

深層学習識別器の効果的な事前学習のための フラクタル画像を用いたデータ拡張

吉岡 雄健¹ 船津 朋和¹ 永岡 隆² 小塚 健倫³ 根本 充貴⁴
山田 誉大⁵ 木村 裕一^{6,7} 石井 一成^{3,5} 波部 斉^{6,7}

近畿大学大学院 総合理工学研究科¹ 近畿大学 生物理工学部 生命情報工学科²
近畿大学 医学部 放射線医学教室 放射線診断学部門³ 近畿大学 生物理工学部 医用工学科⁴
近畿大学 高度先端総合医療センター PET分子イメージング部⁵ 近畿大学 情報学部 情報学科⁶
近畿大学 情報学研究所⁷

1. はじめに

深層学習では事前学習用データセットがモデルの精度を左右することが知られており、事前学習用データセットは大規模化する傾向にある。大規模データセットの作成には人的コスト、時間的コストがかかり、十分な量の画像を含むデータセットを作成することが困難であるため、事前学習には一般公開されている自然画像で構成された大規模データセットを用いることが多い。なかでも、一般物体認識のためのモデルの事前学習や評価のために用いられる ImageNet (IN) [1]は約 1400 万枚のラベル付き画像を有しており、画像分類や画像認識において広く利用されている。

症例数の少ない医用画像を深層学習で識別する際には、十分なデータを用意することは困難であるため、深層学習を用いた従来手法では、自然画像ベースのデータセットで事前学習を行った後、医用画像でファインチューニングを行う方法や、GAN を用いた類似画像生成によるデータセットの拡張などが行われている。しかし、自然画像と医用画像には異なる特徴があり、医用画像分類のための事前学習を自然画像で行うことが適切であるかは疑問である。また、GAN によるデータ拡張にも限界がある。

本研究では、深層学習を用いた医用画像分類のための事前学習にフラクタル画像を使用したデータ拡張を提案し、フラクタル画像の効果的なデータ拡張方法を検討する。

2. フラクタル画像を用いた事前学習

片岡ら[2]は数式ベースで画像生成からアノテーション、事前学習までを行う数式駆動型自己教師あり学習 (FDSL) を提案し、自然画像分類のための事前学習にフラクタルデータベース (FDB) を使用したモデルの分類精度が自然画像で事前学習を行なったものと同様であることを示した。また、自然画像で構成されたデータセット問題を解決し、理論上無限にデータセットの拡張が可能である。さらに、事前学習を効率的に行うため、各クラス 1 枚のみの画像で事前学習を行う One-instance Fractal Database (OFDB) も提案されている。[3]

2.1 フラクタル画像の生成

IFS (Iterated Function System) を用いて図形を構成する点群 $X = \{x_1, x_2, \dots, x_K\}$ を求めることによりフラクタル画像を生成する。*IFS* は完備距離空間 X において変換 $\omega_i: X \rightarrow X$ とそれに対応する確率 p_i の集合によって定義される (式 (1))。この時の N は組 (ω_i, p_i) の数を表す。

$IFS = \{(\omega_1, p_1), (\omega_2, p_2), \dots, (\omega_i, p_i), \dots, (\omega_N, p_N)\}$ (1)
二次元ユークリッド平面上のフラクタル画像生成時の変換 ω_i は式 (2) である。

$$\omega_i(x) = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} + \begin{bmatrix} e_i \\ f_i \end{bmatrix} \quad (2)$$

p_i は $\sum_{i=1}^N p_i = 1$ の確率であり、式 (3) によって求められる。

$$p_i = \frac{|\det A_i|}{\sum_{i=1}^N |\det A_i|} \quad \text{where } A_i = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \quad (3)$$

初期値から確率 p_i に従って変換 ω_i とするアフィン変換 (式 (2)) を座標 $x_{(t-1)}$ に適用し、新たな座標 x_t を求める作業を $K-1$ 回繰り返すことで、点群 $X = \{x_1, x_2, \dots, x_K\}$ を生成する。これを繰り返し、1000 枚の画像を生成する。

Data Augmentation Using Fractal Images for Effective Pre-Training of Deep Learning Classifier.

†YUKEN YOSHIOKA, TOMOKAZU FUNATSU, Graduate School of Science and Engineering, Kindai University.

‡TAKASHI NAGAOKA, TAKENORI KOZUKA, MITSUTAKA NEMOTO, TAKAHIRO YAMADA, YUICHI KIMURA, KAZUNARI ISHII, HITOSHI HABE, Kindai University.

2.2 データ拡張

本節ではFDB、OFDBのデータ拡張方法について示し、OFDBの効果的なデータ拡張を提案する。

2.2.1 FDB

2.1節で生成した1000枚の画像に対し、IFSのパラメータセットの変更を25回、画像回転を4回、3×3のパッチパターンの変更を10回行うことで、各クラス1000枚ずつ、計100万枚の画像を生成する。

2.2.2 OFDB

2.1節で生成した1000枚の画像のみを使用し、学習時にPyTorchライブラリのRandom affine、Color Jitterを使用することによりデータ拡張を行う。

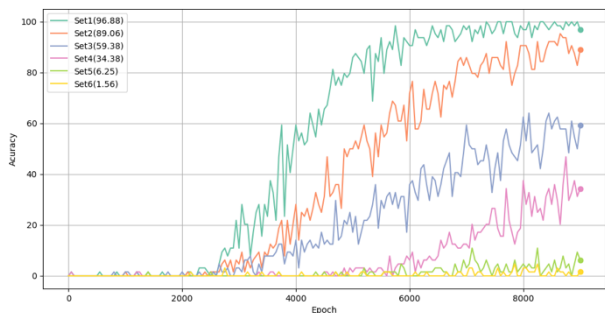


図1:事前学習における正解率

図1はRandom affineの角度、平行移動、拡大縮小の値を変更し事前学習を行った際の正解率である。Set1では、9000回以内に学習が収束しており、モデルが過学習を起こしている可能性が考えられる。そこで、Set1から順にアフィン変換のパラメータの振れ幅を大きくすることで、クラス内で類似度の低い画像を学習するため、事前学習の難易度を上げることができる。これにより、過学習を防ぎ、モデルの汎化性能の向上が期待できる。

3. CT画像識別への適応

事前学習を行ったモデルを使用しCT画像識別のためのファインチューニングを行う。

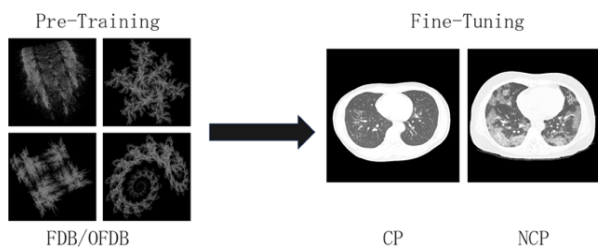


図2:識別器の学習の流れ

3.1 データセット

近畿大学病院で撮影された市中肺炎(CP)と、COVID-19肺炎(NCP)を使用し2クラス識別を行う。それぞれ使用する枚数は、CPが7143枚、NCPが6651枚である。これらの画像は、スライスごとにアノテーションされており、同一患者の画像が学習データとテストデータに混在しないことを考慮した上で均等に分割する。

3.2 実験結果

モデルごとの10分割交差検証の結果を図3に示す。OFDBを事前学習に使用したモデルでは、自然画像で事前学習を行ったモデル(IN)と比較するとPrecisionを除き統計的有意差無く識別可能であり、FDBを事前学習に使用したモデルと比較すると、全ての評価指標において有意差なく識別することができた。また、OFDBではFDBの10分の1の学習量で同等の精度を出しており、計算コストにおいて優れている。

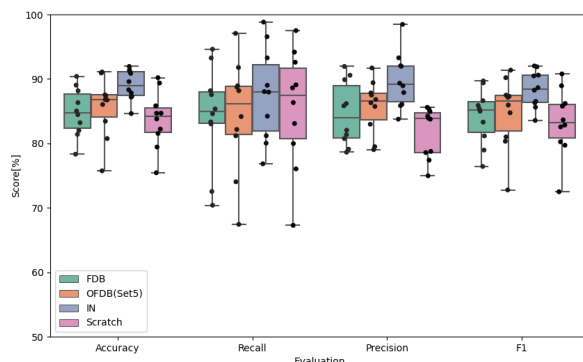


図3:10分割交差検証によるモデルの精度比較

4. おわりに

FDSLは自然画像データベースにおける問題を解決し、学習の規模、難易度を自在に変更できる点において優れている。今後の展望として、フラクタル画像生成時のパラメータ、ハイパーパラメータの調整により、さらなるモデルの性能向上が期待できる。本研究の一部は科研費JP23K11158の補助を受けて行った。

参考文献

- [1] Jia Deng et al., ImageNet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255, 2009.
- [2] Hirokatsu Kataoka et al., Pre-Training Without Natural Images. International Journal of Computer Vision(2022) 130:990–1007.
- [3] Ryo Nakamura et al., Pre-training Vision Transformers with Very Limited Synthesized Images. International Conference on Computer Vision, 2023.