5W-09

# ニュートンフラクタル画像による事前学習効果

近江俊樹* 中村凌† 片岡裕雄† 井上中順* 横田理央‡

*東京工業大学 †産業技術総合研究所 ‡東京工業大学学術情報国際センター

## 1 背景と目的

近年, 画像認識では, 深層学習がスタンダードになっており, 高い認識性能の学習をするためには大規模データセットが必要となる. しかし, 画像データセットを作成する際には多くの問題点がある. 例えば, 医療用画像など, 画像を大量に用意することが困難な場合がある. また, 著作権の問題, 内容の偏り, 倫理的問題について確認する必要がある. さらに, 画像のカテゴリを人の手で正確にラベリングすることが求められる. また,Vision Transformer というモデルでは, 大規模なデータセットを用いることで性能が向上することが確認されているが,JFT-300M/3B に代表されるように, 停止されたものや公開されていないものがある [1].

片岡らは, Iterated Function System(IFS) により生成された 2 次元フラクタル画像データセットである FractalDB と, 数式から生成された画像とその生成パラメータに基づいて付けられた教師ラベルを用いて事前学習モデルを構築する Formula-Driven Supervised Learning(FDSL) を提案した [3]. さらに,FractalDB を 3 次元に拡張し, ランダムな視点から 2 次元に投射した画像を用いる ExFractalDB と, 複数の複雑な輪郭線を描画し, 輪郭形状表現に特化した RCDB というデータセットを提案した [2]. これらのデータセットを用いた FDSL により, 自然画像の画像認識タスクの事前学習に効果があることが実証されている.

しかしながら, FractalDB と ExFractalDB にはフラクタル形状を表現できていない画像も含まれており, RCDB は凸多角形のみを扱っているため画像表現が乏しく, 現状の FDSL に使われるデータセットは多様な画像表現ができていないため性能が不十分だと考えられる. そこで, 本研究では, 多様なフラクタル形状の表現ができ, 十分な画像表現のあるデータセットとして, ニュートン法を用いて生成した画像データセットの NewtonFractalDB(NFDB) を提案し, その事前学習効果を検証する.

## 2 Newton Fractal

式 (1) の反復公式により方程式の根を計算するニュートン法を用いて生成された画像が Newton Fractal である.

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} \tag{1}$$

ここで, $f(z)$ は複素関数であり, $z_n$ は複素数である. さらに式 (2) は緩和係数 $a$ を導入した修正ニュートン法である.

$$z_{n+1} = z_n - a\frac{f(z_n)}{f'(z_n)} \tag{2}$$

ここで, $a$ は複素数である. 式 (2) において $a = 1$ の特別な場合が式 (1) となるため, 本研究では式 (2) を用いて Newton Fractal 画像を生成する.

Newton Fractal 画像は, 画像の各ピクセルに相当する複素数を初期値として, ニュートン法を適用して求まる近似根の種類によって各ピクセルを分類し, その領域を描画する. 例として $f(z) = z^3 - 1, a = 1$ の Newton Fractal 画像を図 1 に示す. $f(z) = z^3 - 1$ は根に $z = 1, \frac{-1+\sqrt{3}i}{2}, \frac{-1-\sqrt{3}i}{2}$ を持ち, 橙, 黄, 赤の領域がそれぞれの根に収束することを表している.

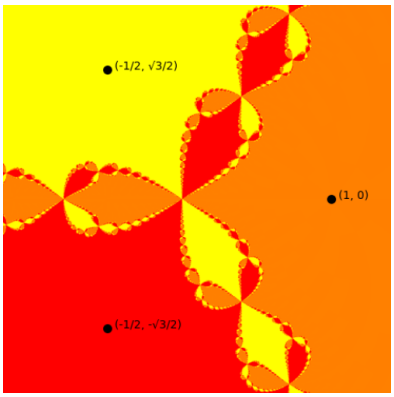


図 1: $f(z) = z^3 - 1, a = 1$ の Newton Fractal

## 3 実験

### 3.1 予備実験

従来の Fracrtal を用いたデータセットはクラス内の画像が 1000 枚である 1k-instance データセット (FractalDB, ExFractalDB) であったが, MIRU2022 の発表において, One-instance 化したデータセット (OFDB, OExFDB) でも十分な事前学習効果が得られることがわかっている. そこで, NFDB を One-instance 化した One-instance NFDB(ONFDB) を用いて実験を行い,OFDB-1k,OExFDB-1k,ImageNet-1k を One-instance にした OIN-1k と比較する.

### 3.2 設定

本実験では Vision Transformer(ViT) の Tiny モデルを用いる. 事前学習で使用するハイパーパラメータは, DeiT [4] で用いられているものをベースとし,Learning Rate を 0.0005, Epochs を 100,000,hflip を 0.5 に変更して行った. ファインチューニングでは DeiT と同じハイパーパラメータで行った. 事前学習効果は, それぞれのデータセットを用いて事前学習を行い, CIFAR10 および CIFAR100 でファインチューニングした精度 (Top-1 accuracy) で評価する.

### 3.3 ONFDB

FDSL により各画像に付与される教師ラベルはニュートン法を適用する関数 $f(z)$ と緩和係数 $a$ に基づくラベルで

**The Effect of Pre-training with Newton Fractal Images**
Toshiki Omi*, Ryo Nakamura†, Hirokatsu Kataoka†, Nakamasa Inoue* and Yokota Rio‡
*Tokyo Institute of Technology
†National Institute of Advanced Industrial Science and Technology
‡Global Scientific Information and Computing Center, Tokyo Institute of Technology

ある. 本実験では簡単のために複素数多項式関数のみを扱い, 方程式の根の数を 10 個に固定し, 生成に用いる 1000 個の関数は固定する.

本実験で使用するデータセットは, 色彩情報が Newton Fractal 画像を用いた事前学習効果にどのように影響するかを調べるために, 上で述べた手法で収束する根で領域を分けて色付けした color 画像, color 画像をグレースケールに変換した grey 画像, 領域の境界のみを白で描画しその他を黒で描画した binary 画像でそれぞれ構成された 3 つのデータセットを用いる. 図 2, 図 3, 図 4 にそれぞれの画像例を示す.

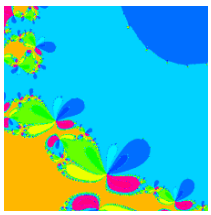

図 2: color

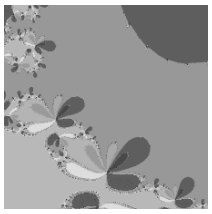

図 3: grey

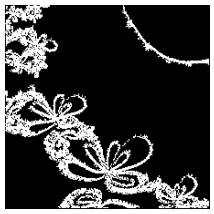

図 4: binary

さらに, 式 2 における緩和係数 $a$ の変化が Newton Fractal 画像を用いた事前学習効果にどのように影響するかを調べるために, $a$ を 0.4 から 2.0 の間の値を取るように変化を加えたデータセット (ONFDB-modified-1k) も用いる. 図 5, 図 6, 図 7 に $a$ を変化させた場合の画像例を示す.

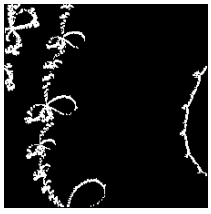

図 5: a=1.0

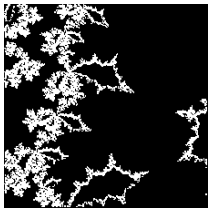

図 6: a=1.8

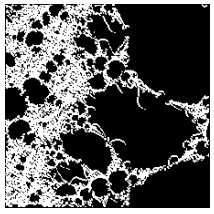

図 7: a=2.0

### 3.4 結果

実験結果を表 1 に示す.

表 1: CIFAR10/100 Finetuning

| Pretrain dataset | Images | CIFAR10 | CIFAR100 |
|---|---|---|---|
| Scratch | | 80.90 | 64.36 |
| OIN-1k | 1,000 | 95.04 | 79.19 |
| OFDB-1k | 1,000 | 96.92 | 83.44 |
| OExFDB-1k | 1,000 | 97.20 | 84.83 |
| ONFDB-1k | | | |
| color | 1,000 | 94.98 | 79.59 |
| grey | 1,000 | 96.99 | 83.62 |
| binary | 1,000 | 97.53 | 85.26 |
| ONFDB-modified-1k | 1,000 | **97.62** | **85.48** |

## 4 結論

実験から, いずれの ONFDB で事前学習した場合でも scratch 学習に比べて大幅な精度向上が見られたため, 事前学習効果があったと考えられる. また, One-instance 学習にも成功していることが分かる. 従来の FDSL と同様の画像表現である白黒画像データセットが最も良い性能を示しており, グレー, カラーの順に性能が下がっていることが分かる.

白黒の ONFDB は, CIFAR10, CIFAR100 ファインチューニングにおいて, 従来手法である OFDB,OExFDB より精度が高く, より多様な画像表現ができていることが考えられる. また, ONFDB-1k-binary より ONFDB-modified-1k が高いことから, 緩和係数を変化させることによる画像表現の変化が事前学習効果に有効であることも示唆される. つまり, Newton Fractal 画像が従来の FDSL データセットよりも多様な画像表現ができており, 事前学習効果に有効であると考えられる.

## 今後の課題

今回実験で用いたデータセットのサイズは 1k-1ins と小規模であるため, より大きな 21k,50k,100k サイズへの拡張や 1k-instance への拡張を行う. また, より小規模化した効率的なデータセットを構築も行う. さらに, 多様な画像表現ができることを利用して, データセットに必要な要素の究明やデータセットの小規模化と高性能化を行う.

## 謝辞

この成果は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の助成事業 (JPNP20006) の結果得られたものです。

## 参考文献

[1] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021.

[2] Hirokatsu Kataoka, Ryo Hayamizu, Ryosuke Yamada, Kodai Nakashima, Sora Takashima, Xinyu Zhang, Edgar Josafat Martinez-Noriega, Nakamasa Inoue, and Rio Yokota. Replacing Labeled Real-Image Datasets With Auto-Generated Contours. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21232–21241, 2022.

[3] Hirokatsu Kataoka, Kazushige Okayasu, Asato Matsumoto, Eisuke Yamagata, Ryosuke Yamada, Nakamasa Inoue, Akio Nakamura, and Yutaka Satoh. Pre-training without Natural Images. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*, November 2020.

[4] Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Herve Jegou. Training Data-efficient Image Transformers & Distillation through Attention. In *International Conference on Machine Learning*, Vol. 139, pp. 10347–10357, July 2021.