

# 教師なし学習による鉛筆画風画像生成

小林 剛豊†

茅 暁陽‡

山梨大学 大学院院医工農学総合教育部†

山梨大学大学院総合研究部‡

## 1 はじめに

本研究ではスタイル・トランスファーを行い、写真から実際の鉛筆画に近い鉛筆画風画像を生成することを目的としている。スタイル・トランスファーの代表的な手法として、参照画像と入力画像を入力し、参照画像のスタイルを入力画像に転写する **Neural style transfer**[1]や、大量のスタイル転写前と転写後の画像ペアを教師データとする **Generative Adversarial Networks (GAN)** [2]を用いる **Pix2pix**[3]が知られている。Neural style transferでは学習済みの **vgg19**[4]から入力画像のコンテンツ(画像の構造)と参照画像のスタイル(画像のテクスチャ)の情報を分離して抽出できることを利用し、ホワイトノイズ画像からスタートし、そのスタイルとコンテンツをそれぞれ入力画像と参照画像に近づけさせていく学習を行うことで参照画像のスタイルを入力画像に転写する。この方法は1枚の参照画像のみでスタイル・トランスファーできる利点をもつが、学習に計算時間を要することに加え、それぞれのスタイルと画像に適した学習パラメータの調整が困難であり、多くの場合期待する効果が得られない問題がある。Pix2Pix ではいったん学習を終えれば、どの入力画像に対しても一定の効果が得られ、高速処理ができるが、モデルの学習に十分な教師ペアを集めることは通常困難である。この2つの手法の欠点を補うために、本研究では教師なし学習が可能な生成ニューラルネットワークとして近年注目を浴びている **CycleGAN**[5]を用いて写真から鉛筆画風画像を生成する方法を提案する。

でモデルの学習を行った。図 2(d)は図 2(c)を参照画像としたときの Neural style transfer の結果である。CycleGAN を用いた方法がより自然な鉛筆画風画像を生成できていることがわかる。

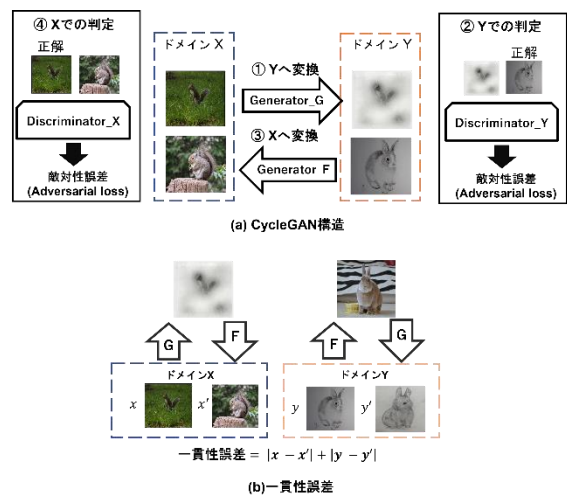


図 1 CycleGAN 概要



図 2 CycleGAN と Neural style transfer の生成結果の比較

## 2 提案手法

### 2.1 CycleGAN による鉛筆画生成

図 1 に示すように CycleGAN では Pix2pix と同様に、生成器による変換画像を判別器によって判定して、判定結果から敵対性誤差を計算して生成器と判別器の学習に用いる。CycleGAN が Pix2pix と異なる点は通常の変換を行うネットワークに加え、逆の変換を行うネットワークも構造に含まれていることである。図 1(a)に CycleGAN の構造を示す。①と②の生成器と判別器で元画像 X を結果画像 Y に変換した後、逆方向の変換を行う③と④の生成器と判別器で Y を X に戻す。逆変換で得られる画像と元画像の差分を一貫性誤差(図 1(b))と言い、この一貫性誤差を用いることで、変換前後の画像ペアを教師データとして用意しなくても、ネットワークの学習が可能である。図 2 に CycleGAN と Neural style transfer による鉛筆画風画像生成結果を示す。学習データとして、ペアとなっていない写真と鉛筆画の画像、それぞれ 150 枚使用し、100epoch

### 2.2 抽象化表現

鉛筆画の特徴の一つとして、画面全体を一様に描かず、モチーフにかかわる重要な部分だけを詳細に描き、それ以外の部分を省略する抽象化という技法がある。抽象化は描く対象にアクセントが付き、受け手に作品のモチーフを効果的に伝えることができる。本研究は顕著性ネットワークを用いて重要領域を自動抽出し、抽象化された鉛筆画風画像の生成を試みる。

Li らはスタイル・トランスファーを行う際に画像内のより顕著な特徴を保存するために、画像中の顕著な(重要な)領域を予測するネットワーク(顕著性ネットワーク)と CycleGAN を組み合わせた **SDP-GAN**[6]を提案した。図 3 に SDP-GAN の構造を示す。顕著性ネットワークに対して、既存の学習済み顕著性マップ生成モデルの結果を教師データとして学習を行い、得られた特徴を CycleGAN の生成器 G と共有させることで、学習が進むにつれて、G は顕著な特徴のみを画像生成に使用することになる。CycleGAN の学習には、2 節

「Generation of pencil style image by unsupervised learning」

† 「Taketo Kobayashi, Graduate School of Engineering, University of Yamanashi」

‡ 「Xiaoyang Mao, Research Department, Graduate School, University of Yamanashi」

で説明した敵対性誤差と一貫性誤差に加え、抽象化誤差も用いる。本研究ではこの抽象化誤差の新しい算出方法を提案することで、抽象化した鉛筆画風画像の生成を実現する。

まず、鉛筆画の場合は顕著な領域は黒い鉛筆で描かれ、省略される領域は白い紙のままであるということを考慮し、顕著性ネットワークから得られる顕著性マップを反転させ、顕著な領域が黒く、それ以外の領域が白である画像を抽象化誤差の計算に使用する。次に、鉛筆画のテクスチャには紙の凹凸によって削られる鉛の量に対応する濃淡のばらつきがあることを考慮し、生成器Gの出力についてはガウシアンフィルタを適用してから誤差の計算に用いる。著性マップSを反転させた画像をS'、生成器Gの出力Oにガウシアンフィルタを適用した結果をO<sub>Ggauss</sub>とすると、抽象化誤差は以下のように定義される。

$$L_{abstract} = O_{gauss} - S' \quad (1)$$

$$O_{gauss}(x, y) = \sum_i \sum_j \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{i^2 + j^2}{2\rho^2}\right) * O(x + i, y + j) \quad (2)$$

$$S'(x, y) = 255 - S(x, y) \quad (3)$$

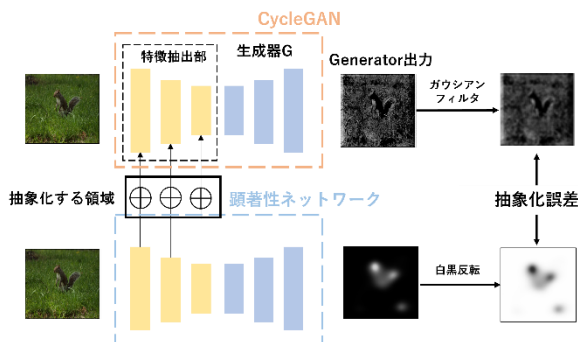


図3 SDP-GAN ネットワーク

実験では学習データとして、CycleGANにはペアでない写真と鉛筆画をそれぞれ150枚、顕著性ネットワークには写真に対して、SalGAN[7]という学習済みモデルを適用し、生成した顕著性マップ150枚を用いた。提案するネットワーク全体で100epochを学習し、生成した鉛筆画風画像の例を図4(d)に示す。提案手法では、顕著性マップ(b)における顕著な領域の周辺だけを詳細に描き、それ以外を省略してスタイル・トランスファーを行うことができていることがわかる。

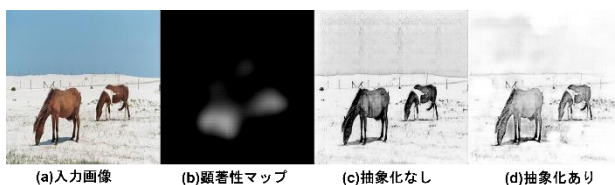


図4 提案手法による鉛筆画風画像

### 2.3 ストローク効果の付与

図5(a)のストローク例に示すように、鉛筆画を描く際には、ストロークの方向を変えながら、物体の形状や質感を表現する。本研究では提案

するネットワークモデルに Neural style transfer のスタイル誤差を加えることでこのようなストローク効果を与えるを試みる。スタイル誤差は式(4)により計算される。ここでG<sub>O</sub><sup>l</sup>, G<sub>S</sub><sup>l</sup>は出力画像Oと参照画像Sのvgg19の各層み込み層lから得られる特徴である。今回はSDP-GANの生成器Gの出力と図5(b)の参照画像からスタイル誤差を計算し、これをストローク誤差とする。ストローク誤差を加えて学習した提案モデルの結果の例とスタイル誤差を用いない場合の結果の例を図5(c), (d)に示す。学習の条件は2.3節と同様である。斜め一方向であるが、ストローク効果が得られていることがわかる。

$$L_{stroke} = \frac{1}{\sum_l} \sum_l (G_O^l - G_S^l)^2 \quad (4)$$

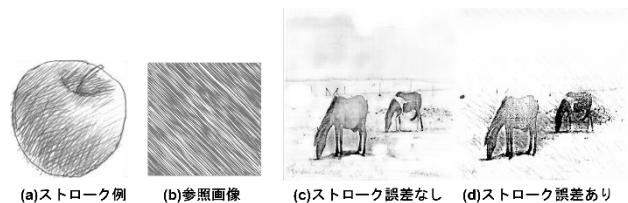


図5 ストローク誤差による学習結果

### 3. おわりに

教師なし学習によりスタイル・トランスファーを行う CycleGAN に顕著性ネットワークを加えることで、画像の顕著性を考慮して抽象化された鉛筆画風画像を生成する方法を提案した。さらに、ストロークを表す参照画像とスタイル誤差を導入することで、ストローク効果を与えることにも成功した。今後の研究では画像の特徴に合わせてストロークの方向や太さなどのパラメータ自動変更できる方法を検討していく予定である。

### 4 参考文献

- [1] Leon A. Gatys, Alexander S. Ecker, "Image Style Transfer Using Convolutional Neural Networks", CVPR 2016
- [2] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, "Generative Adversarial Networks", NIPS 2014
- [3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks", CVPR 2017
- [4] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", CVPR 2015
- [5] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks", ICCV 2017
- [6] Ru Li, Chi-Hao Wu, Shuaicheng Liu, Jue Wang, Guangfu Wang, Guanghui Liu, Bing Zeng, "SDP-GAN: Saliency Detail Preservation Generative Adversarial Networks for High Perceptual Quality Style Transfer", IEEE transaction on image processing 2021
- [7] Junting Pan, Cristian Canton-Ferrer, Kevin McGuinness, Noel E. O'Connor, Jordi, Elisa sayrol, Xavier Giro-i-Neito, "SalGAN: visual saliency prediction with adversarial networks", CVPR 2017