

### CNNに基づくAnimeGAN改良の提案

王, 爽 / Ou, sou

---

(出版者 / Publisher)

法政大学大学院理工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学研究科編

(巻 / Volume)

65

(開始ページ / Start Page)

1

(終了ページ / End Page)

3

(発行年 / Year)

2024-03-24

(URL)

<https://doi.org/10.15002/00030736>

# CNNに基づく AnimeGAN 改良の提案

## PROPOSAL FOR ANIMEGAN IMPROVEMENT BASED ON CNN

王爽

Ou sou

指導教員 李磊

法政大学大学院理工学研究科応用情報工学専攻博士前期課程

In this paper, we propose a method to address the discoloration issue in the AnimeGAN model. Our methodology employs Convolutional Neural Networks (CNN) to sustain the chromatic integrity of the original image. According to the experimental results, the discoloration problem of AnimeGAN has been rectified, successfully preserving the color of the original image.

**Key Words** : NST, AnimeGAN, CNN, Lab

### 1. はじめに

近年、実写画像をアニメスタイルの画像に変換するスタイル変換が注目されている。本稿では、畳み込みニューラルネットワーク (CNN) を利用して Lab 色空間で着色することにより AnimeGAN の変色問題を改良する。

### 2. NST と VGG ネットワーク

#### (1) NST

NST (ニューラル・スタイル・トランスファー) とは、深層ニューラルネットワークを使用することで画像変換を行うことである。

#### (2) VGG ネットワーク

VGG ネットワークは CNN の一種で、2014 年に提案された。16 層のバージョンと 19 層のバージョンがあり、当時の CNN としては圧倒的に多くの層が重なっている。VGG ネットワークは基本的には畳み込み層とプーリング層から構成される深層ネットワークで、大規模なデータセットでの学習に成功しており、それによって特徴を捉えることができる。VGG ネットワークの提案により画像のスタイル変換を実現できるようになった。具体的には参照画像のコンテンツと別の画像のスタイルの特徴を抽出し、それぞれは元の画像と比較してスタイル損失とコンテンツ損失を計算し、総合的な損失を最小化にする。最終的に生成される画像は、元画像の一般的な内容を保持する上に、参照画像のスタイルを付加することもできるようになる。

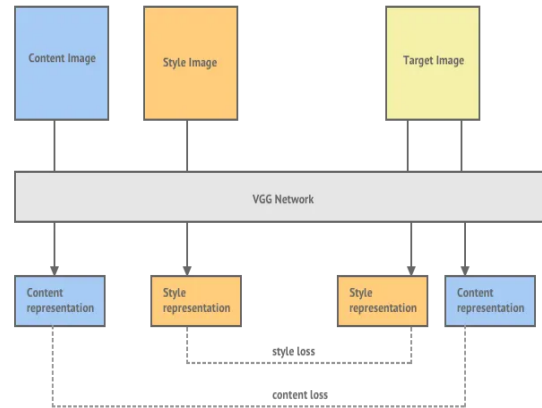


図 2-1 VGG ネットによる画像のスタイル変換

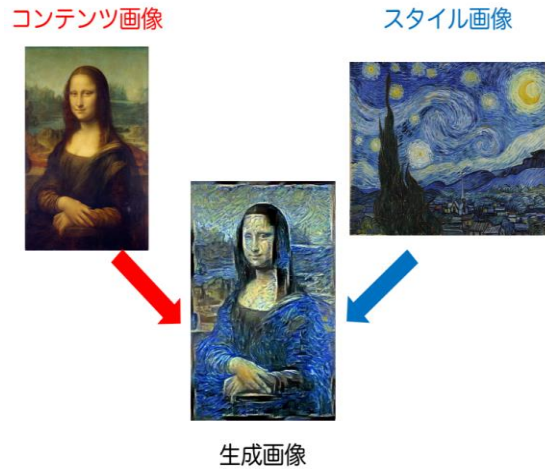


図 2-2 画像のスタイル変換

### 3. AnimeGAN

#### (1) アーキテクチャ

AnimeGAN は現実の写真をアニメ風に変換する軽量モデルである。GAN モデルの一種として生成ネットワーク (Generator) と識別ネットワーク (Discriminator) と呼ばれる 2 つのネットワークが競い合う仕組みを持っている。生成ネットワークは VGG19 ネットワークによって抽出される特徴を元にスタイル変換を行い、できるだけ識別ネットワークを欺くようにアニメ風の画像を生成する。識別ネットワークは生成されたアニメ風の画像と実際のアニメ画像とを区別するために学習する。

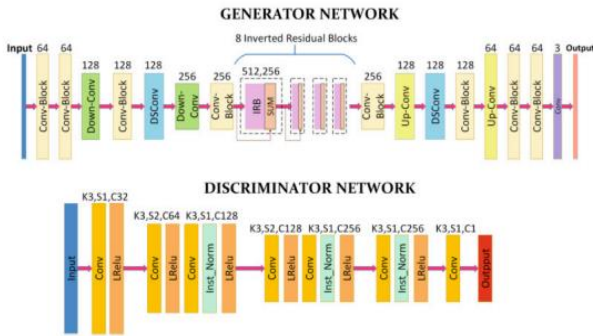


図 3-1 AnimeGAN のアーキテクチャ

#### (2) 損失関数

##### a) 生成ネットワーク

対抗損失 (Adversarial Loss) : 対抗損失は生成ネットワークが生成した画像か本物の画像かを識別ネットワークに区別されないように学習する。

コンテンツ損失 (Contents Loss) : VGG19 の畳み込み層の出力を使用し、生成された画像と元画像との間でコンテンツ損失が計算し、生成画像が元画像と構造的に類似しているように学習する。

グレースケールスタイル損失 (Grayscale Style Loss) : よりジャンプなテクスチャーとラインを備えるためである。グレースケールスタイル変換において、VGG19 の畳み込み層の出力を使用してスタイル損失が計算することができる。

色再構築損失 (Color Loss) : 元の画像の色を維持するために、YUV 色空間で色再構築を行う。

##### b) 識別ネットワーク

識別ネットワークの損失関数において、対抗損失はできるだけ生成された画像かリアルな画像かを区別できるように学習する。また、AnimeGAN は CartoonGAN で提案されたエッジ促進損失を組み込むことで、生成される画像によりはっきりとした輪郭を与える。同時に、グレースケール化の影響により生成画像がグレースケールになることを防ぐため、識別ネットワークはグレースケール対抗損失も提案した。

### 4. CNN に基づく AnimeGAN 改良の提案

AnimeGAN はグレースケールアニメ画像を用いてトレーニングしたため、生成された画像が変色や暗くなる傾向がある。従来の方法は YUV 色空間で色再構築とグレースケール対抗損失を使用して元画像の色を維持したが、効果が良くない。そこで、今回は人間の視覚を近似しますよう設計された Lab 色空間を利用する。まずは入力画像を L (明度), a と b (色度) に分離する。その後は生成ネットワークに L チャンネルだけを入力し、出力したグレースケールのアニメ画像をトレーニング済みの CNN に入力し、元の a と b チャンネルを予測して付加し、色付けられるアニメ風の画像が生成される。

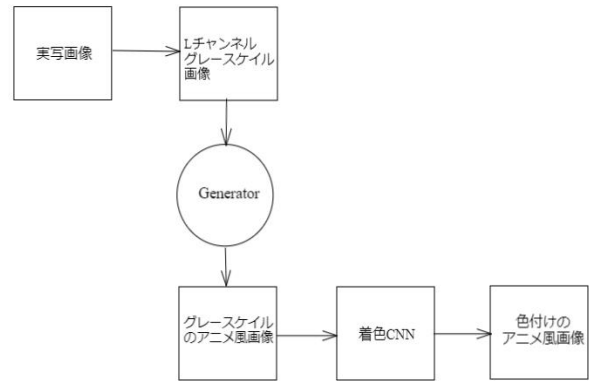


図 4 アーキテクチャ

P: リアル写真のグレースケール画像

A: アニメ画像

X: グレースケールアニメ画像

Y: エッジ抜きグレースケール化アニメ画像

コンテンツ損失関数とグレースケール対抗損失の数式は (41) に示す。

$$L_{con}(G, D) = E_{p_i \sim S_{data}(p)} [\|VGG_i(p_i) - VGG_i(G(p_i))\|_1]$$

$$L_{gra}(G, D) = E_{p_i \sim S_{data}(p)}, E_{x_i \sim S_{data}(x)} [\|Gram(VGG_i(G(p_i))) - Gram(VGG_i(x_i))\|_1] \quad (41)$$

Generator 総損失は (42) に示す。

$$L_G = \omega_{adv} E_{p_i \sim S_{data}(p)} [(G(p_i) - 1)^2] + \omega_{con} L_{con}(G, D) + \omega_{gra} L_{gra}(G, D) \quad (42)$$

Discriminator 総損失は (43) に示す。

$$L(D) = \omega_{adv} [E_{a_i \sim S_{data}(a)} [(D(a_i) - 1)^2] + E_{p_i \sim S_{data}(p)} [(D(G(p_i)))^2] + 0.1 E_{y_i \sim S_{data}(y)} [(D(y_i))^2] \quad (43)$$

### 5. 実験結果

トレーニングデータは AnimeGAN と同じデータセットを使用する。



図5 結果(左:写真, 中: AnimeGAN, 右: 改良後)

## 6. 評価

### (1) 主観評価

AnimeGAN の出力結果と改修後の出力結果に対し、100 人に対してアンケート調査を行う、結果は以下の Table 6-1 に示す。

Table 6-1 アンケート

| モデル/<br>ImgNo | AnimeGAN<br>が良い | CNN に基づく<br>AnimeGAN が良い |
|---------------|-----------------|--------------------------|
| 1             | 24              | 76                       |
| 2             | 17              | 83                       |
| 3             | 33              | 67                       |

### (2) 客観評価

改修前と後の二つの手法において、それぞれ Lab 色空間での原画像と出力画像との誤差を比較することで、知覚的な誤差評価に近い結果を得ることができる。今回は CIEDE2000 の評価方法を利用して差異を計算する。CIEDE2000 公式は下記の式で示す。

$$\Delta E_{00} = (\Delta L')^2 + (\Delta C')^2 + (\Delta H')^2 + R_T(\Delta C')(\Delta H') \quad (61)$$

$\Delta L'$  は 2 つの色の間の明度の差、 $\Delta C'$  は 2 つの色の間の色度の差、 $\Delta H'$  は 2 つの色の間の色相の差、 $R_T$  は補正パラメータで、 $(\Delta C')(\Delta H')$  は色相と色度の相互作用による補正である。 $\Delta E_{00}$  は NCD 誤差 (Normalized Color Difference) と呼ばれる。今回は入力画像と改修前後の出力結果との比較を行い、それぞれの画像にいくつか位置がマッピングするサンプルのピクセル抽出、NCD 誤差を計算して平均値を取った結果は以下の表に示す。

Table 6-2 NCD 差異

| モデル/<br>ImgNo | AnimeGAN<br>の $\Delta E_{00}$ | CNN に基づく<br>AnimeGAN の $\Delta E_{00}$ |
|---------------|-------------------------------|--|
| 1             | 19.39                         | 5.37                                   |
| 2             | 18.56                         | 4.23                                   |
| 3             | 20.45                         | 4.76                                   |

主観評価のアンケートの結果からみると、CNN に基づく AnimeGAN は良い評価である。客観評価に基づき、入力画像の色を保持方面からすると、CNN に基づく AnimeGAN の改良バージョンは元の AnimeGAN より優れている。

## 7. まとめ

本稿では、実写画像をアニメ風の画像を高速的に変換する AnimeGAN モデルの変色問題を改善した。Lab 色空間で三つの成分がそれぞれ独立している特徴を利用し、L チャンネルを入力し、生成された画像を CNN による着色することにより変色問題を改善する。検証実験では今回は AnimeGAN と同じデータセットを使用する。結果の評価について、主観評価はアンケートを取り、客観評価は改良前後の生成画像と元画像の色の差異の計算は CIEDE2000 を使用した。両方とも良い結果を得られた。

謝辞：本研究を遂行するにあたり、多大なるご指導を頂いた法政大学李磊教授に厚くお礼を申し上げます。さらに、助言を頂いた本研究室の方々にも感謝いたします。

### 参考文献

- 1)Jie Chen, Gang Liu, Xin Chen "AnimeGAN: A Novel Lightweight GAN for Photo Animation." ISICA 2019:
- 2)Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. 2017 : arXiv:1703.10593
- 3)Yang Chen, Yu-Kun Lai, YongJin Liu.CartoonGAN: Generative Adversarial Networks for Photo Cartoonization, 10.1109/CVPR2018.00986
- 4)Phillip Isola; Jun-Yan Zhu; Tinghui Zhou; Alexei A. Efros. "Image-to-image translation with conditional adversarial networks" 2017: arXiv:1611.07004
- 5)<https://www.freecodecamp.org/news/colorize-b-w-photos-with-a-100-line-neural-network-53d9b4449f8d/>
- 6)[https://qiita.com/masaka\\_programming/items/fa90ca0275a62594706b](https://qiita.com/masaka_programming/items/fa90ca0275a62594706b)
- 7)<https://medium.com/data-science-group-iitr/artistic-style-transfer-with-convolutional-neural-network-7ce2476039fd>
- 8)Hu Haonan "Proposal For AnimeGAN improvemen Based On HSVcolorspace"2022<http://doi.org/10.15002/00025375>
- 9)Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio"Generative Advirsarial Networks" arxiV:1406.2661
- 10)<https://blog.negativemind.com/2019/06/22/generative-adversarial-networks/>