

# 深層学習を用いた 人物イラストの顔の自動補完

山本大貴<sup>1</sup> 植村 匠<sup>1</sup> 尾島修一<sup>1</sup>

**概要:** 人物イラスト描画のコンピュータ支援技術の開発を最終目標として、まず、顔のパーツの自動生成を試みた。顔を描く技術の中でも、顔のパーツ（目・鼻・口等）を描く技術は習得に時間がかかることが知られている。本研究では、自然画像の欠損箇所を自動生成する飯塚らの手法と、ラフ画から線画を自動生成するシモセラらの手法をもとに、顔パーツ以外の線画イラストから顔のパーツを自動生成する技術を提案する。結果、ボケてはいるが、まつ毛や目など大きいパーツは再現できた。

**キーワード:** イラスト描画, 画像生成, ニューラルネット, 画像補完

## Automatic Face Completion of Human Illustration Using Deep Learning

DAIKI YAMAMOTO<sup>†1</sup> TAKUMI UEMURA<sup>†1</sup>  
OJIMA SHUICHI<sup>†1</sup>

**Abstract:** Our goal is to develop the computer aided technology for drawing human illustrations. We first attempt to automatically generate faces. Among the face drawing techniques, the technique of drawing face parts( eyes, nose, mouth, etc.) takes time to learn. In this study, we propose a method for automatically complementing faces in human illustrations based on the method of Iizuka et al., which automatically generates missing parts in natural images, and the method of Simocera et al., which automatically generates line drawings from rough drawings. As a result, large parts such as eyelashes and eyes can be reproduced, but blurred images are outputted.

**Keywords:** Illustration, Image Synthesis, Neural Networks, Image Inpainting

### 1. はじめに

絵を描く技術は習得が難しく、習得できても多くの手間と時間がかかる。近年では、紙と鉛筆、ペンだけでなく、ペンタブレットや、イラスト作成ソフト、イラスト投稿ウェブサイトの普及が進み、絵を描く人々（以下、イラストレータ）の数も増えてきている。また、TwitterをはじめとするSNSでは、初心者でも気軽にイラストを投稿できることから、コミュニティが拡大を続けている。機械学習の分野においても、線画イラストから自動的に着色を行ったり、ラフ画を線画に変換したりなどの研究がなされている。

これらの新しいデバイス、サービスの普及に伴い、デジタルイラストを描き始める人が増えてきている。また、イラスト初心者に対するウェブ講座なども充実してきていることから、今後ますますイラストレーターの人口が増えることが予想される。

現在、プロ、アマチュア、趣味レベル問わず、多くのイラストレータがいるが、主にプロ未満のイラストレータが絵を描く技術を学ぶ際に最初にぶつかる関門が、顔を描く技術である。顔のパーツや形は角度によって大きく形

状が変化するためである。そのほかにも多くの関門が存在するが、その関門を通過するために多くの時間を費やす必要がある。また、アニメーション業界でも求められる技術や作業量の割に報酬が少ないという理由から、人が集まらず人手不足が問題になってきている。

これらの問題を解決するために、さまざまな作画作業の自動化を目指して、機械学習を用いたコンピュータ支援技術の開発が進められている。本研究では、顔のパーツを作成するための支援技術の検討を行う。

### 2. 関連研究

漫画やイラスト線画を自動的に着色する手法や、ラフ画を線画に変換する手法など数多くの手法が提案されている。さらに、それに伴い、nico-opendata[1]などのイラスト・漫画データセットが配信されている。また、近年では、敵対生成ネットワークを用いて、より精度を高くする手法が提案されており、高い成果を収めている[2]。以下、そのうちのいくつかを紹介する。

#### (1) 白黒漫画に自動着色する手法

白黒漫画と着色の基準となる参照画像を入力してカラ

<sup>1</sup> 崇城大学  
Sojo University

一漫画を生成する。入力する漫画は漫画全ページを対象とし、加えてカラーの参照画像を入力することで、参照画像に含まれる色を用いてすべてのページで同じキャラクターに対して一貫した着色を行う[3]。また、ユーザーがさらにカラードットを加えることで、色付け結果をインタラクティブに変更できる。しかし、色がはみ出てしまう場合があり、その都度ユーザーの調整が必要である。

### (2) ラフ画像を線画に変換する手法

畳み込みニューラルネットワーク(以降 CNN)[4]を用いてラフ画を線画に自動変換する[5]。3種類の畳み込み層から構成されるネットワークで、複雑なラフ画と線画の対応を学習することで、ラスター形式の様々なラフスケッチを良好に線画化することができる。任意のサイズやアスペクト比を持つ画像を入力として扱うことができ、出力する線画も入力画像と同じサイズになる。また、線画からラフ画に変換することも可能である。

### (3) 敵対生成ネットワークの識別器を用いた手法

飯塚らは CNN を用いて、写真におけるシーンの大域的かつ局所的な整合性を考慮した画像補完を行う[6]。また、飯塚らは畳み込みの手法に Dilated Convolution[7]を用いた。これは畳み込みを行う際、より大域的な情報を用いることで、より広範囲の情報を考慮した畳み込みを行うための手法である。補完ネットワークと大域識別ネットワークと局所識別ネットワークの3つのネットワークがある。2つの識別ネットワークをだますように補完ネットワークを学習させることで、全体の整合性が取れ、かつ局所的にも自然な補完画像を出力することができる。

## 3. 提案手法

本研究では、顔のパーツを作画するための支援技術として、顔のパーツの自動生成を行う。すなわち、顔のみを切り抜いたイラストの顔ありとなしのセットで学習させた補完ネットワークに、未知の顔のないイラストを入力することで顔を自動生成する。そのために、まず、飯塚らの提案したネットワークを用い、従来手法[4]での写真の入力を人物イラストの線画の入力に変更し、顔のパーツ生成の精度を検証する。

具体的には、既存の人物イラスト線画を入力し、学習させた全層畳み込みニューラルネットワーク(Fully Convolutional Networks: FCN)[8]に顔を空白の矩形でマスクした未知の人物イラスト線画を入力して、顔を補完させた画像を出力させる。また、FCNを線画に対応したシモセラらのネットワークに変更した場合の実験も行った。さらに、既存の顔のある人物イラスト線画を入力して学習させた識別器に、FCNで補間した線画または、未知の顔のある人物イラスト線画を入力し、入力された画像が本物か偽物かを

2値で識別させ、その結果をFCNに反映させてFCNをさらに学習させることで精度向上を図る手法を提案する。提案手法の概要を図1に示す。

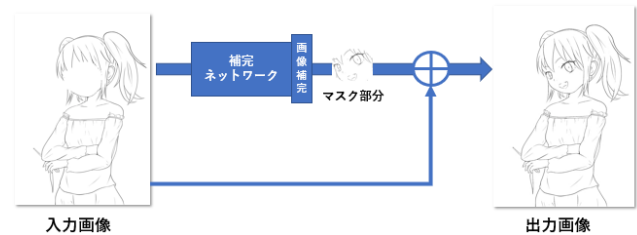


図1 提案手法の概要

Figure 1 Overview of the proposed method.

本研究で用いるニューラルネットワークの構成を表1及び図2に示す。表1の識別器の入力画像サイズは、前者が大域識別ネットワーク、後者が局所識別ネットワークである。また、入力層と出力層以外の層の後に、BatchNormalization[9]層および、DropOut[10]を加えた。

表1 ニューラルネットワークの構成  
 Table 1 Configuration of the neural network.

	生成器	識別器
ネットワーク構造	FCN	畳み込み層, 全結合層
入力画像サイズ	256×256 [pixel]	256×256 [pixel], 128×128 [pixel]
学習方法	教師あり学習	
更新方法	ミニバッチ学習	
誤差関数	平均二乗誤差	2値交差 エントロピー
最適化手法	Adam	
活性化関数	LeakyReLU, Tanh	

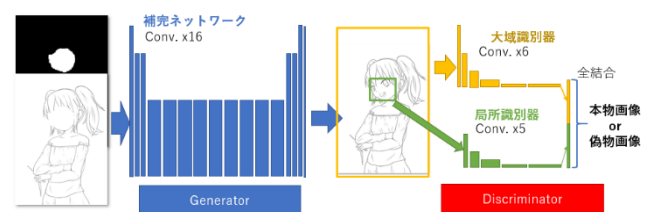


図2 提案手法のニューラルネットワークの構成

Figure 2 Configuration of the neural network.

## 4. 実験と結果

本研究では学習を3段階に分けて行う。1段階目は、生成器単独での学習を行う。入力データに顔のない人物イラスト線画を、教師データに顔のある人物イラスト線画を用いる。2段階目では、識別器単独での学習を行う。ここで、

顔のある人物イラスト線画と顔のない人物イラスト線画をそれぞれについて入力し学習させる。3段階目では生成器と識別器を交互に学習を行う。識別器については2段階目と同様に学習させる。生成器は、入力データは1段階目と同様に入力するが、出力を識別器に入力し、生成された画像か教師データの画像かを判断させる。この時、生成器は識別器を騙すように学習させる。

#### 4.1 実験用データ

本実験のデータセットは顔のない人物イラスト線画とマスク画像と顔のある人物イラスト線画のセットからなる。このデータセットの詳細を表2に示す。学習のためのデータセットの作成は、市販の画像編集ソフトウェア（以下編集ソフト）を用いて行った。なお、元となったイラスト線画はインターネットより無作為に収集した。

表2 データセットの詳細

Table 2 Dataset details.

Training Set	317 枚
Test Set	95 枚
画像サイズ	256×256 [pixel]
濃淡諧調数	8 [bit]

顔のない人物イラスト線画の作成方法は以下の通りである。まず元画像である顔のある人物イラスト線画を編集ソフトで開き、レイヤーを新規作成し、白色(R:255, G:255, B:255)で顔の部分をマスクする。

マスク画像の作成方法は、顔のない人物イラスト線画を作成した後、マスクレイヤーの下に背景が黒(R:0, G:0, B:0)のレイヤーを作成する。

なお、データセット内の画像はすべて読み込み時にアスペクト比を維持したまま 256×256 pixel にリサイズした。

#### 4.2 実験手順

本実験では、まず生成器を飯塚らの FCN で実験（以降実験 1）した後、精度向上のため、線画に特化したシモセラらの生成器に変更して実験した（以降実験 2）。なお、学習率は実験 1 では生成器、識別器ともに 0.001 に、実験 2 では 0.02 に設定した。また、バッチサイズは実験 1 で 8、実験 2 で 16 に設定して実験を行った。

#### 4.3 実験結果

各実験中 1000 epoch 毎にテストデータを入力し、実際に画像の補完を行った。

なお、識別器を含めた学習結果は、あらゆる入力に対して意味のないノイズ画像を出力するのみとなった。そのため、以下の実験結果はすべて生成器の出力のみとなる。

実験 1 での出力結果を図 3 に、実験 2 での出力結果を図 4 に示す。なお、画像は左から順に入力画像、出力画像、正解画像となっている。

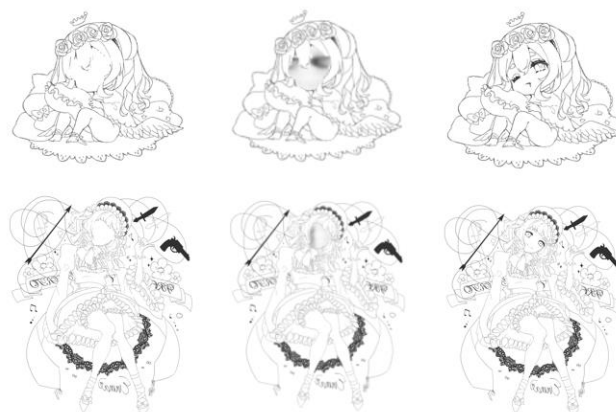


図3 実験1の出力結果

Figure 3 Result of experiment 1.



図4 実験2の出力結果

Figure 4 Result of experiment 1.

提案手法による実験の結果、実験 1 では目の部分に黒い霧のようなものが現れるのみで、顔とわかるような模様は現れなかった。実験 2 では目や鼻の位置に黒い模様が現れ、口や鼻はないものの、まつ毛や瞳の輪郭がある程度再現される結果も得られたが、全体的に精細感のない画像となった。また、いずれの実験においても識別器はノイズ画像を出力するのみであった。

#### 4.4 考察

この結果の原因として、識別器は自然画像などの RGB カラー画像の濃淡で表現しているデータに対しては識別しやすいが、イラスト線画は大部分が白背景の上に濃淡ではなく線で表現されたデータに対しては識別しにくいと推測される。

### 5. まとめと今後の課題

本研究では、顔のない人物イラスト線画に顔を自動補完する手法についての検討を行った。飯塚らとシモセラらの手法を応用し、CNN を用いて顔のパーツのない人物イラスト

ト線画に顔のパーツを自動補完する手法を提案した。

図3及び図4に示す実験の結果を比較すると、実験1で用いた飯塚らのFCNより、実験2で用いた線画に最適化されたFCNを用いたほうが、より顔らしい模様が出力されていることがわかる。すなわち、FCNに線画に最適化されたモデルを用いることで、はっきりとした顔は再現されなかったが顔らしい模様を出力することができたと考えられる。今後は、識別器を線画に対応したものに変更し、精細感を向上させることが課題となる。

## 参考文献

- [1] Ikuta, H. and Ogaki, K. and Odagiri, Y.. Blending Texture Features from Multiple Reference Images for Style Transfer. SIGGRAPH Asia 2016 Technical Briefs. 2016, no.15, p. 1-4.
- [2] Ian J. Goodfellow, et al.. Generative Adversarial Nets. Proc. NIPS 2014. 2014, p. 2672-2680.
- [3] Furusawa, C. et al. Comicolorization: Semi-Automatic Manga Colorization. SIGGRAPH Asia 2017 Technical Briefs. 2017, no.12, p. 1-4.
- [4] 岡谷 貴之. 機械学習プロフェッショナルシリーズ 深層学習. 講談社, 2015.
- [5] Edgar Simo-Serra, et al.. Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup. ACM Transactions on Graphics. 2016, vol.35, no.4, p. 121:1-121:11.
- [6] Iizuka, S. et al.. Globally and Locally Consistent Image Completion. ACM Transactions on Graphics. 2017, vol.36, no.4, p. 107:1—107:14.
- [7] F. Yu and V. Koltun. Multi-Scale Context Aggregation by Dilated Convolutions. International Conference on Learning Representations. 2016.
- [8] Jonathan Long, Evan Shelhamer, Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. CVPR2015. 2015.
- [9] Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Proceedings of the 32<sup>nd</sup> International Conference on Machine Learning. 2015, vol.37.
- [10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. 2014, vol.15, p. 1929-1958.