

事前学習における FractalDB の性能評価

小笠原大貴¹
Daiki Ogasawara前田英作¹
Eisaku Maeda東京電機大学¹

1 まえがき

画像分類を行うモデルの Pre-Training(PT) には大規模な自然画像データセットが必要とされている。しかし、大規模なデータセットはそれを作成するための労力や正解ラベルの一貫性の欠如、画像の著作権の保護が不十分である可能性、被写体の偏りといった問題を抱えている。これらの諸問題の解決策として片岡らは、フラクタル画像の生成とアノテーションを自動で行う FractalDB を用いて PT を行う手法を提案し、一部の画像分類タスクにおいて既存手法に匹敵する性能を達成した [2]。本研究では、片岡らの研究で使用されたもの以外の Fine-Tuning(FT) 用データセットを使用して PT における FractalDB のパフォーマンスを評価する。

2 実験

2.1 FractalDB の分布の解析

本実験では 10 クラスの FractalDB のクラス内標準偏差と各クラスの平均ベクトル同士の距離を解析した。クラス内標準偏差を図 1、各クラスの平均ベクトル同士の距離を図 2 に示す。解析の結果、各クラスのクラス内標準偏差にはあまり大きな差はなく、各クラスの平均ベクトルには偏りがあることが分かった。



図 1: クラス内標準偏差

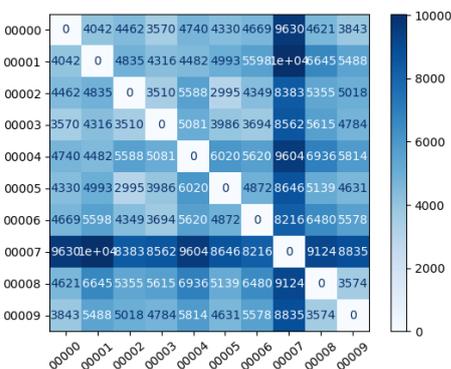


図 2: 各クラスの平均ベクトル同士の距離

2.2 FractalDB の性能評価

2.2.1 データセット

本実験では、PT 用データセットとして ImageNet, FractalDB を使用し、FT 用データセットとして CIFAR-10, CIFAR-100, FashionMNIST, KMNIST [1], K49 [1], MNIST を使用した。

2.2.2 モデルの PT

ImageNet で PT したモデルは、学習済みのモデル¹を利用した。FractalDB での PT は片岡らの論文 [2] を参

¹<https://download.pytorch.org/models/resnet50-19c8e357.pth>

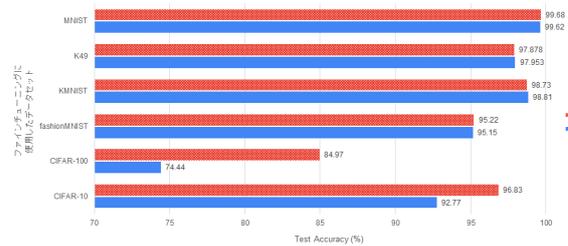


図 3: FT 後の正解率

考にしてパラメータを決定した。

2.2.3 モデルの FT

ImageNet と FractalDB で PT したそれぞれのモデルを複数の自然画像データセットで FT した。FT を行う際のパラメータは PT と同様である。

2.2.4 結果

ImageNet, FractalDB で PT したそれぞれのモデルを複数のデータセットで FT した後のテストデータに対する正解率を図 3 に示す。

図 3 の結果より CIFAR10/100 での FT においては ImageNet は FractalDB と比較して高いパフォーマンスを発揮するが、その他のデータセットでの FT においては差は非常に小さいことが確認できた。

3 考察

- CIFAR-10/100 で FT を行った場合、ImageNet の PT における正解率が FractalDB と比較して高いのは、ImageNet と CIFAR-10/100 のクラスの一部が共通しているため正解率が高くなるのが理由の一つであると推測される。
- KMNIST, K49, MNIST のような自然画像でないデータセットで FT を行う場合、PT における ImageNet と FractalDB のパフォーマンスが拮抗していることから、FractalDB は自然画像でないデータセットの方が相性がいい可能性がある。

4 今後の課題

- 今回の実験では ImageNet と FractalDB の比較を行ったが、今後は PT を行わない場合の精度も含めて比較をすることで、PT における FractalDB のパフォーマンスをより正確に解析することができると考えている。
- 自然画像でないデータセットの種類を増やして実験を行い、FractalDB と相性が良いタスクの特徴を引き続き検討したい。

謝辞

本研究は JSPS 科研費 JP119H01134 の助成を受けた。

参考文献

- [1] Tarin Clanuwat, et al. Deep learning for classical japanese literature, 2018.
- [2] Hirokatsu Kataoka, et al. Pre-training without natural images. In *ACCV*, 2020.