

# ニューラルネットワークを用いたアニメ原画に対する自動彩色手法の開発

南谷 大輔† 米澤 弘毅‡

名城大学理工学部情報工学科† 名城大学情報工学部情報工学科‡

## 1 はじめに

近年、アニメ文化の広がりによりたくさんのファンが増え、多くのアニメが製作されている。また、コンピューター技術の発達により 3DCG などの手法が取り入れられ、作業の全てを手作業で行わず、コンピューターを使って作品が製作されるようになってきた。手作業の工程をコンピューターが行うことで、必要な労力を削減しようとアニメ業界でも様々な取り組みが行われている。

ニューラルネットワークを用いた自動彩色については、マンガを対象にした研究が多く行われているが、アニメを対象とした研究はあまり行われていない。本研究では、入力としてアニメーションの線画とカラーの参考画像を与え、彩色された画像を出力するような転移学習手法の開発を目標とする。

## 2 先行研究

### 2.1 Tag2Pix[1]

この手法は、タグ情報から線画の彩色を行う手法である。GAN(敵対的生成ネットワーク)が基になっている。タグ情報には、「red hair」のように彩色したいパーツと対応する色の情報が紐づけられている。タグ情報は Generator の Decoder 部分にて色情報として線画と結合される。結合には彼らが開発した SECat block が使われている。

### 2.2 参考画像を用いた構造認識によるスケッチのカラー化[2]

この手法は、参照画像から線画の彩色を行う手法である。GAN(敵対的生成ネットワーク)が基になっている。参照画像から色情報を抽出した後、Generator の Decoder 部分にて色情報として線画と結合される。色情報との結合には Adaptive Instance Normalization[3]が使われている。

## 3 提案手法

### 3.1 提案手法の概要

提案するネットワーク図を図 1 に示す。提案手法では、参考画像から色情報を生成し、畳み込みされた入力画像との結合を行う。色情報との結合が4回行われた後で、出力画像が生成される。画像のピクセル数は入力画像、参考画像、出力画像ともに512×512である。Discriminator と損失関数は Tag2Pix と同じものを使用した。

### 3.2 色情報の生成

色情報はカラーの参考画像から生成される。参考画像は変形された後にネットワークへ入力され、入力画像と参考画像の構図が異なる場合でもきれいに彩色させる。参考画像の変形処理はネットワークの学習時のみに行われ、テスト時には行われない。

### 3.3 入力画像と色情報の結合

入力画像と色情報の結合は Decoder の部分で行われる。色情報との結合には Tag2Pix で提案されている SECat block を使用した。

## 4 検証結果

### 4.1 使用するデータセット

ネットワークの学習には、Anime Sketch Colorization Pair データセット[4]を利用した。このデータセットには、カラー画像とそれに対する線画がセットになっており、全部で17769枚の画像がある。その中から14424枚をネットワークの学習用、3545枚をテスト用として利用した。

### 4.2 出力画像の検証

出力画像の検証にはネットワークの学習で使用しなかった3545枚の画像を使用した。

最初に図1のネットワークにおける色情報の次元数を64次元、128次元、256次元と変えたときに出力画像がどのように変化するかを検証した。出力画像を見てみると、64次元の場合は、違った色で彩色されることが多い。256次元の場合は、違った色で彩色されることは少ないが鮮やかさがなく、肌色や灰色で彩色されることが多かった。128次元の場合は、違った色で彩色さ

Towards automatic colorization with reference-guided neural network

†Daisuke NANYA · Meijo University

‡Kouki YONEZAWA · Meijo University

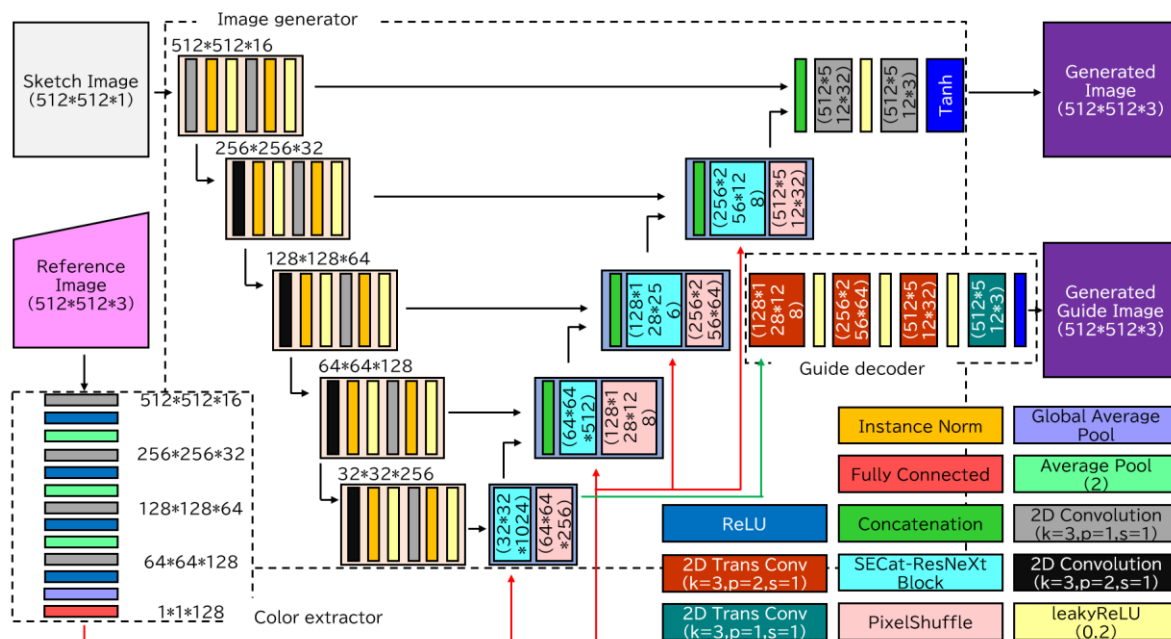


図 1 提案するニューラルネットワークの構成

れることは少なく鮮やかさも保っており、128 次元が最も自然な彩色を行えると結論付けた。

次に各手法における定量的評価を行った。Liu らの手法、pix2pix、提案手法の FID [5] を表 1 に示す。Liu らの手法は[2]の論文の中にあつた値を記載した。また、その論文で使用されているデータセットを用いて各手法における FID の測定を行った。pix2pix は参考画像を入力することができないため、入力画像として線画のみをネットワークに入力した。pix2pix と提案手法は Epoch 数を 40 で学習した。

表 1 各手法における定量的評価

手法	FID
Liu らの手法 ([2]より)	27.99
pix2pix	35.80
提案手法(色次元数:64)	16.15
提案手法(色次元数:128)	15.68
提案手法(色次元数:256)	16.11

正解画像がないため FID を用いた定量的評価はできないが、最後に入力画像と参考画像の構図が異なる場合の出力画像を確認する。入力画像と参考画像の構図が似ている場合は、パーツに合った色を入力画像に彩色することができた。一方で構図が似ていない場合は、キャラクターの部分に同じような色で全体的に色が付いてしまい、彩色に失敗した。

## 5 おわりに

本研究では、アニメ原画の彩色を目的としてネットワークの構築を行った。参照画像の変形

処理を行うことで、似ている構図であれば彩色ができることを確認できた。今後は、構図が似ていなくても彩色が行えるようなネットワークの構築を目指したい。

## 参考文献

- [1] H. Kim, H.Y. Jhoo, E. Park, S. Yoo, Tag2Pix: Line Art Colorization Using Text Tag With SECat and Changing Loss, in: 2019: pp. 9056–9065. [https://openaccess.thecvf.com/content\\_ICCV\\_2019/html/Kim\\_Tag2Pix\\_Line\\_Art\\_Colorization\\_Using\\_Text\\_Tag\\_With\\_SECat\\_and\\_ICCV\\_2019\\_paper.html](https://openaccess.thecvf.com/content_ICCV_2019/html/Kim_Tag2Pix_Line_Art_Colorization_Using_Text_Tag_With_SECat_and_ICCV_2019_paper.html) (accessed April 20, 2022).
- [2] X. Liu, W. Wu, C. Li, Y. Li, H. Wu, Reference-guided structure-aware deep sketch colorization for cartoons, *Comp. Visual Media*. 8 (2022) 135–148. <https://doi.org/10.1007/s41095-021-0228-6>.
- [3] Xun Huang, Serge Belongie, Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization, (n.d.). <https://arxiv.org/pdf/1703.06868.pdf>.
- [4] Taebum Kim, Anime Sketch Colorization Pair, Kaggle.Com. (n.d.). <https://www.kaggle.com/datasets/ktaebum/anime-sketch-colorization-pair>.
- [5] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium, in: *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. <https://proceedings.neurips.cc/paper/2017/hash/8a1d694707eb0fefe65871369074926d-Abstract.html> (accessed April 21, 2022).