

## 画像分類における各種 GAN の教師データ拡張性能の比較 Comparison of GANs for Data Augmentation Performance in Classification

星 将仁<sup>†</sup>      小飼 敬<sup>†</sup>      蓬萊 尚幸<sup>‡</sup>  
Masahito Hoshi      Kei Kogai      Hisayuki Horai

**Abstract:** 敵対的生成ネットワーク(GAN)は、画像を生成するための一般的な手法の一つである。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。そこで、画像に対する学習において教師データを増やすために GAN が広く用いられている。本研究では、画像分類を例にとり、各種の GAN の性能を比較検討する。具体的には、画像分類器を固定し、各 GAN が生成する各世代の画像データを入力としたときの分類精度を測定する。

**Keywords :** GAN, 画像分類, データ拡張

### 1. まえがき

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)とよばれる技術である。深層学習は機械学習の技術の一つであり、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。深層学習の中でも中心となる技術が畳み込みニューラルネットワーク(CNN)である。CNN は特に画像認識の分野で高い性能を発揮する。CNN は教師データを元に学習を行う教師あり学習である。しかしその教師データに偏りがある、もしくはデータ数自体が少数である場合には正しく学習することができず認識の精度が低下してしまう問題がある。

本研究では認識の精度を向上させるために、敵対的生成ネットワーク(GAN)を用いて教師データと類似の画像を生成し教師データを増やす事を目指す。

### 2. 従来技術

#### 2.1 データ拡張

教師データを増やす方法としてデータ拡張という手法が存在する。医療画像分析など、データへのアクセスが難しいケースにおいて有効な手法である。オリジナルの画像に対して、画像の反転、切り抜き、回転などの Geometric Transformation、ピクセル値の操作や Gray Scale 変換などの Color Space Transformation、画像の先鋭化やぼかしなどの Kernel Filters、SamplePairing や RICAP などの Mixing Images、画像の一部をランダムに消す Random Erasing といった入力画像に対し処理を施す手法や、Icing on the Cake[1]や特徴量空間におけるデータ拡張などの Feature Space Augmentation、DeepFool や One Pixel

Attack などの Adversarial Training、生成モデルである GAN を用いた GAN-based Data Augmentation、Style Augmentation[2]などの Neural Style Transfer といった深層学習を用いたデータ拡張の手法がある[3]

データ拡張には、特徴点にあまり変わりのないデータを多く学習したために発生する過学習や、実際にはありえないデータなど品質の悪いデータで学習を行うことによって反対に精度が低下してしまうといった問題が存在する。

#### 2.2 敵対的生成ネットワーク(GAN)

画像を生成する手法として GAN[4]が存在する。GAN は Generator と Discriminator の 2 つのネットワークで構成され、それらを交互に競合させながら学習を進める(Fig. 1)。また学習には大量のデータを必要としないため、GAN の学習を容易に行うことができる。GAN を使用することで入手することが困難なデータや、実際のデータに類似したデータを生成することができる。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 などがあり、I2I 変換(Image-to-Image Translation)では CycleGAN、StarGAN、FUNIT などがある。また GAN は画像生成のタスク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タスクに応用され、生成データの種類も多様化している。

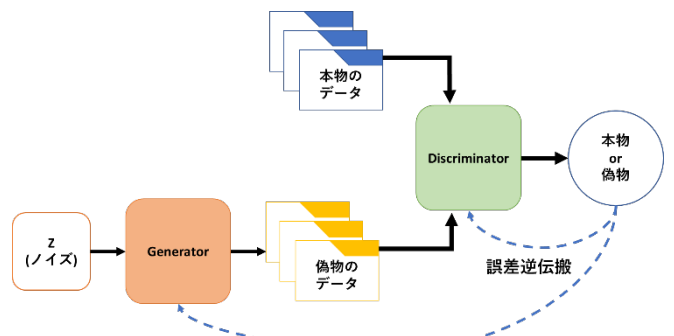


Fig.1 GAN's conceptual diagram

### 3. 本研究の狙い

本研究では従来行われていたデータ拡張の部分を

<sup>†</sup> 茨城工業高等専門学校

National Institute of Technology, Ibaraki College

<sup>‡</sup> 国立高等専門学校機構本部

National Institute of Technology, Headquarters

GAN による教師データ拡張に置き換えて CNN の学習を行う手法を提案する (Fig.2)。単に学習を終えた Generator が生成する画像を教師データとして使用するだけでなく、学習途中の Generator が生成する画像を教師データとした際の画像分類の精度も併せて測定する。データ拡張では評価対象を意識しながらデータの増しを行う必要があったが、GAN であれば評価対象を意識する必要もなく教師データと類似のデータを大量に生成することが可能である。また様々な GAN のアーキテクチャが提案されており、生成できるデータの種類も音楽や動画、3D オブジェクトなど多様化している。

本研究で GAN によって生成された画像が教師データとして有効であることが立証されれば、画像以外の分野に対しても GAN を用いてデータ拡張を行うことができるようになると思われる。そこで GAN を用いることで従来のデータ拡張で生じる問題を回避しつつ、CNN の高い画像認識精度を得られるかを検証することが本研究の目的である。

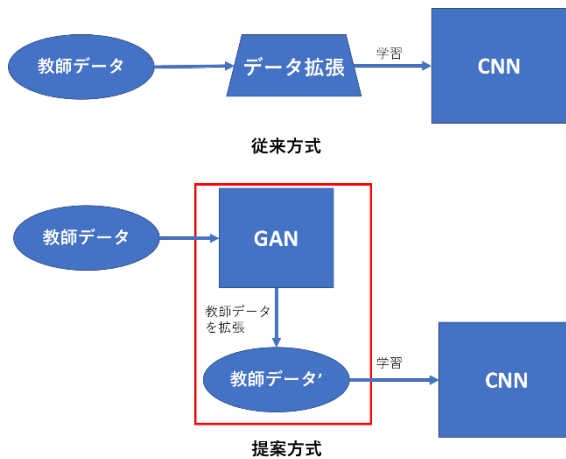


Fig.2 Proposed method

#### 4. 今後の予定

本研究の提案方式の評価は次のように行う。手書き文字データセット (MNIST) を用いて実験を行う。

MNIST データセットの中からランダムに取り出した画像セットを教師データとする。その教師データに対して以下の 4 通りについて認識精度を比較する。

1. データ拡張を行わない
2. 従来のデータ拡張を行う
3. 提案方式でデータ拡張を行う
4. GAN で生成された画像のみ用いる

CNN が評価する対象のデータはすべて同一のデータで行う。実験 3・4 では conditional GAN (cGAN) [5] を用いて画像生成を行う。cGAN はクラスラベルなどの情報を追加することで、生成データを制御できるようにした GAN である。Fig4 は cGAN による MNIST 画像の生成結果である。cGAN を用いることで人の手で生成データに対してラベル付けをする作業をなくすることができる。MNIST データセットについて提案方式が有効であることが確認したのち、一般物体認識においても同様に評価を行う。使用するデータセットは CIFAR-10 及び CIFAR-100 データセットを使用する。MNIST データセットは色情報がグ

レースケールの 1 チャンネルのみであるが、一般物体はカラー画像であり、さらに形状も複雑であるため情報量が増加し MNIST 画像生成よりも難しいと考えられる。同様の評価を CycleGAN [6] を用いて行う。cGAN は潜在空間から本物に近い画像を生成するように学習するが、CycleGAN は画像 A から画像 B への変換を学習する。提案方式について画像生成、I2I 変換の教師データ生成方法の違いによって CNN の認識精度に差が生じるのかを検証する。さらに、本研究の提案方式の適用分野拡大のために画像以外のデータについて有効であるかの検討を行う。

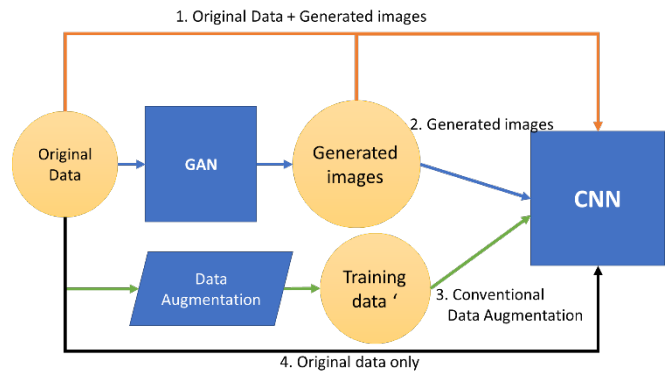


Fig.3 Method of verification



Fig.4 Generated images by conditional GAN

#### 6. 参考文献

- [1] Tomohiko Konno, Michiaki Iwazume: Icing on the Cake: An Easy and Quick Post-Learning Method You Can Try After Deep Learning, 2018, arXiv:1807.06540 [cs.LG], 2018
- [2] Philip T. Jackson, Amir Atapour-Abarghouei, Stephen Bonner, Toby Breckon, Boguslaw Obara: Style Augmentation: Data Augmentation via Style Randomization, 2018, CVPR Workshop, 2019
- [3] Connor Shorten, Taghi M. Khoshgoufar.: survey on Image Data Augmentation for Deep Learning, Journal of Big Data, 2019
- [4] Goodfellow, I. J., Pouget-Abadie, J., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative adversarial nets, NIPS, 2014
- [5] Mehdi Mirza, Simon Osindero: Conditional Generative Adversarial Nets, 2014, arXiv:1411.1784 [cs.LG], 2014
- [6] Zhu, J., Park, T., Isola, P., Efros, A. Alexei.: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, ICCV, 2017