

小規模データに基づく GAN トレーニングに関する研究

Research on GAN training based on small data

魏 正航

指導教員 亀田 弘之

東京工科大学 コンピュータサイエンス学部 亀田研究室

キーワード：敵対的生成ネットワーク, データ拡張, 機械学習

1. はじめに

時代とともに技術はめざましい成長をしている。現時点、様々な分野で作業の自動化が進んでおり、将来的には人間の仕事のほとんどが機械に置き換わると言われている。そのなかで重要となる技術は深層学習(Deep Learning)である。深層学習は画像、音声、自然言語などの広い分野の機械学習タスクにおいて、人間の認識レベル以上の精度を達成し、様々な分野において大きな成功を取って、従来の技術では画像認識などの解決が難しかった問題を容易に解決できるようになるため非常に注目されている。

深層学習の中で中心となる技術が敵対的生成ネットワーク(GAN)である。GAN は特に画像生成の分野で高い性能を発揮する。今日までの研究により非常に多くの GAN のアーキテクチャが提案されており、鮮明な画像や多様性のある画像を生成することが可能になった。特に、GAN によって、新たに特徴を持ったデータが作り出せるので、深層学習を補うことができる。千葉大学の研究によると、GAN を導入したら畳み込みニューラルネットワークは、正解率が上昇することを確認した。また GAN は画像生成から始まったものであるが、自然言語処理や音声処理など様々な分野にも応用されている。

高品質な GAN をトレーニングするには、通常は 5 万枚から 10 万枚のトレーニング用の画像が必要となります。しかし、多くの場合、研究者が数万枚または数十万枚の学習サンプル画像を自由に使え

ることはほとんどできない。さらに、著作権やプライバシー問題を引き起こす恐れもある。そのため、新しいアルゴリズムを開発し、小規模なデータで高品質な GAN をトレーニングすることができれば、GAN をより広く使用できるようにすることができる。共に、深層学習の発展を促すこともできる。

そこで、本研究では必要な学習データを減らすために、小規模データで GAN をトレーニングすることを旨とする。本研究の目的は以下となる。

- データ拡張を用いた GAN を提案し、通常の GAN よりも生成データの品質が向上し、モデルの精度を改善すること。
- データ拡張を導入された各種の GAN の性能を比較検討する。

2. 画像分野におけるデータ拡張

画像分野で十分な数の学習データが集められない場合、様々な手法でデータ拡張を行う。データを水増しのために頻りに用いられている方法は主に二種類がある。

幾何変形は画像分野におけるデータ拡張は訓練画像に対し特定の変換を適用することでデータを増量し学習データ空間を補間する手法であり、画像認識の精度向上や学習の正則化のために長い間用いられてきた。これまでに画像データセットに対する数多くのデータ拡張手法が提案されており、CIFAR-10 や ImageNet においては random crop, horizontal flip, rotate, translate などの変形操作および color shift, whitening などの色変換などがベースライン手法として用いられてきた。

転移学習はもう一つよく用いられた手法であり。画像分野と自然言語処理分野でデータ拡張の手法として小規模な学習データで学習を行う前に、Common Object in Context(COCO) や ImageNet などの大規模なデータセットで学習を行う。学習済みの検出モデルの重みを初期値として学習を始めることで、小規模な学習データでも高い精度を出すことが可能になる。

3. 本研究の概要

ここでは、提案手法であるデータ拡張手法の全体を説明し、それぞれに関して紹介する。

3.1 拡張セットの構築

画像認識の学習において、最も多く利用されているのが、MNIST データセットである。MNIST データセットは、0 から 9 までの数字画像の集合であり、28×28 画素の白黒画像となっている。また、訓練画像 60,000 枚、テスト画像 10,000 枚から構成されている。

本研究は MNIST に対して、幾何変形の従来手法の 16 種類の中で Cutout と Translate X/Y を選んだ。また、MNIST の特性を考慮した、以下 2 つの拡張手法を加えた。

- Random Erasing：ランダムに一部を消すフィルタ
- BalloonAdd：ランダムに吹き出しを添付するフィルタ

MNIST セットをランダムに 10,000 枚を取り出し以上の 4 つ手法で拡張を行い、50,000 枚の訓練画像拡張セットを構築した。

3.2 拡張セットを用いた GAN のトレーニング

現在では様々な GAN のアーキテクチャが提案されており、代表的なものとして画像生成では DCGAN、StyleGAN、SinGAN などがある。本研究では、各種の GAN に対して、データ拡張を導入し、小規模なデータで GAN トレーニングの性能を比較検討する。従来のデータ拡張は主に学習データに対して拡張を行う。本研究提案手法は、まず Discriminator に対して、入力されたデータを拡張を行う。また、Generator が生成する画像をそのままデータとして Discriminator に入力するだけ

でなく、学習途中の Generator が生成するデータも拡張を行う。

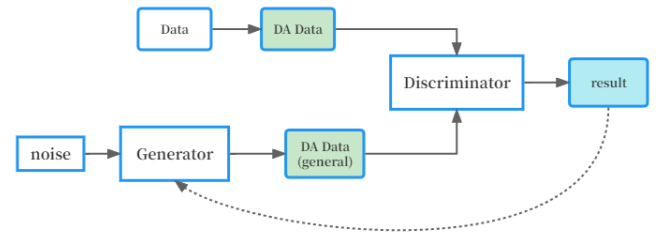


図 1. 処理の概要

実験として 3.1 構築された拡張データを用いて GAN トレーニングを行った。50,000 枚の拡張データを採用し、エポックを 30,000 に設定した。結果は以下のようなになる。

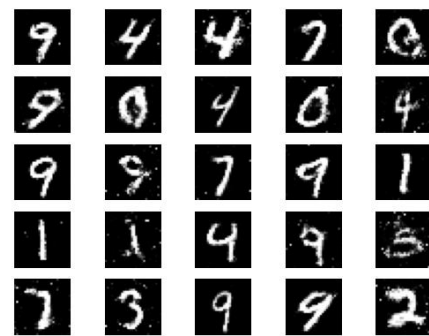


図 2. 処理結果の例

3.3 提案方式を評価

画像分類では、最高の検証精度を提供するモデルチェックポイントを使用し、GAN の場合はできない。今研究は評価指標を 2 つに設定する。一つ目は GAN 生成したデータの品質、GAN に高品質の画像を生成してもらいたい。二つ目は生成した画像の多様性、トレーニングデータセットに固有の画像を生成する必要がある。本研究はフレッシュ開始距離 (FID) 使用して GAN の品質を評価する。

6. おわりに

本研究では、GAN のトレーニングのデータセットの拡充に有効な拡張操作を研究した。限られたデータセットを使った場合の GAN トレーニング問題に対して改善手法を提案した。