

論文リスト

学籍番号 20C1119 森田大雅

2023 年 3 月 10 日

1 DCGAN

<https://arxiv.org/abs/1511.06434>

2016 年と前のものだが当時は畳み込みを使った教師あり学習が幅広く使われているのに対し、畳み込みを使った教師なし学習はあまり注目されていなかった。そこで筆者らは、教師ありとなしの中間的なモデルと作り、畳み込みを使った教師なし学習の橋渡しをしようと考えた。本論文の提案は、GAN を学習することで良い画像表現ができるという使い道があるが、特徴抽出器として生成器と識別器の部分を再利用すれば、教師ありのタスクにも使える。そこで次のような実験を行っている。CIFAR-10 というデータセットを用いて、特徴を学習する k-means との比較を行った。単一層で一斉に特徴抽出を行えるよう上手く調整されたもの、複数の層に拡張したものなどと、DCGAN と分類器 LSVM を組み合わせたものを比較した。

本論文では、識別器が出力した特徴がラベル付けはされていないがそれと同じように使うことができるのかを調べている。CIFAR-10 のデータセットを使い、識別の精度を調べる実験を行っている。実験より、有用であることが示された。本論文は GAN の新しいモデルの提案にグルーピングできる。

2 LSGAN

<https://arxiv.org/abs/1611.04076>

論文が書かれた当時の標準的な GAN は識別器 (Discriminator) に活性化関数として数 Sigmoid 関数を、損失関数として交差エントロピーを用いる。しかし損失関数が勾配消失問題という問題を引き起こし、学習が上手く行かない恐れがある。この問題を解消するために本論文の LSGAN が作られた。論文によると標準的な GAN は DCGAN のことを指し、このモデルの D の損失関数を最小二乗誤差に変えること提案をしている。これにより、生成画像のクオリティが高くなることと、学習の安定性が定量的に示されている。以下の 2 つの実験を行っている。

1. いくつかのデータセットで、標準的な GAN(DCGAN と EBGAN) と LSGAN の比較
2. LSGAN と標準的な GAN との安定性を比較

データセットは LSVN という 4 つの場面 (寝室、協会、ダイニングルーム、会議室) の画像のデータセットを用いている。実験 1 は生成画像を可視化したシンプルなのである。実験 2 は 4 つのパターンに分けて安定性を評価している。G の batch Normalization (パッチ正規化) を取り除いた場合と、G と D の両方の batch normalization を取り除いた場合。また、optimizer (最適化?) に Adam と RMSProp をそれぞれ用いた場合である。

* 本論文の文献 [12] によるとモデルの性能には optimizer の選択が重要になっているため、文献 [34] の Adam と論文 [35] の RMSProp という 2 つの optimizer を用いて安定性を評価している。(おそらく文献 [12] で [34] と [35] が用いられたからそれに同じことをしてモデルの性能を評価していると思われる)

実験 1 では 4 つの場面 (寝室、協会、ダイニングルーム、会議室) のデータセットを用いた。実験結果から LSGAN が他の GAN よりも生成画像の質が良かった。実験 2 では安定性を評価するために、2 つの optimizer と G、D の batch normalization を取り除いた 4 つの場合分け実験を行った。*BNgd は G、D の Batch Normalization を取り除いたことを意味する。

1. BNg wuth Adam
2. BNg with RMSProp
3. BNgd wuth Adam
4. BNgd with RMSProp

実験結果から全体的に LSGAN の方が画像の質が良く、また Adam よりも RMSProp が、そして RMSProp の中でも BNgd より BNg が他のよりも安定して学習できることが示された。よって LSGAN の BNg with RMSProp が最も安定性があると示されている。本論文は GAN のモデルの改良にグルーピングできる。

3 On the Evaluation of Generative Adversarial Networks By Discriminative Models

<https://arxiv.org/pdf/2010.03549.pdf>

GANs によって画像などをよりリアルに生成できるようになったが、殆どが可視化することによって質を評価している。しかし、画像の分野を超えると視覚的に評価するのが難しい。そこで定量的に評価できる仕組みが必要であり、それによって画像の分野でも異なる GANs のモデルの比較などに使用することができる。よく使用される評価指標に IS(Inception Score) と FID Score(Frechet Inception Distance Score) があるが、どちらも決定的な欠点がある。IS は ImageNet のデータセットを使った学習済みモデルである。非効率的なものと正しく機能するように改良しなければ評価を行えない。FID は IS の欠点の大部分を改善したものだが、画像の分野でしか評価が行えない。

この論文では生成器の部分に WGANs (WGANs は DCGAN と同じ構造) を、識別器の部分に Siamese nueral networks(SNNs) を用いている。生成器の部分は適宜変更しているが、識別器に SNNs を用いるモデルを今回提案している。

画像の分野、GANs のモデル評価、画像以外の分野に適用ができるように実験を行う。

1. quality(人間の感覚に合うように質) を評価

2. 本モデルが頑強であることを示すために GANs で発生しやすい問題 (今回は以下の 3 つ) が起こっているかを調査
 - (a) mode dropping(モード崩壊)
 - (b) mode invention(訳不明)
 - (c) intra-mode collapse
3. いくつかの GANs のモデルを使用し、FID と SDS の比較
4. 画像以外の分野でも適用

つまり 1 と 2 で画像の分野、3 で GANs のモデル評価、4 で画像以外の分野に適用である。4 つ目だが、今回は医療分野 (脳の活動) のデータを使用している。生成器の部分に以下の 3 つ

1. VAE(Variational Autoencoder)
2. medGAN
3. corGAN

を、識別器に

1. SDS(Siamese distance score)
2. MMD(Maximum Mean Discrepancy)

を用いて比較を行っている。結果として生成器は VAE、識別器は SDS を用いた場合が最も良い点数だった。結論として、1 と 2 から IS より、3 と 4 から FID よりも優位であることが示されている。本論文は GAN の新しい評価指標の提案にグルーピングできる。

4 Geometry Score

<https://arxiv.org/abs/1802.02664>

GAN の研究でよく取り上げられる問題が生成画像の質を定量的に評価することと、どんなモデルでも適切にモード崩壊を検知することである。本論文では生成データの geometrical properties(幾何的性質) を計算し、実際のデータと比較することで、GAN の性能を測る。manifold hypothesis(多様体仮説) という幾何学の分野を用いている。なぜこの仮説を用いるのかは直感的と書かれている。データ分布が多様体に対応しているように見えるようなことが書かれている気がするが、私は上手く読み込むことができず分からなかった。実験結果とこの直感がよく一致しており、得られて結果から様々なモデルでの比較に使用することができるのと。本論文は GAN の新しい評価指標の提案にグルーピングできる。

5 APA

<https://arxiv.org/abs/2111.06849>

最新の GAN は大量のデータを必要とし高画質の画像を生成するが、中にはプライバシーや著作権などの問題が発生する可能性のあるデータもある。そのときに必ずしも大量のデータを揃えられるとは限らない。そこで

データに制限があってもそれなりの画像を生成できるこのモデルが提案された。近年の研究からデータが量、質のどちらかだとしても不十分の場合、識別器が過学習を起こすことが分かっている。そこでこのモデルには Data Augmentation と呼ばれるデータの増強に加えて、生成器が識別器の過学習を防ぎ力を持つように工夫がされている。

生成過程は以下のようになっている

1. サンプル画像を生成器へ入力し、偽画像が生成される
2. 生成した偽画像も時々、確率 p で真の画像として識別器に与える
3. 識別器が予測する過学習の度合い λ に応じて、確率 p を上下げする

実験は GAN でよく用いられる IS (Inception Score) と FID (Frechet Inception Distance) と呼ばれる 2 つの評価指標を用いて評価を行っており、実験結果から StyleGAN2 よりも良かったことが示されている。

6 StyleGAN3

<https://nvlabs.github.io/stylegan3/>

途中

7 Improved Techniques for Training GANs

GAN の学習はナッシュ均衡を見つけること一つの要因である。しかし、ナッシュ均衡は特別な場合のみアルゴリズムが存在する。連続型の変数で、次元数も高いためなかなか収束しない。一般的にはナッシュ均衡を見つけるよりも、勾配降下法を用いて価値関数の小さな値を見つけるようモデルが作られている。ナッシュ均衡を求めることに専念したら、ずっと探し続ける可能性があるので収束できるように工夫する。それが「Feature matching」と「Minibatch discrimination」の二つである。

途中