

# 超解像を用いた画像生成の高速化

東京都立小石川中等教育学校 高校2年 小森陽向

## ■研究内容

GAN(敵対的生成ネットワーク)を用いた画像生成は近年急激に身近なものとなっている。しかし個人で画像生成を行う場合、その学習と生成には長時間を要する。本研究では超解像(画像を高画質にするGANの一種)を利用し、学習の高速化を行った。

## ■研究方法

研究用の画像として広く利用されているMNIST(手書き数字の28x28画像データ)を利用した。

方法1: 28x28(元のサイズ)、14x14、10x10のピクセル数で画像生成を行い、学習時間を測定した。

方法2: 画質を落とした14x14と10x10の画像を超解像を用いてそれぞれ画素数を上げた。

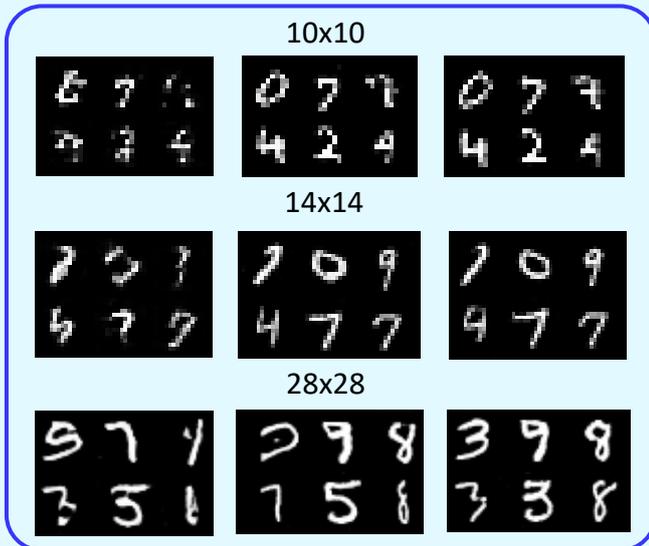


今回の実行環境はGoogle Colabで、PythonのPyTorchでDCGANモデルをそれぞれ作成した。また超解像には自作モデルではなくだれでも利用できるという観点から学習済みモデルのSwinIRを用いた。

## ■研究結果1

詳細な条件は以下に示す

- バッチ数は固定し、エポック数は100回とした。
  - ネットワークの層数は固定するのが難しかったため、10x10と14x14が7層、28x28が10層とした。
  - 損失関数はバイナリクロスエントロピーを用いた。
- 以下がそれぞれの生成画像である。



左から1エポック、50エポック、100エポック終了後の画像。10x10の画像は画質が低すぎて数字を判別しづらい。

※「エポック」とは学習回数の単位である。

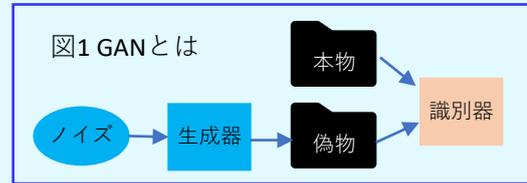
## ■参考文献

『GANディープラーニング実装ハンドブック』秀和システム

『【超解像】SwinIRを使って画像を高解像度化してみる』 <https://tt-tsukumochi.com/archives/4716>

## ■GANとは何か

GANは画像を生成するためのニューラルネットワークである。



生成器が偽物を作り出し、それを識別器が'偽物'と判別することにより、'偽物'と判別されづらい画像を生成することが出来る。

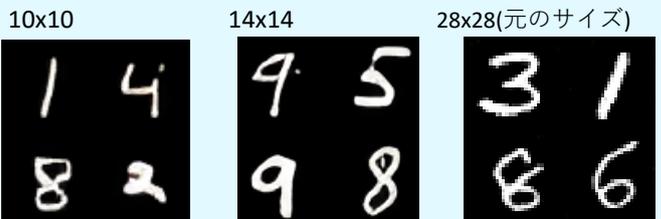
## ■研究結果2

それぞれのデータセットでかかった時間は以下の通りである。

画質	画素数	層数	実行時間
10x10	100	7層	2470s
14x14	196	7層	2592s
28x28	784	10層	3462s

この表からも分かるように、約4倍のピクセル数だと1.35倍の学習時間となっているものの、層数が同じ10x10と14x14ではそこまで大きな差は見られなかった。

また超解像を実行したのが以下の画像である。



10x10は輪郭がいびつな形になってしまったが、14x14は元画像と比較しても遜色無い結果になった。

## ■考察・展望

およそ4分の1程度の画素数の画像であれば、超解像を用いることで、本来の画像とほぼ同じ精度の画像が生成できた。一般的にはGANのモデルにおいては、精度を良くするためには画素数を上げることが重視される。しかし画素数を落としたとしても、超解像を効果的に利用することによって、精度の高い画像が生成できた。これによって個人的なPCでも短時間で画像の生成を高品質に行うことができる。またMNIST以外のカラー画像データでも同じように学習時間を短縮できるのかを研究してみたい。