

画像認識の学習用画像生成

玉川学園高等部 1年 原田 悠世

研究動機

私は昨年、ロボカップという大会に出場したのをきっかけに、機械学習を用いた文字認識の研究を行ってきた。

昨年は学習するための**画像データ不足の問題で十分な性能が発揮できなかった**と考えられた。そこで、データを増やすデータ拡張を行おうと考えた。

一般的には学習データを増やすにはデータ拡張という手法を用いることが多い。これは画像を加工して似た画像を増やすものである。

しかし、データ拡張では加工により文字の特徴を損なう可能性がある。そこで、**画像を生成できるGANを用いてデータ拡張を行う**ことにした。

GANとは

GANとは、機械学習の手法の一つで、**生成器と識別機が交互に学習を行い学習する**学習手法である。

識別機は生成器が作り出した画像と本物の画像を分類し、生成器は識別機を騙せられる画像を生成する。

学習した生成器で推論を行うことによってノイズという画像の元から**偽物の画像を生成することができる**。

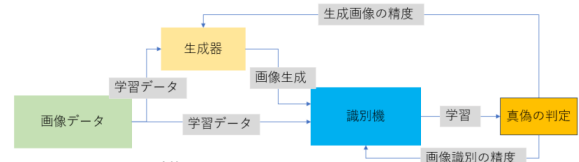


図1 GANの構造

実験方法

今回の実験方法は右図の工程で実験を行った。

今回のモデルの学習は**0~9手書き数字**の画像群から枚数を決めて学習用データとして使用した。また、**今回使用したGANは全て同じで、900枚の画像で作った**。

各モデルの評価については0~9までの数字ごとの**分類の正答率と正答率の平均**から評価を行った。

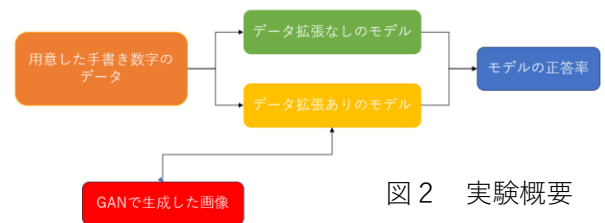


図2 実験概要

実験1 生成画像の検証

(目的)

GANで生成された画像がデータセットとして有用かを調べるため、GANで生成された**画像の割合が多い場合でも学習に悪影響を及ぼさないか**を調べる。

(条件)

・100枚の画像でデータ拡張なしのモデルを作成し、同じ100枚の画像と100枚の生成画像でデータ拡張ありのモデルを作成する。

(結果・考察)

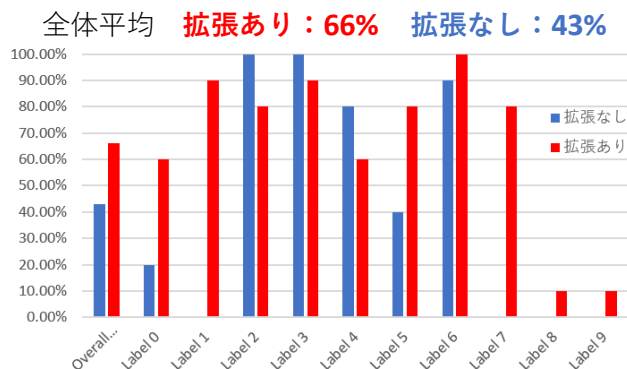


図3 実験1の正答率

図は縦軸が正答率、横軸は左から正答率の平均と各数字の正答率を表している。

結果として**一部を除いて正答率が全体的に上がった**。下がった部分は**過学習の抑制の結果だ**と考えられる。

実験2 データ拡張の検証

(目的)

全てのモデルの学習用画像を統一することでデータ拡張として有用かどうかを調べる。

(条件)

・GANも含め、データ拡張分の画像を除き全て同じ900枚から学習を行う。拡張分は100枚とする。

(結果・考察)

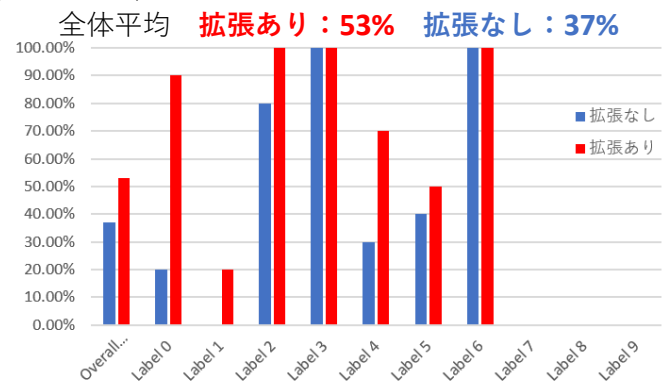


図4 実験2の正答率

図は図3と同じ方法で結果を表している。

結果としてデータ拡張を行ったほうが**正答率が高くなった**。これは手書き数字とGANの相性が良かったからだと考えられる。また、実験1と比べ、正答率が落ちたのは、手書き数字の画像なので、**独特な数字の画像が増えた為、精度が下がった**と考えられる。

考察・結論

実験1の結果より、データ拡張ありの方の全体的な正答率が高かったことから**GANによる生成画像は有効であると考えられる**。また、実験2の結果から**データ拡張を行ったモデルのほうが全体的な正答率の向上が起こった**ことから、**GANによるデータ拡張は有効だといえる**。

参考文献

EMNIST: an extension of MNIST to handwritten letters Cohen, G., Afshar, S., Tapson, J., & van Schaik, A. Wed, 1 Mar 2017, PyTorch <https://pytorch.org/>, TensorFlow <https://www.tensorflow.org/>, Generative Adversarial Networks Ian J. Goodfellow Tue, 10 Jun 2014 <https://arxiv.org/abs/1406.2661>