

< 修 士 論 文 >

画像生成モデルの適応的パラメータ
入れ替えを用いたイラストの画
風変換

(要 旨)

滋 賀 大 学 大 学 院
デ ー タ サ イ エ ン ス 研 究 科
デ ー タ サ イ エ ン ス 専 攻

修了年度：2022年度

学籍番号：6021124

氏 名：中江 剛之

指導教員：飯山 将晃

提出年月日：2023年1月11日

1. はじめに

画像変換は、画像のドメインを別の画像のドメインに変換する手法である。例えば、実世界の風景の写真を入力として、その構造を保ったままモネが描いたような風景画のように変換する処理が挙げられる。近年はこの画像変換を応用して、表情を変換するアプリケーションや人間の顔をアニメ風に変換するアプリケーションなどが登場し、エンターテインメント分野での活用が進んでいる。この活用は実世界の画像だけに限った話ではなく、イラストや漫画等にも需要がある。例えば下書きのラフを仕上げの状態まで変換することで、イラストを描く手間を省いたり、描く際の参考資料にすることも可能になる。

イラスト同士の画像変換では、画像に映る顔の特徴を維持したまま画風を変換することでアイデンティティを維持することが重要になる。もし変換時に被写体の顔の特徴が変わると描写されている人物が変化し、肖像画風にするという目的が達成できなくなる。イラスト同士での画像変換でも同じことが言え、画風とアイデンティティの要素が重要となる。しかし変換器を用いる既存手法の画像変換では、変換時にアイデンティティを喪失する問題があるため、アイデンティティを維持しつつ画風を変換することができない。

2. StyleGAN2 のパラメータ入れ替えについて

変換器を用いる方法以外に、画像生成モデルである StyleGAN2 の生成器のパラメータを入れ替えて画像変換を実現する手法がある。StyleGAN2 の生成器は多層構造になっており、入力に近い低解像度の生成層ほど生成画像の構造を制御し、出力に近い高解像度の生成層ほど生成画像のテクスチャを制御する特徴となっている。

この特徴を応用するとそれぞれの生成画像の特徴を引き継いだ画像を生成でき、画像変換を行える。例えば、生成器の構造が同じ少年漫画の画像生成器 A と少女漫画の画像生成器 B があり、この二つから新しい生成器 AB を作成することを考える。新しい生成器 AB の低解像度層を少年漫画の画像生成器 A の低解像度層のパラメータに設定し、生成器 AB の高解像度層に少女漫画の画像生成器 B の高解像度層のパラメータを設定すると、生成器 AB の生成画像は少年漫画のキャラクターだが画風は少女漫画の新たな画像を生成できる。この仕組みを応用した画像変換を Swapping と呼ぶ。またこの変換はパラメータを入れ替えるだけでなく、混合することでも変換が可能である。

3. 提案手法

本研究では StyleGAN2 のパラメータを入れ替えずに混合することで画風変換を行い、さらにグリッドサーチを用いてパラメータの混合比を自動で調整し、各画像に最適な画風の変換を見つける。この手法を Mixing と呼ぶ。

パラメータの混合比の探索はアイデンティティと画風の評価を行う損失関数の組み合わせで行われ、変換画像候補を損失関数に入力し最も小さい値を出力した画像を変換画像として出力する。アイデンティティの評価は LPIPS で行われ、スタイルの評価は作品を分類する分類器の尤度と変換目標のドメインとのラベルの誤差を計算したもので行われる。

また本研究では StyleGAN2 の生成器から 256×256 サイズの画像を生成し変換を行うため、生成器は $4 \times 4 \cdot 8 \times 8 \cdot 16 \times 16 \cdot 32 \times 32 \cdot 64 \times 64 \cdot 128 \times 128 \cdot 256 \times 256$ の層から成り立っている。これらの層に、それぞれ混合比率候補 $[0, 0.25, 0.5, 0.75, 1]$ の 5 つの探索を行うが、全ての層に対して 5 つの候補によるグリッドサーチの探索を行うと膨大な探索量になるため、探索量削減のため探索する層の比率は 16×16 と 32×32 と 64×64 の層に限定した。

4. 実験結果

本研究の実験では Manga109 と呼ばれる漫画データセットを使用し、「プリズム・ハート」と「Oh!われら劣等生徒会」の 2 作品の顔画像を用いて変換を行った。先行研究との比較では、CycleGAN・U-GAT-IT・Swapping と提案手法 Mixing との比較を行う。

本研究はアイデンティティを維持しつつ画風を変換することを目的とするためアイデンティティの維持に関する評価と、画風の評価、そして画像変換の品質の評価を行う。しかしこれらの評価は定量的にできないため、ユーザースタディによる定性的な評価を行う。本研究のユーザースタディは、アイデンティティの評価 10 問・画風の評価 10 問・変換品質の評価 10 問の 3 セクションで構成されており、合計 30 問のアンケートとなっている。

このアンケートを 30 人に取った結果、アイデンティティの評価では提案手法 Mixing は Swapping よりアイデンティティの維持で劣る結果となったが、画風の評価・生成画像の評価では他の手法と比べ Mixing が最も良い評価を得る結果となった。

5. 終わりに

本研究では StyleGAN2 のパラメータの混合による画像変換を行った。さらに最適なパラメータの混合比を探索し、最適なアイデンティティを維持した画風変換を画像ごとに行う手法 Mixing を提案した。実験結果によると本研究の提案手法は、アイデンティティの維持では既存手法 Swapping に劣るものの、画風の変換では既存手法 Swapping より良い評価を得ることができ、変換の品質の評価も高いことがわかった。