

視点依存性から着想を得た 深層学習画像分類のためのデータ刈り込み戦略

長野 駿介[†] 熊田 孝恒[‡] 中島 亮一[§] 中原 裕之[¶]
京都大学[†] 京都大学[‡] 京都大学[§] 京都大学・理化学研究所[¶]

1 はじめに

データ刈り込みとは、テスト精度を大きく犠牲にせず学習データセットを削減する手法であり、深層学習においても有効である。これは増大する計算およびインフラコストの削減の手段として重要な役割を担っている。深層学習の画像認識における多くの既存のデータ刈り込みメトリックは、その効果を実現するためにモデル学習を必要としている¹⁻³が、これには時間とリソースがかかる。そこで、人間のオブジェクト認知特性（視点依存性）に着目し、オブジェクトを観察する視点によって変わる画像特徴に基づくメトリックを提案する。このメトリックはモデル学習を必要としないため、効率的かつ迅速にデータ刈り込みを行うことができる。

本研究では視点に依存して変化する特徴量として、①画像に映るオブジェクトのパーツ数、②画像に投影されるオブジェクトの面積というメトリックを利用する。これは、映るパーツ数または面積の大きさと、画像がもつ情報量すなわち分かりやすさが相関するという仮説に基づく。①は、200種類の鳥類からなるデータセット CUB200-2011⁴（鳥の15個の身体のパーツについて、それぞれが映っているかどうかのアノテーションされている）を用いて、このアノテーションをパーツ数のメトリックとする。②は、ShapeNetの3Dモデルの複数視点からのレンダリング画像による自作データセットを用いて、画像に映るオブジェクトの面積をメトリックとする。

最適なデータ刈り込み戦略は初期データの量に依存するといわれている⁵。具体的には、初期データが豊富な場合には、より難しい例を残し、初期データが貧弱な場合には、より簡単な例を残すことが推奨される。面積のメトリックにおいては、データサイズに応じて最適な戦略が変わるかについても検証する。

実験により、これらのメトリックは画像分類タスクにおいてランダム刈り込みよりも優れた性能を示した。

刈り込み戦略	精度
ランダム戦略	0.8339 ± 0.0026
最少パーツ戦略	0.8310
最多パーツ戦略	0.8366

表1 刈り込み結果のベンチマーク表。CUB200-2011のデータセットのResNet-50による画像分類の精度（刈り込みなしの精度:0.845）。最高値は太字で強調している。ランダム刈り込みに関しては、スコアの標準偏差を記載している。

2 パーツ数をメトリックとしたデータ刈り込み

2.1 実験設定

CUB200-2011の画像分類タスクはPytorchライブラリを通じて、ImageNet21kで事前学習したResNet-50⁶を使用する。各モデルは、基本学習率 = 0.001、重み減衰 = $1e-4$ 、モメンタム = 0.9のSGDを用いて、バッチサイズ = 64の1枚のNVIDIA RTX4090 32GBグラフィックカードの単一ノードで学習させる。すべてのモデルは、95エポックで学習される。データ刈り込みには、3つの戦略を適用する。各クラスからランダムに10%ずつを削除するランダム戦略、各クラスで画像に映るパーツ数が少ない下位10%を削除する最少パーツ戦略、パーツ数が多い上位10%を削除する最多パーツ戦略である。CUB200-2011のオリジナル学習データセット5994枚から、データ刈込によって5394枚まで削減される。これらの戦略を使用して、CUB200-2011のテストセット5794枚における画像分類（200クラス分類）の精度を比較する。

2.2 結果と考察

実験結果を表1に示す。最多パーツ戦略は、最少パーツ戦略またはランダム戦略よりも、モデルのテスト精度が高い結果を示している。

本実験は事前学習済みのモデルを使うことにより、初期データが豊富なケースと考えられる。そのため、難しい例を多く残す最多パーツ戦略が最良となる結果は私たちの想定と整合的であり、このケースでのデータ刈り込みが有効であることを示す。

3 面積をメトリックとしたデータ刈り込み

3.1 実験手順

本実験ではShapenetの3Dモデルのレンダリング画像の自作データセットを用いて、ResNet-18アーキテクチャで画像分類タスクを行う。Shapenetのモデル学習は、Pytorch

Data Pruning Strategy Inspired by Viewpoint Dependence for Image Classification Tasks in Deep Learning

[†] Shunsuke Nagano, Kyoto University

[‡] Takatsune Kumada, Kyoto University

[§] Ryoichi Nakashima, Kyoto University

[¶] Hiroyuki Nakahara, Kyoto University, RIKEN

戦略/データサイズ	250	500
ランダム戦略	0.6486 ± 0.0119	0.7608 ± 0.0054
最小面積戦略	0.6517 ± 0.0109	0.7607 ± 0.0078
最大面積戦略	0.6493 ± 0.0094	0.7562 ± 0.0069

表2 小さな初期データセットの刈り込み結果のベンチマーク表

戦略/データサイズ	1000	2000
ランダム戦略	0.8338 ± 0.0049	0.8677 ± 0.0049
最小面積戦略	0.827 ± 0.0060	0.8629 ± 0.0041
最大面積戦略	0.8357 ± 0.0025	0.8703 ± 0.0029

表3 大きな初期データセットの刈り込み結果のベンチマーク表

ライブラリを通じて、標準的な ResNet-18 を使用して行う。モデルは ResNet-18 の基本学習率=0.001, モメンタム=0.9 の SGD (確率的勾配降下法) を用いて、バッチサイズ=4 の 1 枚の NVIDIA RTX4090 32GB グラフィックカードの単一ノードで学習させた。すべてのモデルは、30 エポックで学習された (7 エポックごとに学習率を 0.1 の係数で減衰)。データセットとして ShapeNet の 3D モデルデータセットから選択された 9 クラスのオブジェクト (ベンチ, 自転車, カメラ, 自動車, イス, 帽子, ヘルメット, ノートパソコン, ソファ) を使用する。様々な視点からレンダリングして、合計 9 万枚の画像 (9 クラス × 100 種 × 100 枚) を生成した。

本実験では、初期データセットのサイズの違いによるモデルパフォーマンスの違いを観察するため、各クラスの学習データサイズを 250 枚, 500 枚, 1000 枚, 2000 枚と変動させ、これら各サイズに対して個別のモデルを学習させた。実行時のランダムな変動を考慮し、全 9 万枚のデータセットをそれぞれのデータサイズごとに分割し、それぞれの分割データセットを使用してモデルを 1 回ずつ学習させることで、各クラス 1000 枚のテストセットにおけるモデルの画像分類タスク (9 クラス分類) の平均精度を算出する。

データ刈り込みには、3 つの戦略を適用する。ランダム戦略、各クラスで画像に映る面積が小さい下位 10% を削除する最小面積戦略、面積が大きい上位 10% を削除する最大面積戦略である。

3.2 結果

小さな初期データセットの場合 (250 と 500 のデータサイズ) (表 2)

500 のデータサイズでは、最小面積戦略が最大面積戦略よりも精度が高く、250 のデータサイズでは最小面積戦略がランダム戦略をも上回っている。

大きな初期データセットの場合 (1000 と 2000 のデータサイズ) (表 3)

最大面積戦略は、最小面積戦略またはランダム戦略よりも、モデルのテスト精度が高い結果を示している。

実験結果は、豊富な初期データセットの場合には面積が

大きい画像 (簡単な例) を刈り込むことでモデルの学習効率と精度が向上することを示した。対照的に、貧弱な初期データセットでは、面積が小さい画像 (難しい例) を刈り込むことが効果的である。これは、面積のメトリックにおいてデータサイズに応じて最適な戦略が変わることを示す。

4 まとめ

本研究は、画像認識における計算インフラコストの増大という問題に対し、視点依存性を利用したメトリックによるデータ刈り込みによる解決策を提案した。実験により、オブジェクトの映る面積とパーツ数をメトリックとすることが画像分類タスクにおいて有効であることが示された。本研究は ShapeNet による自作データセットと CUB-200-2011 の 2 つのデータセットに限定して検証した。しかしながら、面積は 3D モデルによるデータ増強へ容易に応用可能であり、パーツ数も適切なアノテーションが行えれば他のデータセットにおいても検証可能であるため、将来的な応用への可能性は非常に大きい。

参考文献

- [1] Vitaly Feldman and Chiyuan Zhang. What neural networks memorize and why: Discovering the long tail via influence estimation. In *Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS'20*, Red Hook, NY, USA, 2020. Curran Associates Inc.
- [2] Mansheej Paul, Surya Ganguli, and Gintare Karolina Dziugaite. Deep learning on a data diet: Finding important examples early in training. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, Vol. 34, pp. 20596–20607. Curran Associates, Inc., 2021.
- [3] Mariya Toneva, Alessandro Sordani, Remi Tachet des Combes, Adam Trischler, Yoshua Bengio, and Geoffrey J. Gordon. An empirical study of example forgetting during deep neural network learning. In *International Conference on Learning Representations*, 2019.
- [4] Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, and Serge Belongie. The caltech-ucsd birds-200-2011 dataset, August 2023.
- [5] Ben Sorscher, Robert Geirhos, Shashank Shekhar, Surya Ganguli, and Ari Morcos. Beyond neural scaling laws: beating power law scaling via data pruning. *Advances in Neural Information Processing Systems*, Vol. 35, pp. 19523–19536, 2022.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.