

A Proposal for Improvement of the AnimeGAN Based on CNN

Shuang Wang and Lei Li

*Major in Applied Informatics, Graduate School of Science and Engineering,
Hosei University, Koganei, Tokyo 184-8584 Japan
E-mail: lilei@hosei.ac.jp*

Abstract

In this paper, we propose a method to address the discoloration issue in the AnimeGAN model. Our method uses Convolutional Neural Networks (CNN) to sustain the chromatic integrity of the original image. We separate the image in the Lab color space, where the components of luminance (L), green-to-red variation (a), and blue-to-yellow gradient (b) exist independently. We use the L channel as the input for the AnimeGAN and color the grayscale image by predicting (a) and (b) channels with CNN. According to the experimental results, the discoloration problem of the AnimeGAN has been rectified, successfully preserving the color of the original image.

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

1 はじめに

1.1 研究背景

漫画は、絵と言葉を組み合わせた形で物語や情報を表現する日本発祥の表現媒体である。最初は江戸時代に絵草紙と呼ばれる絵入りの説話や物語が広く流通したが、昭和時代に漫画が大衆文化の一環として定着し、漫画作家も多く登場し、少年漫画や少女漫画が隆盛を迎えました。現在は日本の漫画は国内外で広く愛され、人々の生活にエンターテインメントを提供する一方、異なる文化の交流も促進できるため、私たちの生活に離れない存在になっている。

しかし、アニメを手動で作成することは非常に時間が長く、作家は手描きアニメーションやデジタル作画のスキルを磨くことは相当に時間がかかる。高品質なアニメ作品を作成するには、線、テクスチャ、色彩、陰影の慎重な考慮が必要であり、これは作品を作成するのが難しく、時間がかかるため、実世界の写真を高品質なアニメに自動的に変換できる自動化技術が求められる。

2014年にVGGはKaren SimonyanとAndrew Zissermanによって提案された。VGGネットワークは畳み込みニューラルネットワーク(CNN)の一種である。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、それによって広範な特徴を捉えることができる。VGGネットワークは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ネットの特徴はシンプルなアーキテクチャを持ち、また3×3のカーネルサイズの小さな畳み込みを多く利用していることである。この3×3のサイズの小さな畳み込みを利用し、チャンネル数を増やすという構造はResNetなどVGGネット以降の様々なモデルで利用されるようになった。