

進捗報告

1 今週やったこと

LSGAN による生成画像の考察

2 生成画像の考察

LSGAN を用いて多義図形生成を試みた。前回 1000 epoch 実行したが、今回は 5000 epoch まで実行することで学習の程度を確認するとともに ViT の識別結果及び attention map を比較した。GAN で生成した画像 50 枚を入力画像とした場合、多義図形に関する ViT の識別率は 1000,2000,3000,4000 epoch 後で 96.0%, 5000 epoch 後で 90.0% となった。

図 1～図 5 に GAN によって 1000 epoch 毎に生成された画像例を示す。また今回は入力に固定されたノイズを用いることで epoch 毎の学習を見た。図 1～図 5 より私の主観ではあるがやはり 5000 epoch 後に関しては他と比べて多義図形性を失いつつあることが見て取れるため正しく識別されているのではないかと思われる。

図 6 に正しく識別された画像の attention map の画像例を示す。図 6 より GAN で生成された解像度が低いいため attention map を適用しているが非常に見にくい状態となっている。導入として今回は LSGAN を用いたが、高解像度な画像を生成できる GAN (PGGAN など) を実装したい。

3 考えていること

前回話に出てきた GAN の Discriminator の部分を ViT に置き換えるというものを実装したい。しかし、TransGAN の論文によると Transformer ベースの Discriminator は弱いことが知られているため、あまり良い結果が出るかはわからないが、多義図形においては機能することも考えられるため実装したい。また、今回多義図形を自動生成したが、解像度が高いとは言えず見にくいためより解像度の高い画像を生成する GAN を実装したい。

多義図形に関する ViT の識別率が思っていた以上にかなり高かったので、GAN に入力画像としている画像が ViT における 訓練データ+評価データ+テストデータを合わせたものを使っているため、訓練データに似た画像が生成された場合、容易に ViT は多義図形と識別できるのではないかと懸念している。後に GAN に対する入力画像を ViT における評価データ+テストデータのみにするによる変化を見ていきたい。(この場合かなり枚数が少なくなってしまう。単に多義図形生成だけであれば訓練画像を含めて構わないと考えている。)

4 今後の方針

多義図形に適した GAN の実装, 発表用資料の作成, スライドの作成



図 1: LSGAN による 1000 epoch 後画像例

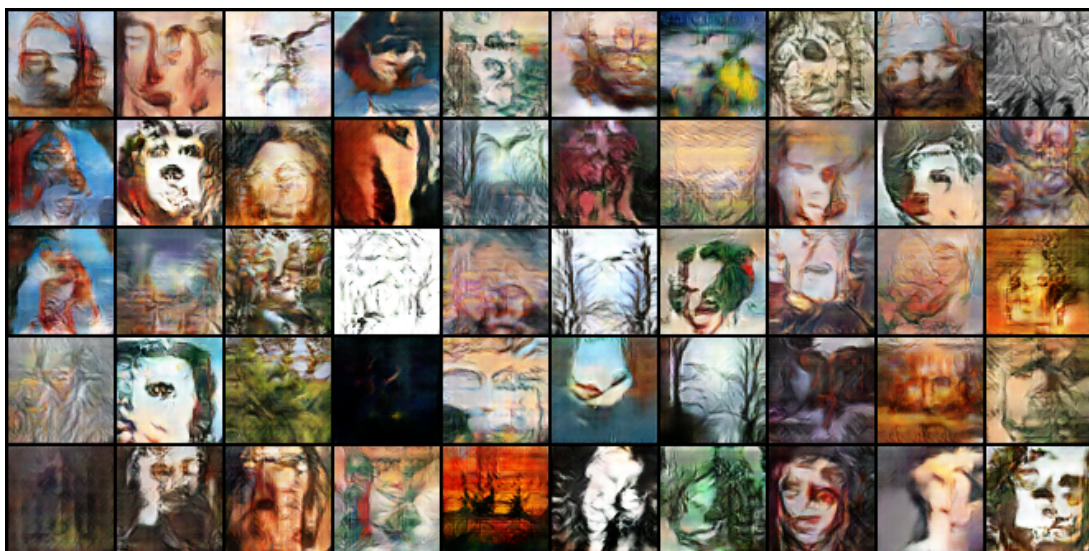


図 2: LSGAN による 2000 epoch 後画像例

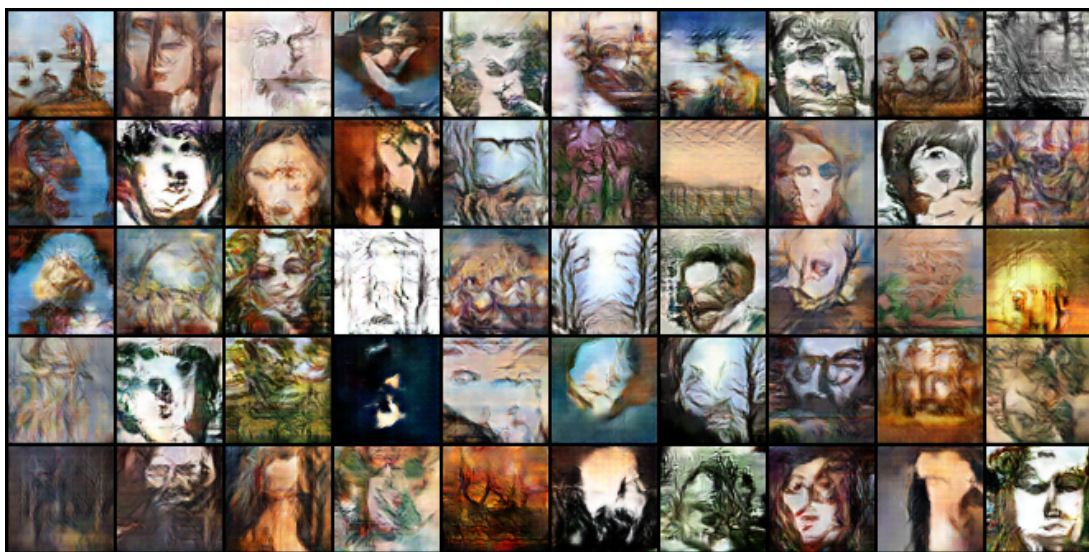


図 3: LSGAN による 3000 epoch 後画像例



図 4: LSGAN による 4000 epoch 後画像例

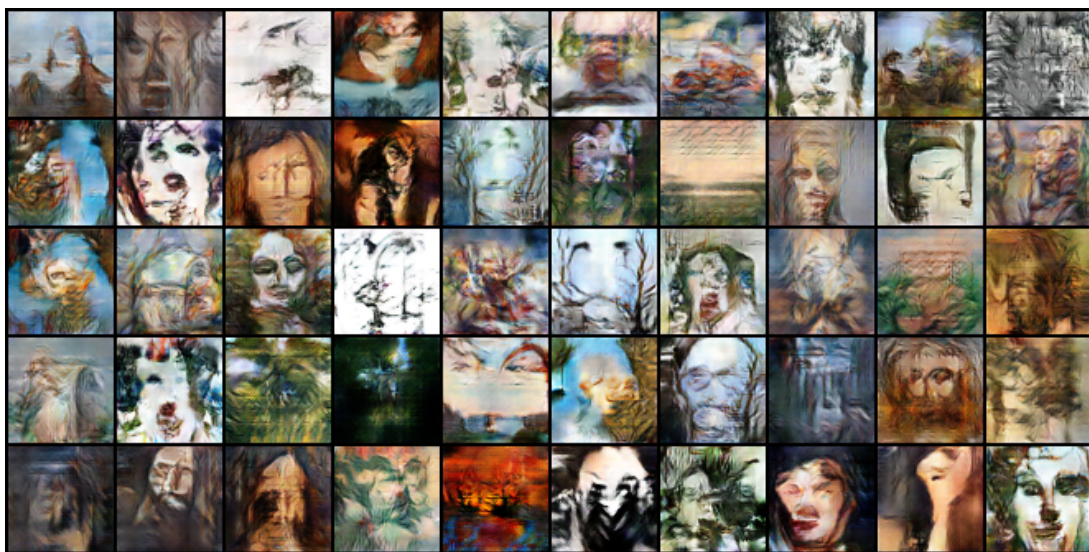


図 5: LSGAN による 5000 epoch 後画像例

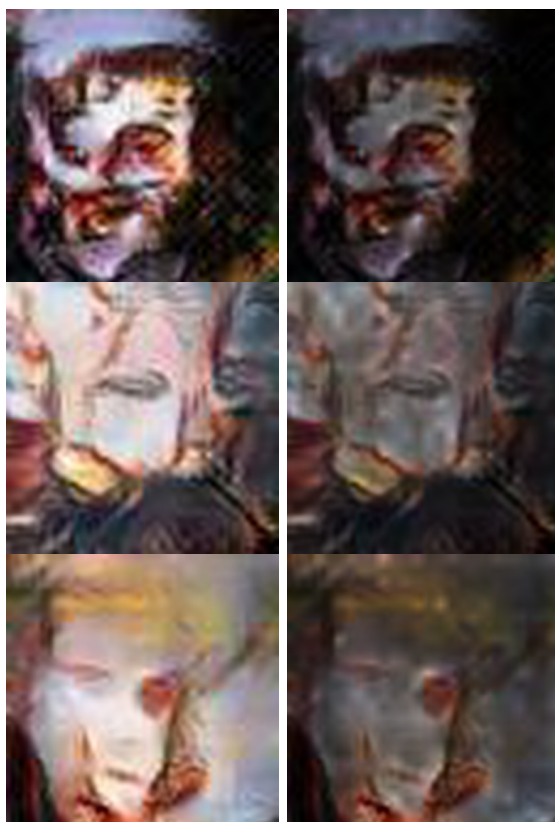


図 6: 左 : GAN による生成画像 右 : attention map 適用画像