

生成画像品質を考慮した CapsGAN によるデータ拡張

Data Augmentation Using CapsGAN

Considering Generated Image Quality

松田 卓也[†] 丸寄 佳奈子[‡] 渡辺 裕[‡]
Takuya MATSUDA[†] Kanako MARUSAKI[‡] and Hiroshi WATANABE[‡]

[†] 早稲田大学大学院 基幹理工学研究科
情報理工・情報通信専攻

[‡] 早稲田大学 基幹理工学部 情報通信学科

[†] Communication and Computer
Engineering, Graduate School of
Fundamental Science and Engineering,
Waseda University

[‡] Department of Communications and
Computer Engineering, Waseda University

Abstract AI's accuracy depends on data quantity. However, some people may not have a lot. Thus, it is convenient if we can train with a few data by incorporating data augmentation. GAN can be used for data augmentation. CapsGAN employs Capsule Network in the discriminator. We propose data augmentation using CapsGAN.

1. はじめに

近年、様々なシステムに AI が導入されつつあるが、AI の精度を最大限に引き出すにはビッグデータが必要になる。世界的大企業はビッグデータを確保しているために精度の高い AI を創出している。しかし、中小企業や個人研究者などはビッグデータを確保できないために、AI で高い精度を出せないことがある。

我々は画像認識を対象とし、画像データの水増し (data augmentation) に焦点を置く。CapsGAN を用いて従来よりも元データに類似した品質の良い画像を生成し、訓練データセットを水増しする。これにより、元データの量が少ない環境において CNN のクラス分類精度の向上を図る。

2. 関連研究

渡部らは DCGAN[1] を用いて動物の類似画像を生成し、生成画像を品種識別用 CNN の学習データセットに追加することで、その CNN モデルの識別精度が上がることを確認した[2]。生成画像を全て追加した場合、生成画像の中で品質の良い画像だけを追加した場合の、どちらにおいても精度が向上することが確認されている。

3. 提案手法

3.1. CapsGAN

関連研究で扱われていた DCGAN は識別器に CNN の構造を適用している。CNN は一般的に中間層において畳み込み層とプーリング層が複数重ね合わされているが、このプーリング層において空間的な位置情報を

喪失している。対して、Hinton らが提案した Capsule Network[3] はプーリング層を減らし、特徴フィルタを複数枚ごとにグループ化してカプセル単位で学習することで、特徴間の位置関係も学習することができる。

Jaiswal らが提案した CapsGAN[4] では、Capsule Network の構造を GAN の識別器にのみ適用している。それゆえ、生成器が生成した画像の特徴間の位置関係が正しくなければ識別器は偽画像と判断し、特徴間の正しい類似画像のみが識別器から真画像と判断されるようになる。同時に、生成器は真画像に似た画像を作るように学習していく。

本研究で扱う CapsGAN の識別器の構造を以下の図 1 に示す。Digit Capsule 層の後に真偽値を返す。生成器の構造は DCGAN[1] と同じである。

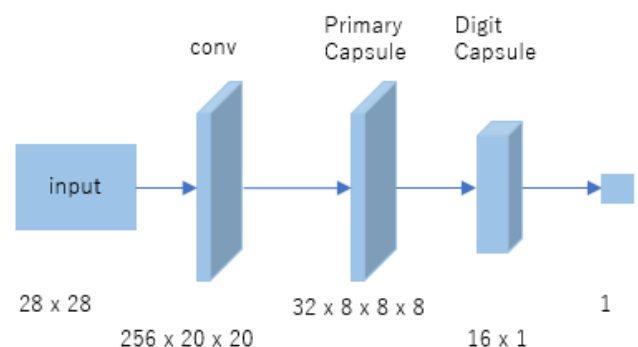


図 1. CapsGAN(識別器)

3.2. 評価

MNIST データを学習済みの CNN モデルで、

CapsGAN を用いて生成した画像の品質を評価する。また、比較対象として DCGAN を用いて生成した画像の品質も評価する。それぞれの GAN が生成した画像を CNN モデルでクラス分類し、正しいクラスに分類される数が多いほど、その GAN が良いモデルであると評価する。評価値は以下の式で表される。

$$\text{評価値(\%)} = \frac{\text{正解分類数}}{\text{クラス数} \times \text{データ数}} \times 100 \quad (1)$$

3.3. データ拡張

本研究では CapsGAN によるデータ拡張を提案する。CapsGAN によって生成された画像は特徴間の位置関係を保持するので、生成画像は構造が正しく品質が良いと推測される。また、CNN は学習データが多いほど精度が高くなるので、学習データセットが少ない環境においてこのデータ拡張が DCGAN によるデータ拡張よりも有用であると考えられる。

4. 実験

4.1. 実験 I

提案手法で述べた GAN のモデルの評価方法に従い、DCGAN と CapsGAN それぞれに対して MNIST データセットを用いて実験を行った。それぞれの生成画像を図 2、図 3 に示し、結果を表 1、表 2 に示す。図 2 の方が背景は白いが、デジタル文字自体は図 3 の方が鮮明である。CapsGAN の方が高い評価値を得た。

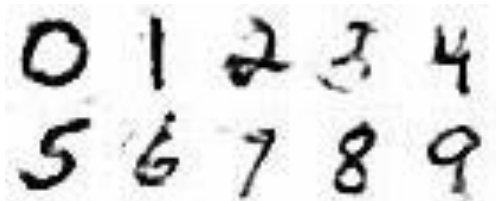


図 2. 生成画像(DCGAN)

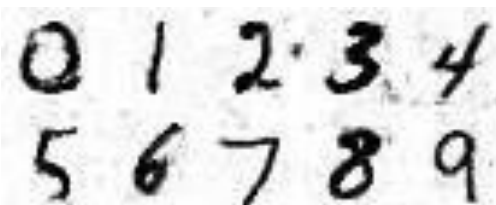


図 3. 生成画像(CapsGAN)

表 1. 実験結果(クラス 0~4 の分類率(%))

	0	1	2	3	4
DCGAN	99.7	99.7	97.3	98.8	96.2
CapsGAN	99.8	99.8	99.6	99.9	99.4

表 2. 実験結果(クラス 5~9 の分類率(%))

	5	6	7	8	9
DCGAN	98.4	99.7	99.5	99.7	99.1
CapsGAN	99.5	99.5	99.5	98.7	99.2

4.2. 実験 II

提案手法に従って CapsGAN と DCGAN それぞれでデータ拡張を行い、データ拡張しない場合と比べて CNN の分類精度の比較を行う。結果を表 3 に示す。

表 3. 実験結果

	データ拡張なし	DCGAN によるデータ拡張	CapsGAN によるデータ拡張
正解率	99.06	98.95	99.17

5. まとめ

本稿では、CapsGAN と DCGAN で画像を生成し、どちらが品質の良い画像を生成しているか評価を行った。さらに、生成した画像を元データに加えて CNN のクラス分類を行った。実験の結果、(1)より CapsGAN の方が DCGAN よりも品質の良い画像を 0.67%多く生成していて、CapsGAN によるデータ拡張を施した CNN のクラス分類が最も高い精度を得た。

文 献

- [1] A.Radford, L.Metz, and S.Chintala: "Unsupervised representation learning with deep convolutional generative adversarial networks", In 4th Int. Conf. on Learning Representations(ICLR), No 67, pp.1-16 (May. 2016)
- [2] 渡部, 渡辺:"生成画像品質を考慮した DCGAN を用いたデータオーギュメンテーション", FIT2017 (第 16 回情報科学技術フォーラム), H-038, pp.273-274, (Sept. 2017)
- [3] Ayush Jaiswal, Wael AbdAlmageed, Yue Wu, Premkumar Natarajan, "CapsuleGAN: Generative Adversarial Capsule Network", <https://arxiv.org/abs/1802.06167>, 2018
- [4] Sara Sabour, Nicholas Frosst, Geoffrey E Hinton, "Dynamic Routing Between Capsules", <https://arxiv.org/abs/1710.09829>, 2017

† 早稲田大学大学院 基幹理工学研究科 情報理工・情報通信専攻 渡辺研究室

〒169-0072 東京都新宿区大久保 3-14-9 早大シルマンホール 401

TEL.03-5286-2509 E-mail: t-matsuda@akane.waseda.jp