

令和 4 年 6 月 20 日現在

機関番号：82626

研究種目：基盤研究(A)（一般）

研究期間：2019～2021

課題番号：19H01134

研究課題名（和文）自然の形成原理に則した深層学習の真相究明

研究課題名（英文）Deep Learning Principle Analysis in Accordance with Natural Law

研究代表者

片岡 裕雄（Kataoka, Hirokatsu）

国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員

研究者番号：70784883

交付決定額（研究期間全体）：（直接経費） 28,700,000円

研究成果の概要（和文）：自然法則を参考にした数式から画像パターンや教師ラベルを自動生成し画像データセットを構築可能な枠組み「数式ドリブン教師あり学習」を提案した。理論的には事前学習フェーズにおいて実画像の代替となる教師ラベル付きの大規模画像データセットを無限に生成することができる。同枠組みは実画像に対して自動で教師ラベルを付与する自己教師あり学習と類似の枠組みであるが、実画像すらも用いないという点でより困難なタスクに挑戦している。現時点では全ての設定において実画像と人間による教師ラベル付の画像データセットの学習には及んでいないが、一部の設定では提案のFractalDBによる学習による精度が超えている設定も確認できた。

研究成果の学術的意義や社会的意義

学術的意義：数式で生成した画像データセットで従来のベースラインと同等まで到達できることを明らかにした。さらに、画像データセットを数式で定義することで、画像認識の原理解析をより容易にした。画像パターンの複雑性が担保され、画像カテゴリは一貫性がある形で割り振られることが重要であった。

社会的意義：従来の画像データセットで問題視されていた、プライバシー侵害・差別的出力・人間のラベル付コストの問題がなくなったため、数式等に権利が含まれていない限りは商用利用が可能なモデルとして提供可能であり、今後AIの社会実装が容易になる。さらに、論文の実験結果が再現できる形でコードや学習済みモデル等を公開するに至った。

研究成果の概要（英文）：This research project has proposed a learning theory named, Formula-Driven Supervised Learning (FDSL) which automatically renders image patterns and their corresponding labels by natural principles such as fractal geometry. Theoretically, the framework enables to generate unlimited number of fractal images based on a natural law. The proposed framework defines a novel learning strategy which is similar to Self-Supervised Learning yet our framework does not require any real images in the pre-training phase. According to the experimental results, the FractalDB pre-trained convolutional neural networks have recorded similar performance rates to de-facto-standard models like ImageNet pre-trained models.

研究分野：コンピュータビジョン

キーワード：Deep Learning Pre-training Transfer Learning Fractal Geometry

1. 研究開始当初の背景

深層学習の成功のためには大規模画像データセットを用いた事前学習により基礎的な視覚機能を学習させる必要があった。事前学習を通して作成した学習済みモデルは、凡ゆる画像認識のタスクに対して転用可能であり、開発の初期段階における強力なベースラインを構築する上では必須の技術となっていた。しかし大規模画像データセットに関して、日本では個人情報保護法や著作権法など画像利用に関する制約や、各企業が独自収集した画像を公開しづらいことが共通の大規模画像データセットの利用を妨げる原因となっていた。さらに、世界的にはプラットフォームである Google が YouTube を、Facebook (現 Meta) が Instagram を買収し、大量の画像や動画の中から二次利用が可能なライセンス (Creative Commons License) の画像・動画を集めて大規模画像データセット JFT-300M (3 億画像データセット)・YouTube-8M (800 万動画データセット)・Instagram-3.5B (35 億画像データセット)として提案してきた。しかし、JFT-300M / Instagram-3.5B に関して論文上は世界最高水準の精度が報告されているものの、非公開であるため社外の研究者による再現が不可能であるという問題を孕んでいた。スタンフォード大学が公開し、デファクトスタンダードとして扱われていた ImageNet-1k/21k データセットも元々不特定多数の画像を Web 上から収集してきたため、プライバシー保護などの保証ができず商用利用が明示的に禁止されていたが、2019 年頃にはデータセット内の偏りが人種差別など不公平な結果に繋がるとの報告が行われ、人物関連のラベルの一部が削除された。このように、国内のみならず、世界的にも大規模画像データセットの利用に関しては社会問題となっており、早急な解決が待たれる状況であった。

2. 研究の目的

画像データセットをめぐる問題を根本から解決するためには実画像を一切用いずに大規模画像データセットを構築し、事前学習により基礎的な視覚機能を学習させることが重要と考える。さらに、深層学習が隆盛したことにより、画像認識の精度は飛躍的に向上したことで「自然画像とは何だろうか？」そして「その画像カテゴリとは？」という本質に迫る問いの解明がより重要になってきたと言える。研究目的として、自然の形成原理に則し、能動的に生成した画像パターンとその画像カテゴリを教師ラベルとした機械学習方法を網羅的に探索することで両者の問いに迫る。自然法則を参考にした数式から画像パターンや教師ラベルを自動生成して大規模画像データセットを構築することで、画像認識システムを社会実装する際の権利関係に依らずデータの利活用ができるのみならず、原理が明らかになっている数式から画像データセットを構築できるため、少なくとも学習データに関しては透明性を担保して原理検証ができると考える。

また、本研究課題にて構築したコードやデータセット、事前学習済みモデル (学習済みパラメータ)などはプロジェクトページなどで広く公開する。この点においては、計算リソースが比較的整っていない場合でも再現できるように整備することが重要である。

3. 研究の方法

実画像や人間の教師ラベルを一切用いず、自然法則を参考にした数式から画像パターンと画像カテゴリ (教師ラベル)を自動生成しながら大規模画像データセットを構築し、深層学習の事前学習モデルを作成する方法論は「数式ドリブン教師あり学習 (Formula-Driven Supervised Learning; FDSL)」として提案した。具体的にはフラクタル幾何により形成される画像パターンを直接画像平面に投影するために、Iterated Function System (IFS)を基本的な関数として描画する。ここで、IFS による画像データセット構築について (1) 画像カテゴリ探索、(2) フラクタル画像生成、(3) 画像インスタンス拡張の順に説明する。

(1) 画像カテゴリ探索: IFS を構築する際には 6 種類の基本パラメータセットをランダム生成して、画像に投影した際に簡単なクオリティチェックを行うことにより画像カテゴリを自動で登録することができる。ここで、簡単なクオリティチェックとは、画像全体に占めるフラクタル領域の画素数の割合を閾値として比較する。基本カテゴリ数は従来法に準じて 1,000 に設定しているため、ベースとなる画像データセットは 1,000 カテゴリとなる。

(2) フラクタル画像生成: フラクタル幾何は自己相似性を持つ形状のことであり、画像を大域的に見た特徴と局所的に見た特徴が類似することを示しており、植物や地形図などにも現れる入り組んだ複雑な形状を生成できるとされている。IFS においては、基本的にはある初期値から開始し、(1) で探索した基本パラメータセットを確率的に変更しながら再帰的に画素を打ち込んでいくことで複雑な形状を描画することができる。

(3) 画像インスタンス拡張：画像データセットを効果的に構築するためには、各カテゴリに対して画像の見え方を多少変える必要がある。フラクタル画像の生成に対して基本パラメータセットの重み変更(25種類)、画像回転(4方向)、3x3ピクセルのパッチパターン変更(10種類)によりばらつきを与える。拡張により各画像1,000(25 x 4 x 10)種類の見え方のパターンを用意することになる。

標準的な画像データセットは100万オアダの画像を含むため、画像カテゴリ1,000種類 x 画像カテゴリあたりの画像インスタンス1,000種類の100万画像のデータセットとなり、数量の面では標準画像データセットに匹敵するものができる。提案データセットはFractalDBと呼称し、標準画像データセットと同様に事前学習に適用する大規模画像データセットとして使用する。

次にFractalDBを用いて深層学習の事前学習を実施、実画像を含むデータセットに対して追加学習を行う。ここで、FractalDBのパラメータチューニングには膨大なパラメータセットの探索が行われたが、事前学習効果を高めるパラメータは主に下記の3種類であった。

- ・画像カテゴリ数 / 画像インスタンス数：画像データセットのスケールに直結するパラメータであり、より大きくするほど精度向上が見られた。比較的画像カテゴリ数を大きくする方が画像インスタンス数を大きくするよりも効率良く精度向上することが明らかとなった。探索の結果、画像カテゴリ1,000種類 x 画像インスタンス1,000種類にするデータセット構成に到達した。

- ・パッチによるフラクタル画像描画：フラクタル画像を生成する際、1x1ピクセルのドットよりも3x3ピクセルのパッチで描画した方が精度向上することが確認された。

- ・重みパラメータ：フラクタル画像を生成する際の基本パラメータセットの重みパラメータを変動させることで画像インスタンスを増加させ、事前学習効果を高める影響があることを明らかにした。

4. 研究成果

本プロジェクトを通じた研究の結果、数式から生成した画像データセットによる事前学習でも画像認識のベースラインを向上させ、標準画像データセットに近い精度まで到達することが明らかとなった。また、深層学習における事前学習においては実画像を用いず、データセットの偏りも生じないため、権利侵害や差別的出力を含むことなく事前学習モデルを構築できると言える。2022年現在、我々からは権利を主張していないため、数式に権利が含まれていない限りは商用利用も可能である。深層学習の事前学習用のデータセットを数式で自動生成するプロジェクトを通して、画像自体は着色に依らず形状の複雑性が担保されていること、画像カテゴリは画像パッチの形状に応じて一貫性がある形で割り振られていれば効果的な視覚機能を獲得できる。

論文などの研究成果としては、コンピュータビジョン分野のメジャーな国際会議 Asian Conference on Computer Vision (ACCV 2020)にて Best Paper Honorable Mention Award を受賞、さらに実験等を拡張した手法が同分野トップジャーナルの International Journal of Computer Vision (IJCV)に採択された。IJCVの論文は、公開からの数ヶ月で30,000回以上も閲覧があるなど、世界的にもインパクトある研究成果となった。さらに、論文に対応するコードや学習済みモデル、データセットの生データやレンダリングコードなどを再現できる形で公開した。論文や資料含め、研究プロジェクト (<https://hirokatsukataoka16.github.io/Pretraining-without-Natural-Images/>) から公開されている。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 Hirokatsu Kataoka, Kazushige Okayasu, Asato Matsumoto, Eisuke Yamagata, Ryosuke Yamada, Nakamasa Inoue, Akio Nakamura, Yutaka Satoh
2. 発表標題 Pre-training without Natural Images
3. 学会等名 Asian Conference on Computer Vision (ACCV) (国際学会)
4. 発表年 2020年

1. 発表者名 Nakamasa Inoue, Eisuke Yamagata, Hirokatsu Kataoka
2. 発表標題 Initialization Using Perlin Noise for Training Networks with a Limited Amount of Data
3. 学会等名 International Conference on Pattern Recognition (ICPR) (国際学会)
4. 発表年 2021年

1. 発表者名 SHAO GUANGXIN、片岡 裕雄、佐藤 雄隆
2. 発表標題 フラクタル幾何学を用いたデータセットの拡張および特性評価
3. 学会等名 画像センシングシンポジウム (SSI2019)
4. 発表年 2019年

1. 発表者名 Hirokatsu Kataoka, Kazushige Okayasu, Asato Matsumoto, Eisuke Yamagata, Ryosuke Yamada, Nakamasa Inoue, Akio Nakamura, Yutaka Satoh
2. 発表標題 Pre-training without Natural Images
3. 学会等名 International Journal of Computer Vision (IJCV)
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Pre-training without Natural Images
<https://hirokatsukataoka16.github.io/Pretraining-without-Natural-Images/>
FractalDB-Pretrained-ResNet-PyTorch
<https://github.com/hirokatsukataoka16/FractalDB-Pretrained-ResNet-PyTorch>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	中村 明生 (Nakamura Akio) (00334152)	東京電機大学・未来科学部・教授 (32657)	
研究分担者	井上 中順 (Inoue Nakamasa) (10733397)	東京工業大学・情報理工学院・助教 (12608)	
研究分担者	前田 英作 (Maeda Eisaku) (90396143)	東京電機大学・システムデザイン工学部・教授 (32657)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------