

SRGAN を用いた超解像データによる画像分類精度の向上

永田研究室 F116022 北野 親良

1. 目的

ニューラルネットワークは現在の機械学習プロセスの主流であり、画像・音声の認識や株価の予測など様々な場面で利用されている。ニューラルネットワークによる画像のカテゴリ分類問題の精度は学習時に使用するデータセットの画質が鮮明である方が望ましいとされている。そこで近年高い注目を集めている Generative Adversarial Network (GAN)を画像の解像度向上に応用した Super-Resolution Using a GAN (SRGAN) を用いて、低解像度のデータセットを高解像に変換（超解像処理）し、その画像データがニューラルネットワークの分類精度に及ぼす影響を検証する。

2. 研究内容

初めにディープラーニング用フレームワークの Keras を用いて SRGAN とその改良版モデルである Enhanced SRGAN(ESRGAN)を Python で実装した。SRGAN の学習では高解像度の画像データセット(HR 画像)とそれを圧縮した低解像度データ(LR 画像)をペアで使用する。SRGAN は LR 画像に超解像処理を施し高解像度(SR 画像)に変換し生成する Generator と、SR 画像か HR 画像かを見破る Discriminator の二つのネットワークから構成され、これらが互いに敵対することで学習を進める。これにより Generator は単体で学習する場合に比べより高いクオリティの超解像を生成できるようになる。今回 SRGAN の学習用データにはスタンフォード大学が公開している `sd10` というデータセットを使用した。このようなプロセスを経て学習した Generator を用いてベンチマーク用低解像度データセットの CIFAR10(飛行機, 車, 犬, 猫などから成る 10 カテゴリの画像データ)を高解像度化した。次に画像のカテゴリ分類を行うシンプルな構造の畳み込みニューラルネットワーク(CNN)を構築した。この CNN は CIFAR10 に含まれる訓練データを SRGAN の Generator で超解像化した画像, ESRGAN の Generator で超解像化した画像, リサイズツールで拡大しただけの画像, オリジナル(未加工)の画像で学習させ、それぞれ CNN1, CNN2, CNN3, CNN4 として保存した。最後に CIFAR10 に含まれるテストデータを用いて、それぞれの学習済み CNN による分類精度(識別率)を比較した。

3. 結果

CNN1, CNN2, CNN3, CNN4 での平均識別率はそれぞれ 77.41%, 78.60%, 76.42%, 75.27%となった。このことから ESRGAN の Generator で生成した画像によって学習した CNN2 が最も高い正解率となり、次点で SRGAN の Generator を使用した CNN1 となった。超解像処理によって画像が持つ情報がより鮮明になったことや、消失していた特徴がうまく復元されたことが識別の良い判断材料となったと考えられ、超解像処理が画像分類問題の前処理として有効に作用することが確認できた。

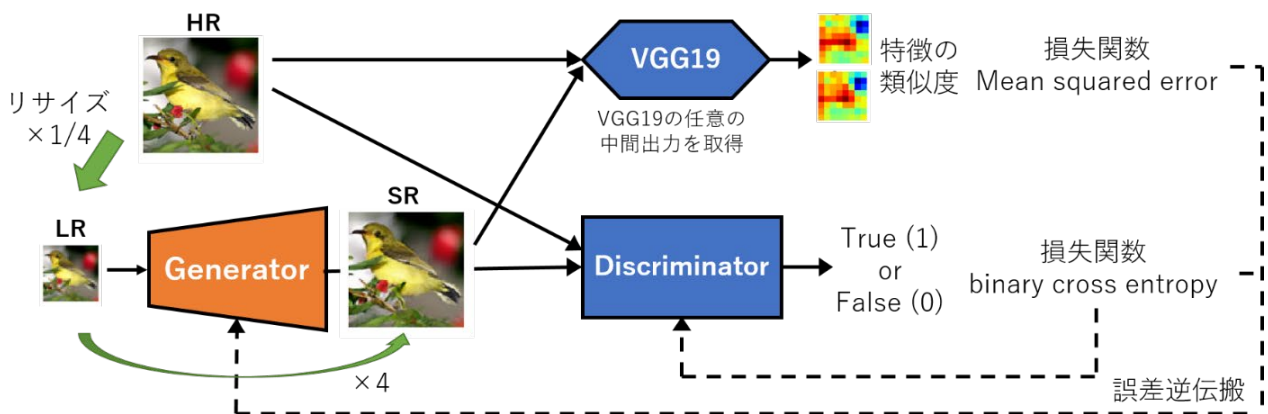


Fig. 1 Data flow in Super-Resolution GAN Architecture.