

DCGAN を用いたデータオーギュメンテーションによる 猫の品種識別について

On cat breed identification by data augmentation using DCGAN

渡部 宏樹[†]

渡辺 裕[†]

Hiroki Watabe[†]

Hiroshi Watanabe[†]

[†] 早稲田大学大学院 基幹理工学研究科

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

Abstract In recent years, Deep Learning has been focused on not only in image recognition but also in various fields. In this paper, we applied DCGAN to CNN's data augmentation. By the experiment, we confirmed the effectivity of this data augmentation and the recognition success rate reaches to 70.6%.

1. はじめに

スマートフォン等の普及により、多くの人が気軽に写真を撮り、SNS 等に載せていくことで多くの写真がインターネット上に溢れている。代表的な被写体としては、犬や猫等のペットが挙げられる。犬や猫等の動物は、品種間の違いが顔のパーツや模様等の局所的な細かな違いしかないことがある。そのため、ある動物の画像から品種の識別をすることは一般物体認識に比べて難しい。

近年、ディープラーニングが画像認識や音声認識、自然言語処理等の様々な分野で注目されている。特に画像認識分野では、畳み込みニューラルネットワーク(CNN)を用いた手法が高い性能を発揮している[1]。CNN はネットワークの構造やパラメータ、学習画像枚数等がその性能に影響を与える。また、画像の自動生成分野では、Deep Convolutional Generative Adversarial Networks(DCGAN)[2]と呼ばれる手法が高精度の画像生成を可能にしている。この手法は CNN で実装された生成機と判別機を共進化させることで高性能の生成機を得ている。

本研究では、猫の顔画像を対象とし、CNN による品種識別実験を行う。さらに、DCGAN を用いて少数の訓練画像を基に猫画像を生成し、訓練データセットに加えるデータオーギュメンテーションを提案し、それにより CNN の識別率が向上することを示す。

2. ディープラーニング

2.1. 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク(Convolutional Neural Network : CNN)は、画像認識分野においてよく用いられるディープラーニングの一種である。基本的な畳み込みニューラルネットワークでは、畳み込み層とプーリング層と呼ばれる二つの層が交互に複数回繰り返される。畳み込み層とプーリング層の繰り返しが終わると、全結合層と呼ばれる層でユニットの全結合が行われる。さらに、クラス分類が目的の場合は、最後に出力層としてソフトマックス層が用いられる。

2.2. DCGAN

訓練データセットからデータ分布を獲得し、画像を生成する生成機と、入力が訓練データセット由来か生成機由来かを判別する判別機を共進化させることで、画像の生成機を得るシステムを Generative Adversarial Networks (GAN)という[3]。GAN は学習を進めることで、判別機が本物の画像か生成された画像かの判断ができない、つまり本物と見分けがつかない精度の画像を生成することを目的としている。生成機は一様分布などからサンプルされた乱数 z を入力として受け取り、これを種として画像を生成する。

この GAN をバッチ正規化およびプーリング層を畳み込み層に置き換えた全畳み込みネットワークを用いた CNN で実装したものを Deep Convolutional Generative Adversarial Networks(DCGAN)という。

3. DCGAN を用いたデータオーギュメンテーション

本研究では 2.2 節で説明した DCGAN を用いて CNN の訓練画像データセットのデータオーギュメンテーションを行う。DCGAN の生成器は、判別機の CNN で本物かどうかの区別がつかない画像を生成する。よって、その画像は DCGAN の訓練画像に類似した CNN 特徴量を持つと推測される。CNN は、一般的に訓練に使う画像枚数が多いほどその性能は高くなる。そこで、訓練画像枚数が十分でないデータセットにおいて、訓練サンプルを増やすのに有用なデータオーギュメンテーションとなると考えられる。

4. 評価実験

4.1. 実験手法

訓練データセットの画像を基に DCGAN を学習させ、学習した DCGAN を用いて画像を生成する。生成した画像を訓練データセットに加え、画像認識用 CNN を学習し、その正解率を、生成した画像を訓練データセットに加えない従来手法と比較する。なお、生成し、データセットに加える画像枚数を各クラス 50~250 枚で変動させながら実験を行った。

本実験では、データセットとして”The Oxford-IIIT-Pet dataset”[4]内の猫の画像から、顔領域を切り出したものを、1 クラスの訓練画像 50 枚、テスト画像 100 枚、全 12 クラスを使用する。

4.2. 生成された画像

生成された画像の例を以下に示す。

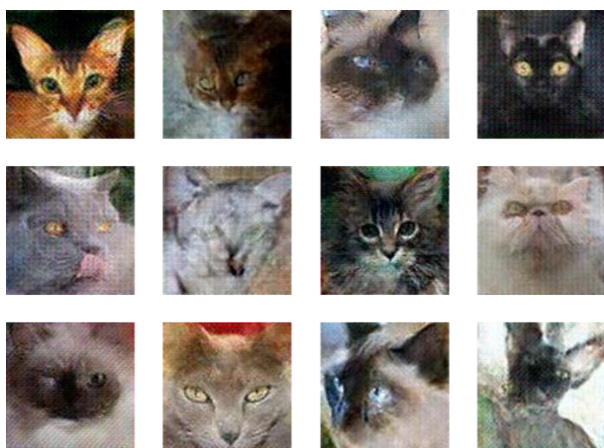


図 1 生成された画像例

4.3. 実験結果

従来手法および提案手法の実験結果を表 1 に示す。従来手

法では 68.9%，提案手法では、各クラスに 200 枚ずつ生成画像を加えた時が最高で 70.6%と 1.7%の正解率の向上が確認できた。

表 1 実験結果

手法	正解率 (%)
従来手法	68.9
提案手法(+50 枚)	68.8
提案手法(+100 枚)	69.2
提案手法(+150 枚)	69.9
提案手法(+200 枚)	70.6
提案手法(+250 枚)	70.3

5. おわりに

本研究では、DCGAN を用いたデータオーギュメンテーションを行い、訓練画像枚数が十分でないデータセットにおけるその有効性の検証を行った。その結果、従来手法の正解率が 68.9%，提案手法の正解率が最大 70.6%と提案手法が有効であることが確認できた。

文献

- [1] 岡谷貴之：“機械学習プロフェッショナルシリーズ 深層学習”，講談社，東京，(2015)
- [2] Radford A, Metz L, Chintala S. : “Unsupervised representation learning with deep convolutional generative adversarial networks” In 4th Int. Conf. on Learning Representations (2016)
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio : “Generative adversarial nets”, In NIPS, p.p. 2672–2680 (2014)
- [4] O. M. Parkhi, A. Vedaldi, A. Zisserman, C. V. Jawahar : “Cats and Dogs”, IEEE Conference on Computer Vision and Pattern Recognition (2012)

† 早稲田大学 基幹理工学部 情報通信学科

〒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: xps169@akane.waseda.jp