

GAN 学習時の画像枚数削減方法の提案

尾崎 優[†]

† 横浜国立大学 理工学部

長尾 智晴^{††}

†† 横浜国立大学 大学院環境情報研究院

1. はじめに

近年, Generative Adversarial Networks (GAN)[1]では現実の写真と区別できないような高品質かつ多種多様な画像の生成が可能になりつつある。しかし, GAN の学習には大量のデータが必要であり, データの収集が大きな課題になっている。

そこで本稿では, 学習画像が少ない状況でもより高品質な画像生成を可能とする学習方法を提案する。

2. 提案手法

GAN の生成画像の品質を左右する要因の一つに潜在変数の次元数がある。そこで GAN 学習時に, 潜在変数の次元数を増減し, 学習データに合わせて最適化する。

損失の大きさから学習の進捗がわかる WGAN-gp[2]の識別器の損失を参考に, 一定 Epoch 毎に次元数を変更し, 学習を再開する。また, 増減の方向が変わった際に増減幅を1減らすことで探索範囲を狭めていく。これを繰り返し, 学習中の次元数の探索を実現する。

3. 実験設定

学習画像には LSUN[3]データセットの tower, bedroom カテゴリ, および CelebA[4]を用いる。実験に用いる枚数は5000枚とする。前処理として64×64にリサイズ, および正規化を行う。

ベースのモデルはWGAN-gpとする。通常のWGAN-gpを比較対象として実験を行う。どちらも学習は200Epoch行う。また, 提案手法は20Epoch毎に次元数の変更を行い, 増減幅の初期値は10とする。どちらも次元数の初期値は100とする。

モデルの評価には Fréchet Inception Distance(FID)[5]を用いる。この指標は値が小さいほど学習画像と類似した画像を生成できているとされる。

4. 実験結果

提案手法と比較手法の FID を表1に示す。また, 提案手法の学習中の次元数変更の様子を図1に示す。

表1より tower において比較手法より改善が見られた。また, 図1より次元数は最後まで比較的大きく変更されている。bedroom と CelebA においては次元数が局所解へ初期収束し, 初期値からの変更が小さかったため改善ができなかったと考えられる。

表1. FID の比較

	WGAN-gp	提案手法	最終次元数
tower	133.7294	129.6151	71(-29)
bedroom	214.8633	221.0722	94(-6)
CelebA	111.9954	129.8148	90(-10)

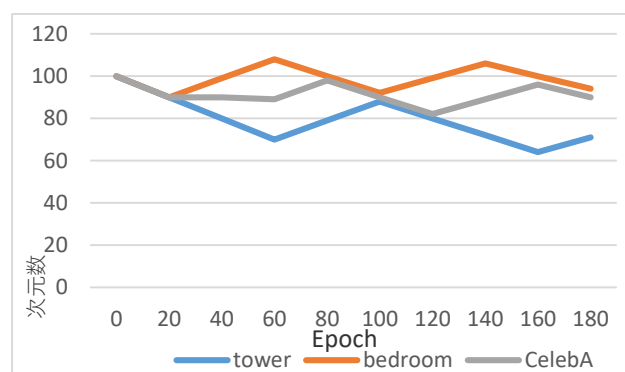


図1. 各学習データにおける次元数変更の様子

5. まとめ

本稿では GAN の学習画像が少ない状況下でも生成を可能とする GAN を提案した。今後は複数のオブジェクトが存在する学習データへの適用, より高解像度の画像を対象とした場合での生成, 生成画像のさらなる品質向上があげられる。

参考文献

- [1] Ian J. Goodfellow, et al. Generative Adversarial Networks. arXiv:1406.2661. 2014.
- [2] Ishaan Gulrajani, et al. Improved Training of Wasserstein GANs. arXiv:1704.00028. 2017.
- [3] Fisher Yu, et al. LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop. arXiv:1506.03365 [cs.CV], 10 Jun 2015
- [4] Ziwei Liu, et al. Deep learning face attributes in the wild. In Proceedings of International Conference on Computer Vision (ICCV), December 2015.
- [5] Martin Heusel, et al. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. arXiv:1706.08500. 2017.