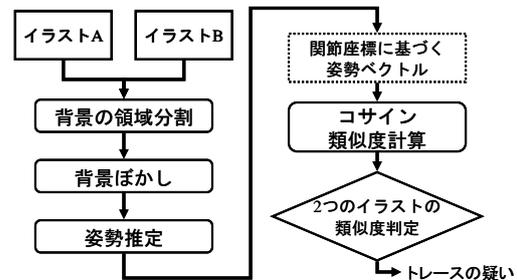


図 1 提案手法の処理の流れ

姿勢推定時には、U-Net で分割された背景領域に対して姿勢推定の障害にならないような処理を施した上で OpenPose に入力する。比較する 2 枚のイラストに対してこれを行い、それぞれにおいて、2 つの接続する関節座標からなる 10 本のベクトル(姿勢ベクトル)同士のコサイン類似度の合計で両イラスト画像がトレースされたものかどうかを判断する。

キャラクター部分の姿勢推定前に背景部分を完全に除去するのが理想的であるが、領域分割が不完全であるとキャラクター領域の一部も背景として除去され、姿勢推定の精度が下がってしまう。そこで、キャラクター領域には膨張処理をかけると同時に背景領域は除去ではなくぼかし処理にとどめる。

類似度判定の際には、同じイラスト画像でも膨張の大きさやぼかし割合を変更して姿勢推定が最も正しく検出したものを使用する。

## 4 学習・評価

背景部分を処理した 2 枚のイラスト画像に対して姿勢推定を行い、それぞれの類似度を算出し、2 枚の姿勢類似度を求める。そして、アンケートでの主観評価も行い、姿勢類似度と比較・評価する。

### 4.1 U-Net の学習

イラスト投稿サイトの Danbooru から“rating=safe”, “solo”, “full\_body”のタグが付与されているイラスト画像を 353 枚収集し、アノテーション付けを行ってデータセットを作成した。これらは 1 枚の画像にキャラクターが単体かつ全身が描かれているイラスト画像である。これらの画像に対して、左右反転のデータ増強を行い、合計

Similarity judgment of character's posture to avoid tracing illustrations

<sup>†</sup> Kota Shindo · Tokyo Denki University

<sup>‡</sup> Katsuto Nakajima · Tokyo Denki University

706 枚を学習用データとした。画像サイズにはばらつきがあるため、学習前にリサイズを行い統一した。学習パラメータを表 1 に示す。

表 1 学習時のハイパーパラメータ

入力サイズ	800 px×1280 px
バッチサイズ	2
エポック数	200
学習率	0.01
訓練画像	636 枚
検証画像	70 枚

#### 4.2 姿勢の類似度判定

姿勢推定に用いる OpenPose により、25 点の関節座標情報が得られる。類似度の算出には西澤らの手法[4]を参考にした。まず、図 2 に示すように取得した座標から隣り合う関節座標同士で、10 本の姿勢ベクトルを計算する。2 枚の画像それぞれ同じ姿勢ベクトル間のコサイン類似度を求め、それらを合計して全体の類似度とする。なお、一方の画像全体が回転すると類似度は下がるという点に注意する必要がある。

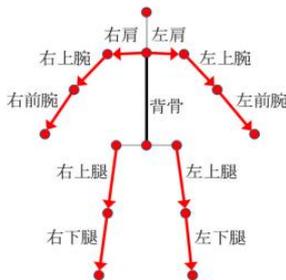


図 2 姿勢ベクトル ([4]より引用)

#### 4.3 主観評価

34 名の被験者に対してアンケート調査を行った。被験者には 2 枚 1 組のイラスト画像 10 組（イラストペア A~J）を提示し、1 枚のキャラクター姿勢が他方のそれをトレースしたものかどうかを判定して貰う。

なお、被験者のイラストとの関わり具合によって評価が変化するかを調べるため、被験者には事前に「普段からイラストの閲覧や投稿を行っているかどうか」を回答して貰った。

#### 4.4 評価結果

イラストペア A~J の 10 組のコサイン類似度と主観評価結果を表 2 に示す。表 2 の左から、イラストペアに対するコサイン類似度、全被験者 34 名中トレースの可能性があると判断した人数の割合、普段からイラストと関わりがあると答えた被験者 21 名中トレースの可能性があると判断した人数の割合を表している。

表 2 コサイン類似度と主観評価 (%)

イラストペア	コサイン類似度	全被験者 34 名	関わりある被験者 21 名
A	83.9	14.7	9.5
B	95.1	35.3	33.3
C	76.4	5.9	4.7
D	<b>90.7</b>	<b>94.1</b>	<b>100</b>
E	79.1	2.9	0
F	74.9	23.5	23.8
G	<b>96.3</b>	<b>85.3</b>	<b>85.7</b>
H	87.0	32.4	23.8
I	89.2	2.9	4.7
J	90.4	17.6	4.7

イラストペアの D と G の一方は過去にトレースとして判断されたイラスト画像であるが、D、G 共にコサイン類似度が高く、アンケート結果も同様に高かったことから、客観的指標であるコサイン類似度が主観的評価に近い指標になり得ると言える。しかし、B、I、J ではコサイン類似度は高いが主観的評価は低く、コサイン類似度だけではトレース可能性の判定には不十分である事も言える。例えば I では共にキャラクターが腕を前に組んで立っている姿勢として類似しているためコサイン類似度は高くなるが、この姿勢はイラストにおいて描かれやすい姿勢であるため、主観評価ではトレースとして判断されなかった。従って、描かれやすい姿勢かどうかを大量のイラスト画像で学習し、コサイン類似度に対する重み調整を行うなどの工夫を考える必要がある。

なお、主観評価において、全被験者 34 名分と普段からイラストと関わりがある被験者 21 名分だけの結果はそれほど大きな差は見られなかった。

#### 5 今後の課題

キャラクターの服装や持ち物、背景の複雑さにより、背景領域の分割が上手く行かないイラストや、関節座標取得がそもそも困難な姿勢のイラストも多くある。そのため、そのような画像でも姿勢推定精度を高めるために、イラスト画像で OpenPose の追加学習を行う事が課題の 1 つである。また、上記の描かれやすいを姿勢に対する客観指標の重み変更、および、今回は対象外としたキャラクターの体の一部だけのイラスト画像のトレース判定も今後の課題である。

#### 参考文献

- [1] Z. Cao, et al., "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," In CVPR2017
- [2] P. Khungum, et al., "Pose Estimation of Anime/Manga Characters: A Case for Synthetic Data," In MANPU2016
- [3] O. Ronneber, et al., "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science, 2015
- [4] 西澤, 他, "姿勢推定を援用した実人型モデルの描画学習支援システム," 芸術学会論文誌, 2019