

画像生成ネットワークを用いた顔画像の日本アニメ風変換に関する検討

T19M059 近澤 勇太

指導教員 三好 力

1. はじめに

本研究では敵対的生成ネットワーク(Generative Adversarial Nets:GAN)を用いて顔画像を日本のアニメ風キャラクターに変換することを目的とする.日本のアニメーションは主に手書きであり,アニメキャラクターは非常に高度な単純化と誇張表現を兼ね備えているため,実写の顔画像をアニメ風に変換するように GAN に学習させることは難しい.そこで顔画像のアニメ化のための新しい GAN の損失関数であるランドマーク損失を提案する.

2. 提案手法

提案するランドマーク損失は CartoonGAN やその改良版である AnimeGAN のモデルに追加する形で使用する.CartoonGAN や AnimeGAN は風景画像を対象としたものであり,コンテンツのテキストチャやエッジをアニメ風に変換することができる.提案するランドマーク損失は,これに加えて顔パーツのアニメ化を行いその位置を補正する損失関数である.ランドマーク損失では,アニメキャラクター用に学習させたランドマーク検出器を使用し,Generator が生成した画像のランドマーク座標と,変換前の正解データのランドマーク座標と比較し誤差を測る.ランドマーク検出器はアニメキャラクターに対してのみ正しくランドマーク座標を計測できるため,ランドマーク損失は顔画像をよりアニメキャラクターの特徴を持った画像に変換し,また,顔のパーツも元画像に近い位置で生成することが可能である.

ランドマークの誤差は Wing Loss を用いて計算し,ランドマークの検出には Deep Alignment Network(DAN)を使用する.ランドマーク損失は次の式で定義する.

$$L_{land}(G, D) = E_{p_i \sim S_{data}(p)} [\|DAN(G(p_i)) - land_{p_i}\|_w]$$

ここで $DAN(G(p_i))$ は生成画像のランドマーク座標を表し, $land_{p_i}$ は元画像のランドマーク座標を表す.

3. 実験

実験には実写の顔画像 6000 枚とアニメキャラクターの顔画像 5697 枚を使用し, CartoonGAN, AnimeGAN, とそれらにランドマーク損失を追加し

た 4 つの GAN で実験を行った.各 GAN の画像生成結果の一例を図 1 に示す.

AnimeGAN とランドマーク損失を追加した Aniime-

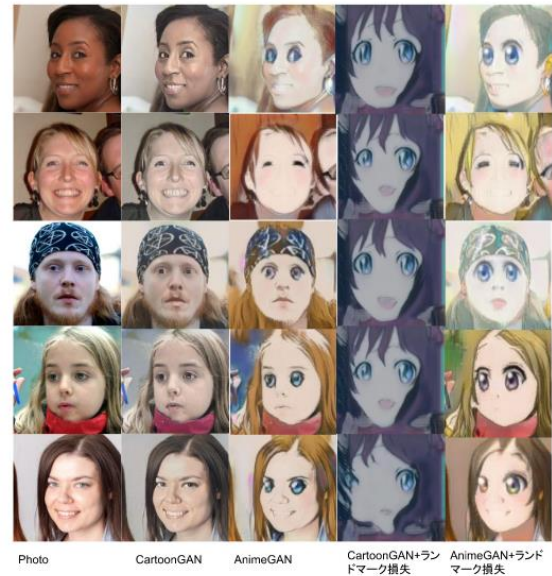


図 1 画像生成結果

GAN を比べると,目がアニメ的に強調され,テキストチャも滑らかになっていることがわかる.また,鼻や口なども簡易かされている.しかしながら, CartoonGAN とランドマーク損失の結果では,アニメキャラクター的ではあるものの同じような画像しか出力されず,元の画像の面影は残っていない.これはランドマーク検出器が,アニメのテキストチャに影響されやすく,実写の顔画像よりも全体的にアニメ風のテキストチャを持った顔画像のほうがランドマーク損失が小さくなるためであると考えられる.

4. おわりに

顔画像をアニメキャラクター風に変換するための新しい GAN の損失関数の提案を行った.提案したランドマーク損失を用いることでよりアニメキャラクターに近い顔画像を生成できることが分かった.しかしながら,各損失の調整が難しく,学習がうまくいかない場合があった.より,安定した学習を行えるように誤差の計算方法などの改善点がかんがえられる.