

小規模なデータで GAN トレーニングに関する研究

G2121019 魏正航

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 を導入さらたが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 に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. 関連研究

2.1 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繁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や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, horizontalflip, 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。[1]

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 Objectin 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。[2]

2.2 GAN の複数データセット同時学習によるデータ拡張 [3]

山口真弥は少量の学習データで深層学習モデルを学習するための GAN ベースのデータ拡張手法としてドメイン融合手法を提案した。GAN を介して目的モデルの学習対象となる目的データセットと、それとは異なるドメインの外部データセットを混ぜ合わせ、目的データセットと外部データセットを1つのネットワーク上でパラメータ共有しながら学習することでデータを生成させる。それによって生成された拡張データを目的モデルの学習に適用し、通常の学習よりも高い予測精度を得られた。

これらの関連研究のように学習データに対するデータ拡張手法は効果的な正則化手法として証明したが、実際に使う時まだ大量のデータが必要。

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

違う学習データ内容に対して、拡張データの効果を研究するため。本研究は画像認識用データセットについて、従来から提供されている MNIST データセット、CIFAR-10 データセット二種類を採用した。

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。

CIFAR-10 データセットは、カラー画像認識の学習において、よく利用されている。CIFAR-10 データセットは、airplane、automobile といった、10 種類の画像の集合である。32×32 画素の RGB 画像となっており、訓練画像 50,000 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：ランダムに一部をマスクするフィルタ
- 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。

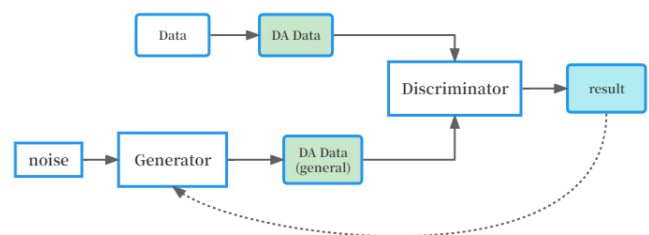


図1. 処理の概要

また、様々な GAN のアーキテクチャが提案されており、アルゴリズムに合わせてデータ拡張導入時点変える。

拡張を調査するために、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。結果は以下のようになる。

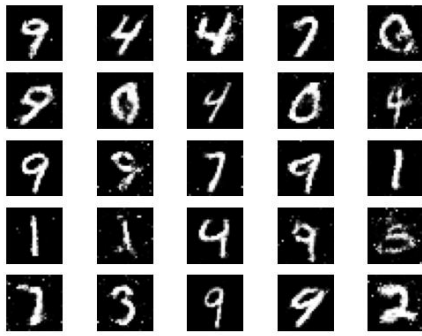


図 2. 処理結果の例

3.3 提案方式を評価

提案方式の評価は次のように行う。各種の GAN に対して、MNIST データセットと CIFAR-10 データセット中で学習データを取り出し、そのデータに対して以下の通りについてデータ拡張の効果と比較する。

1. データ拡張を行わない
2. 学習データ拡張を行う
3. Generator 生成したデータも拡張を行う

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。本研究はフレイッシュ開始距離 (FID) 使用して GAN の品質を評価する。

4. 研究計画

表 1. 研究のスケジュール

	8月	9月	10月	11月	12月	1月	2月
拡張データ作成							
GAN トレーニング							
評価実験							
論文の完成							
学外発表							

5. 進捗状況

- MNIST データセットの拡張セットを完成した。
- 拡張セットを用いた典型的な GAN のトレーニングを完成した、今後は DCGAN、StyleGAN などのトレーニングを進む予定

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。

参考文献

- [1]水野充大, 岡留剛, 系列データにおける特徴空間上での疑似訓練データ生成, 情報処理学会講演論文集, 1884-197X, 2020-09-11, 6p
- [2]三本拓未, 四宮友貴, 果実位置検出における半教師あり深層学習とデータ拡張の適用, 日本知能情報フェジィ学会, 18820212, 2021-37-0, p241-244
- [3]山口真弥, 敵対的生成ネットワークの複数データセット同時学習によるデータ拡張, DEIMForum 2019, P1-114
- [4]山下明博, AI 研究のための手書き動物イラストデータセットの開発, Journal of Yasuda Women's University, 0289-6494, 2020-02-28, 48, p247-256
- [5]河野曜平, GAN を用いたデータ拡張, IPSJ SIG Technical Report, Vol.2017-CVIM-207, 14
- [6]Masahito Hoshi, Kei Kogai, Hisayuki Horai, Comparison of GANs for Data Augmentation Performance in Classification, 2020 年情報科学技術フォーラム (FIT), 3, p149-150
- [7]菅野, 怜, 鍵和田, 聡, 2 段階の画像生成を活用した偏りのあるデータセットに対する実践的な data augmentation, 第 82 回全国大会講演論文集, 2020-02-20, 1, p101-302
- [8]吉田航, 寺内光, 森直樹, 岡田真, TDGA AutoAugment を用いた漫画データセットにおけるデータ拡張手法の提案, 人工知能学会全国大会論文集, JSAI2021, 0, 3D2OS12b02-3D2OS12b02
- [9]Philip T. Jackson, Amir Atapour-Abarghouei, Stephen Bonner, Toby Breckon, Boguslaw Obara: Style Augmentation: Data Augmentation via Style Randomization, 2018, CVPR Workshop, 2019
- [10]Connor Shorten, Taghi M. Khoshgoufar: survey on Image Data Augmentation for Deep Learning, Journal of Big Data, 2019