

進捗報告

1 今週行ったこと

sklearn の勉強をした。

猫の画像認識の論文「DCGAN を用いたデータオギュメンテーションによる猫の品種識別について」を読んだ [1]。この研究では、猫の顔画像を対象とし、CNN による品種識別実験を行い、DCGAN を用いて少数の訓練画像をもとに猫画像を生成し、訓練データセットに加えるデータオ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。猫の顔画像のデータセットとして” The Oxford-IIIT-Petdataset” [2] ※1 内の猫の画像から 1 クラスの訓練画像 50 枚、テスト画像 100 枚、全 12 クラスを使用した。研究の結果、従来手法（データオギュメンテーションを行わないとき）の正解率が 68.9 %、提案手法の正解率が最大 70.6 %（200 枚の画像を加えたとき）と提案手法が有効であることが確認できた。

※1 [2] は猫と犬の識別および品種を識別する研究であり、この研究で作成されたデータセットは犬 25 種、猫 12 種※2 がそれぞれ約 200 枚ずつある。

※2 猫 12 種はアビシニアン、ベンガル、バーマン、ボンベイ、ブリティッシュショートヘアー、エジプシャンマウ、メインクーン、ペルシャ、ラグドール、ロシアンブルー、シャム、スフィンクス

1.1 用語

データオギュメンテーション：訓練データの画像に処理を加えることで、データ数を増やす方法。

CNN：畳み込みニューラルネットワーク。基本的な CNN 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繰り返される。この二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。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。

GAN：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、入力に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。

DCGAN：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。

2 今後の課題

- 先行研究である、「Cat and dogs」[2] をよく理解し、より自分のしたい研究に近い論文を読む（または猫でなくても動物の個体識別の論文を読む）

参考文献

- [1] 渡部宏樹, 渡辺裕. Dcgan を用いたデータオギュメンテーションによる猫の品種識別について. 映像情報メディア学会年次大会講演予稿集, Vol. 2016, pp. 24D-5, 2016.
- [2] O. M. Parkhi, A. Vedaldi, A. Zisserman, and C. V. Jawahar. Cats and dogs. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3498-3505, 2012.