

未学習物体の分類手法に関する一検討

○渡邊響輝（日大工・情報・院生） 田中宏卓（日大工・情報）
和泉勇治（日大工・情報） 加瀬澤正（日大工・情報）

1. まえがき

近年、機械学習の技術向上により、画像中の物体認識技術は目覚ましく向上している。しかしながら、そのほとんどは既に学習された物体を分類するための技術・手法であり、我々が日常的に行っている、未知の物体についての分類問題は未解決のままである。

そのため、このような問題の解決を目的として、最近の研究では、画像を用いて、複数の階層において同時に認識するための手法に関する検討が進められている。この手法が確立できれば、未学習の物体であっても、その物体が上位階層のどのカテゴリーに属するのかを推論したり、少ない画像を用いた学習のみにより、その物体を認識対象として新たに追加したりすることができる。

そこで、本研究では、既存ネットワークモデルがどの程度上位階層についての分類能力を持つのかを調べ、未学習物体の分類に適したネットワークモデルの検討を行う。

2. 複数階層のカテゴリー分類を行うニューラルネットワーク

物体は階層的なカテゴリー構造を持っている。猫を例に取って考えるとき、そのカテゴリーの構造は、たとえば図1のように表すことができる。以下では、図中の最下層、すなわち「猫」を含む階層を小分類と呼ぶことにする。また、「哺乳類」を含む階層を中分類、「動物」を悩む階層を大分類と呼ぶことにする。

3. 実験1

3.1 実験方法

実験1では、既存のニューラルネットワークモデルが、潜在空間において、上位階層の共通概念の表象である大分類をどの程度表現できているのかについて調べる。具体的には、潜在空間のベクトルとして、各モデルの出力層の直前層の特徴ベ

クトルを抽出し、大分類のクラスごとに平均ベクトルと標準偏差ベクトルを求める。そして、それらに基づき、大分類の各クラスに属するベクトルの分離度合いを調べる。

3.2 実験条件

今回想定する階層構造を表1に示す。実験では、小分類のカテゴリーを推論するだけでなく、表1の大分類3クラス、中分類14クラスのいずれであるかをそれぞれ出力するように、出力層のみをファインチューニングした Inception v3²⁾, MobileNet v2³⁾, NasNetMobile⁴⁾, ResNet50⁵⁾, の4種類のモデルを使用する。モデルの概略構成を図3に示す。

また、今回の実験では、学習には使用していない未知物体を集めたテスト用のデータセットを入力として用いることにする。小分類クラス一つにつき、テスト用画像は50枚を使用する。サイズは112×112とし、カラー画像とする。

3.3 実験結果

ファインチューニング後の各モデルの、大分類の各クラスに属する物体の特徴ベクトルの平均ベクトルを用いて算出した大分類におけるクラス間距離を表3に、標準偏差ベクトルのノルムを表4に示す。ここで、標準偏差ベクトルのノルムとは、当該クラスに属する物体の特徴ベクトルを用いて算出した多次元標準偏差ベクトルのノルムである。

また、クラス間距離を、それを与える両クラスの標準偏差ベクトルのノルムの和で除算した値を表5に示す。この値を分離度と呼ぶことにする。

3.4 考察

今回実験に使用したモデルの中で、特に分離度が大きくなったのは MobileNet v2 であった。一般に、クラス間距離が大きいほど各クラスが潜在空間上で分かれており、分散が小さいほど分類結果のばらつきが小さくなると考えられるので、分離

度は大きくなるほどクラス間の分類が明確にできていることになる。

4. 結言

本研究では、複数階層のカテゴリ分類を行うニューラルネットワークの構築を目的として、既存のネットワークモデルに複数階層の出力を持たせ、その潜在空間における特徴ベクトルの分離度を調べた。

また、上記のネットワークモデルと筆者が提案したネットワークモデルを用いて大分類・中分類階層における認識性能検証をし、提案モデルの正答率が最も高いことを確認した。

今回の実験を参考にし、今後は潜在空間に大分類・中分類の特徴ベクトルがクラスターを形成するようなネットワークモデルの構築方法を検討していく予定である。

参考文献

- 1) 渡邊響輝, “複数階層においてカテゴリ推論を行うニューラルネットワークモデル”, 令和3年度卒業研究発表会(日本大学工学部情報工学科) C-1-1.
- 2) Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, “Rethinking the Inception Architecture for Computer Vision”, CVPR 2016.

3) Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam,

“MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”,
<https://arxiv.org/abs/1704.04861>.

4) Barret Zoph, Vijay Vasudevan, Jonathon Shlens, Quoc V. Le, “Learning Transferable Architectures for Scalable Image Recognition”,
 CVPR 2018.

5) Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian SunDeep, “Residual Learning for Image Recognition”,
 CVPR 2016.

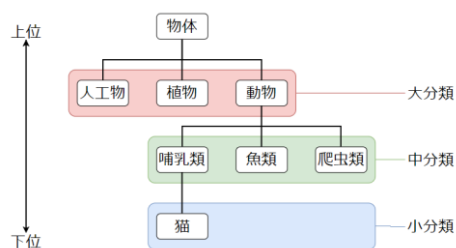


図1 階層的なカテゴリーのイメージ

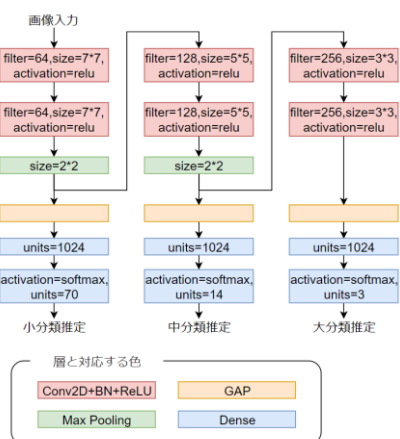


図2 提案ネットワークモデルの構成

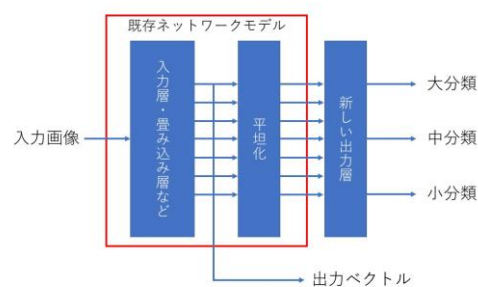


図3 ファインチューニングモデルの概略構成

大分類	中分類	小分類
動物	哺乳類	ボーダーテリア, ペルシャ猫 バグ, 豚, シマウマ
	爬虫類	コモドドラゴン, ナイルワニ, ニシキヘビ, アカウミガメ, エリマキトカゲ
	魚類	金魚, シビレエイ, オニカマス, ホホジロザメ, クマノミ
	鳥類	クジャク, ワシ, ニワトリ, ダチョウ, コンドル
	昆虫	コガネグモ, テントウムシ, ハエ, バッタ, トンボ
植物	果物	パイナップル, レモン, ザクロ, イチジク, バナナ
	野菜	キャベツ, キュウリ, イチゴ, トウモロコシ, ピーマン
	キノコ	キノコ, マイタケ, ハラタケ, シロシメンタケ, スッポンタケ
	花	ヒナギク, たんぽぽ, バラ, ヒマワリ, チューリップ
人工物	建築物	刑務所, 高架橋, レストラン, 灯台, 噴水
	乗り物	新幹線, マウンテンバイク, 旅客機, ミニバン, タクシー
	家具	本棚, 四柱式ベッド, 机, ロッキングチェア, 扇風機
	道具	シャベル, ハンマー, 手斧, 定規, ボールペン
	衣類	Tシャツ, ジーンズ, ブラジャー, セーター, ジャケット・コート

表1 カテゴリーの持つ階層構造

大分類	中分類	小分類
動物	哺乳類	キットギツネ,レッサーパンダ,ゴリラ
	爬虫類	イモリ,キスイガメ,ウィップテール・リザード
	魚類	ウナギ,ダツ,アカエイ
	鳥類	フクロウ,メキシコカワガラス,鴨
	昆虫	ハンミョウ,オサムシ,蜂
植物	果物	オレンジ,グラニースミス,パラムツ
	野菜	ブロッコリー,ズッキーニ,カボチャ
	キノコ	シャグマ,アミガサタケ,ボルチーニ,テングタケ
	花	アジサイ,ハイビスカス,ユリ
人工物	建築物	納屋,ダム,城
	乗り物	飛行船,モーターボート,リムジン
	家具	食器棚,折りたたみいす,洗濯機
	道具	傘,モップ,カップ
	衣類	毛皮のコート,トレーナー,ボンチョ

表2 テスト用データセットの構成

クラス間距離	動物～植物	動物～人工物	植物～人工物
Inception v3	30.51	5.46	36.39
MobileNet v2	0.06	0.06	0.28
NasNetMobile	3.94	3.45	2.42
ResNet 50	0.00	0.00	0.04

表3 大分類におけるクラス間距離

分離度	動物～植物	動物～人工物	植物～人工物
Inception v3	0.0175	0.0029	0.0177
MobileNet v2	0.3503	0.0676	0.3224
NasNetMobile	0.0069	0.0098	0.0067
ResNet 50	0.0683	0.0031	0.0670

表5 分離度

標準偏差ベクトルのノルム	動物	植物	人工物
Inception v3	957.32	1122.77	933.85
MobileNet v2	0.43	0.37	0.50
NasNetMobile	170.32	179.92	183.18
ResNet 50	0.30	0.30	0.31

表4 標準偏差のノルム

正答率	動物	植物	人工物	全体
提案モデル	0.529	0.822	0.713	0.688
Inception v3	0.000	0.000	1.000	0.333
MobileNet v2	0.443	0.448	0.283	0.391
NasNetMobile	0.108	0.853	0.136	0.366
ResNet 50	0.456	0.025	0.828	0.436

表6 大分類クラスの正答率

正答率	哺乳類	爬虫類	魚類	鳥類	昆虫	果物	野菜	キノコ	花	建築物	乗り物	家具	道具	衣類	全体
提案モデル	0.167	0.107	0.420	0.173	0.313	0.380	0.380	0.827	0.647	0.667	0.533	0.400	0.340	0.180	0.395
Inception v3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.071
MobileNet v2	0.107	0.113	0.033	0.067	0.013	0.080	0.527	0.000	0.040	0.093	0.093	0.007	0.080	0.033	0.092
NasNetMobile	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.071
ResNet 50	0.047	0.020	0.033	0.380	0.093	0.000	0.033	0.000	0.013	0.113	0.033	0.260	0.313	0.467	0.129

表7 中分類クラスの正答率