

styleGAN 技術を用いた顔画像からの衛生マスクの除去

T170539 美濃部 和也

指導教員 三好 力 教授

1. はじめに

新型コロナウイルスの感染拡大の伴い、衛生マスクの着用する人が増加したことで街頭写真にはマスク姿が多く写り込むようになった。マスク姿からは表情が読み取りにくいことで不安感を抱いてしまい、写真全体の印象が下がってしまう。一方で、衛生マスクをプライバシー保護の目的でも使用されることがあり、安易に素顔を晒すことはできない。そこで本研究では、実在しない人物の顔画像を生成できる StyleGAN を用いて顔画像からマスクの除去と顔の特徴変換した画像の生成を検討する。

2. 提案手法

GAN は訓練データにない物体の生成は困難である。マスクを着用した顔画像の場合、特徴が大きく生成画像に不自然さが残る。そこで、低解像度に変換して曖昧な画像として処理することを考えた。本研究では StyleGAN のエンコーダーである pixel2Style2pixel (pSp) の超解像手法を用いて低解像度顔画像から画像を生成する方法を提案する。

3. 実験

画像から顔部分を切り取り 256×256 にリサイズした顔画像を一度低解像度変換し、pSp の超解像手法を用いて顔画像を生成した。6 倍、32 倍にそれぞれダウンサンプリングすることで異なる解像度の生成を検証した。

写真から顔を検出するとき、マスクを着用していると精度が落ちる。そこで dlib が提供する畳み込みニューラルネットワーク(CNN)を採用した顔検出器を用いて従来手法と精度を比較した。また、この手法で検出した顔からランドマークの検知を行い、補正した顔画像に対して同様に低解像度に変換し画像生成を行った。

低解像度に変換した場合マスクの色が反映され生成画像に影響を及ぼす。そこで、pSp が提案するマルチモーダル合成手法を用いてスタイル変換を行う。肌や髪の色、または顔の部位をランダムな別人の特徴と組み合わせ、顔画像を生成した。

4. 結果

画像([1]より出典)から CNN モデルで検出したトリミングした顔画像を図 1 に示す。図 1 を 8 倍にダウンサンプリ

ングした顔画像を図 2 に示す。図 2 を pSp の超解像手法に入力し生成した顔画像を図 3 に示す。また、スタイル変換した顔画像を図 4 に示す。



図 1. 顔画像

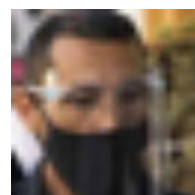


図 2. 低解像度画像

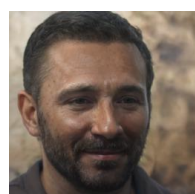


図 3. 超解像手法で生成した顔画像

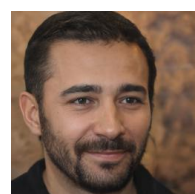


図 4. Style Mixing で生成した顔画像

5. 考察

pSp の超解像手法を用いた場合、解像度が異なる場合でもマスクを除去し、マスクで隠れている部分の表現ができたが、肌の色はマスクの影響が強くみられた。生成画像は顔の輪郭や髪形は入力画像に似た特徴がみられ、目などの細かな特徴は入力画像と異なったため、別人と認識できた。CNN モデルの顔検出精度は従来手法よりも低解像度のマスク姿の検出で高い精度を示した。顔検出や検知したランドマークに若干の誤差が確認できたが、画像の生成は可能であり精度に大きな差はなかった。スタイル変換では主に口と肌の色に対し別の特徴を組み合わせることで、肌の色などのマスクの影響による特徴を抑え、独自性を追求することができた。潜在変数はランダムで決まり、生成画像の評価も人間が行うため、特定の特徴に変換するにはコストがかかる。

出典元

[1] MATT OTT. "US services sector grows for seventh consecutive month.", Seattlepi, (2021), (参考:2021-1-16). <https://www.seattlepi.com/news/article/US-services-sector-grows-for-seventh-straight-15852746.php>