

令和 5 年 6 月 16 日現在

機関番号：12601

研究種目：若手研究

研究期間：2019～2022

課題番号：19K20290

研究課題名（和文）群表現論に基づく少数動画からの高精度な特徴学習

研究課題名（英文）Feature Learning from Few-shot Videos Based on Group Representation Theory

研究代表者

椋田 悠介（Mukuta, Yusuke）

東京大学・先端科学技術研究センター・講師

研究者番号：50830874

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：共変性を持ったNeural NetworksであるEquivariant Neural Networksに対するモデル構造の拡張と学習手法を提案した。モデル構造の拡張として、画像の局所情報をまとめ上げて一つの全域特徴を作り上げる特徴量コーディング手法の共変な拡張を提案した。新たな学習手法として、Equivariant Neural Networksに対して自己教師付き学習を行うための共変な損失であるequivariant pretext labelsとinvariant contrastive lossを提案した。

研究成果の学術的意義や社会的意義

提案手法はDeep Neural Networksに対して共変性の事前知識を陽に組み込むためのEquivariant Neural Networksを活用するための枠組みになっている。提案手法により通常のDeep Neural Networksより高性能でかつ性質がより明らかな認識が行えることが期待される。理論的にも不変な特徴量の有効性の理由の一端となる性質を証明し、不変な特徴量に対する理解の促進につながった。

研究成果の概要（英文）：We proposed an extension of the model structure and the learning method for Equivariant Neural Networks, which are the Neural Networks model that considers equivariance. As an extension of the model structure, we proposed an equivariant extension of the feature coding method, in which local features in an image are summarized to form a single global feature. As a new learning method, we proposed equivariant pretext labels and invariant contrastive loss, which are equivariant losses for self-supervised learning that matches the structural restriction of Equivariant Neural Networks.

研究分野：Machine Learning

キーワード：Invariance Feature Learning Machine Learning Computer Vision

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

(1) Convolutional Neural Networks 技術の発展により画像認識や動画認識の精度は飛躍的に発展している。一方でデータから全ての特徴量を自動的に獲得する Convolutional Neural Networks には性能を発揮するために学習データが大量に必要であることや、学習によって得た特徴量の性質を理解することが難しいといった問題点が挙げられる。

(2) そこで本研究では Deep Neural Networks に不変性の考え方を取り入れた Equivariant Neural Networks に着目する。Equivariant Neural Networks は入力データに対する不変性を Neural Networks の構造に埋め込む手法であり、不変性の事前知識を活用することにより (i)Neural Networks の構造に制約が加わることでデータ数に対して性能の高い識別機が学習可能になる、(ii)学習されたモデルに対して不変性を持っていることが保証されている、といった利点がある。

(3) 本研究では Equivariant Neural Networks のより広範な応用に向け、モデル構造の拡張や効率の良い学習手法の提案に取り組む。

2. 研究の目的

(1) 本研究では Equivariant Neural Networks に対して共変性を考慮した新たなレイヤの提案やモデル学習手法を提案することにより、認識精度の高い特徴抽出モデルの構築方法を確立することを目的とする。

3. 研究の方法

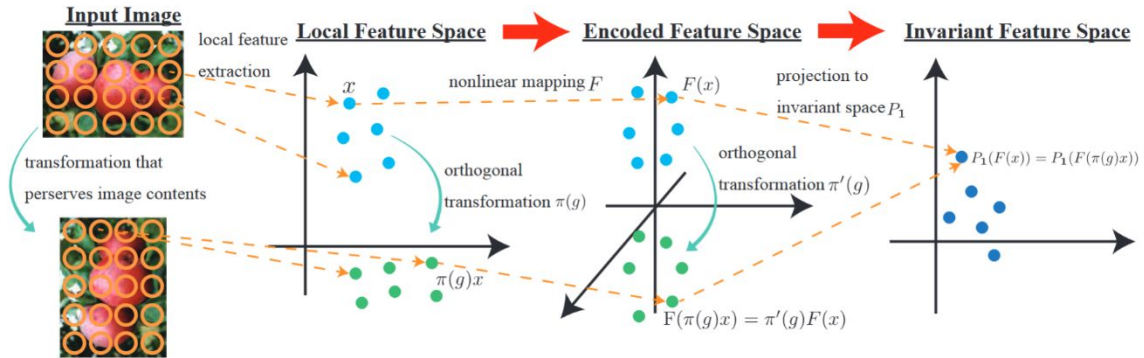


図 1 提案する共変な特徴量コーディング手法のフレームワーク

(1) Equivariant Neural Networks のモデル拡張に取り組む。Equivariant Neural Networks において最終的に得られた特徴量をまとめ上げる **global pooling** 層は不変性に対応する変換群方向において平均や最大値を取る操作を行うという単純な物であった。そこで本研究では古典的な画像認識において局所的な情報をまとめて大域的な特徴を構成する手段であった特徴量コーディングの手法を適用することを試みる。この時に重要なのは特徴量コーディング手法を考える変換群の影響の計算が容易になるように拡張することである。そのために本研究では入力情報の積に対応するテンソル積に着目し、既存特徴コーディング手法をテンソル積に書き直し、その後対応するテンソル積から不変成分を取り出すという形で新たな特徴量のまとめ上げ手法を提案した。

(2) Equivariant Neural Networks の新たな学習手法を提案した。教師無しの特徴学習手法として **self-supervised learning** と呼ばれる手法はデータに対する事前知識から **pretext task** と呼ばれる仮想的なタスクを人手で作成し、そのタスクを識別モデルに学習させることによって特徴抽出器を獲得する手法であり、良い事前知識を用いることにより教師ラベル無しでも優れた特徴抽出器を学習できることが知られている。しかしながら自己教師付き学習を Equivariant Neural Networks に適用する時には Equivariant Neural Networks の共変性が問題になる。すなわち Equivariant Neural Networks は共変性を持った関数しか学習できないため、

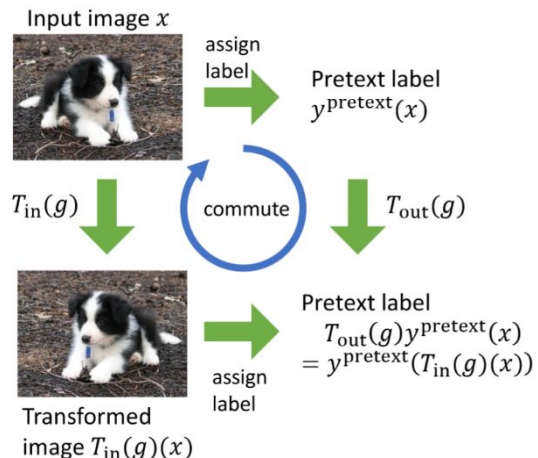


図 2 equivariant pretext label のイメージ

pretext task が共変性を破る時その最適解が学習できなくなるという問題がある。そこで共変性を保ったまま学習を行うために pretext task を制御する手法として、pretext labels の空間に変換が働いており入力に対する変換と共変になるという equivariant pretext labels の概念と、入力データの一部に変換をかけてもデータ間の類似度が変わらないように要請する invariant contrastive loss の概念を提案した。提案した概念に基づき既存の自己教師付き学習に対応する pretext task に対してその共変な拡張を提案した。

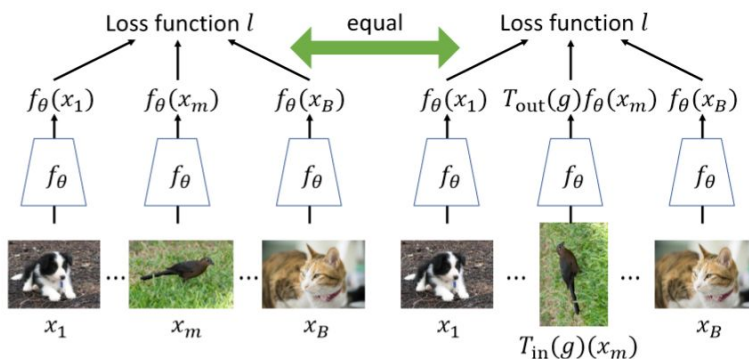


図 3 invariant contrastive loss のイメージ

4. 研究成果

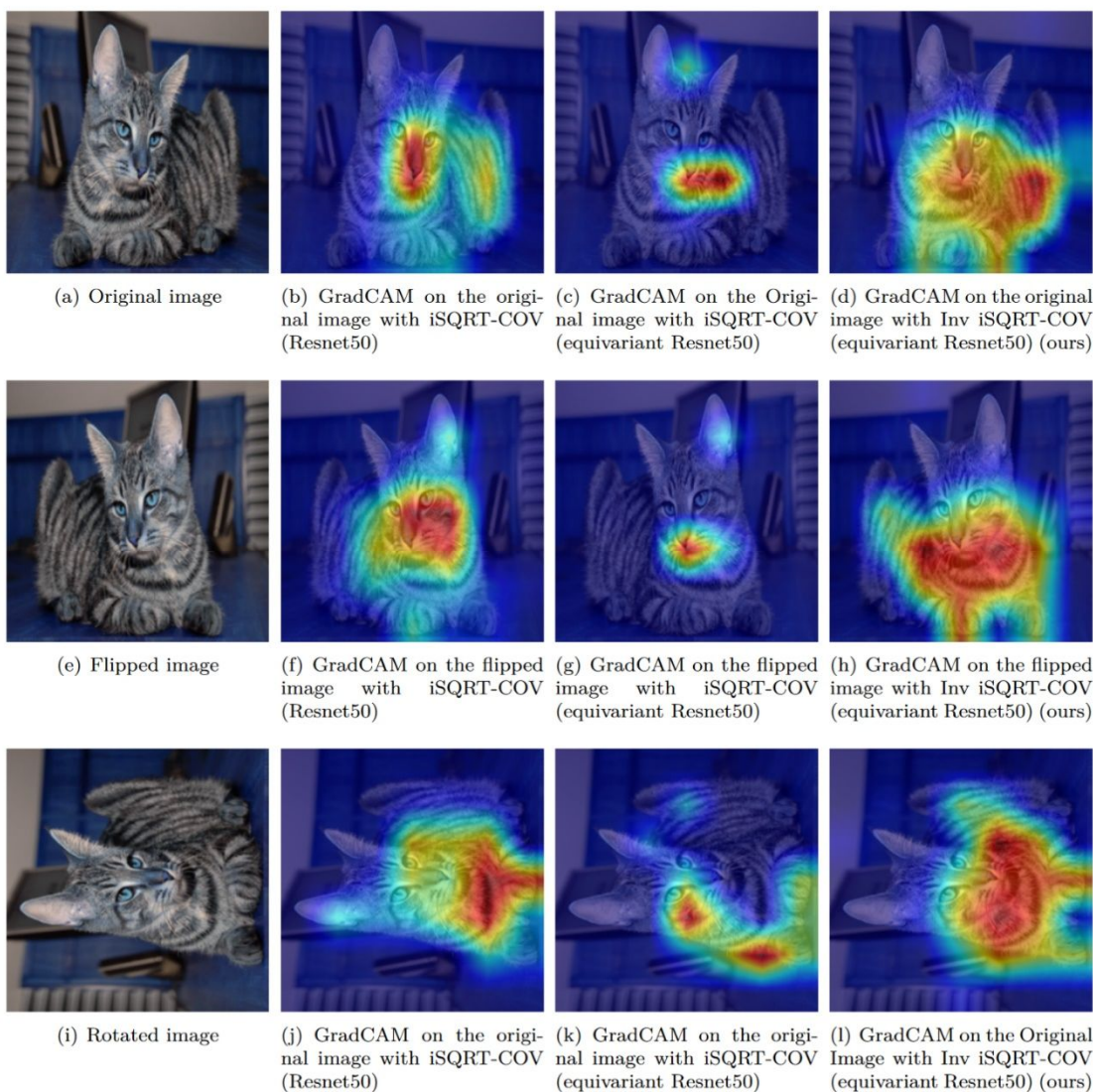


図 4 提案手法により獲得した特徴量の attention map。提案手法は入力に変換をかけても比較的 attention map の形状を保っている。

(1) Equivariant Neural Networks のモデル拡張に関して、既存コーディング手法である Bilinear Pooling, VLAD, VLAT といった手法に対してその共変な拡張手法を提案した。同時に不変な大域特徴と data augmentation のある種の等価性を示し、不変性の有用性を確かめた。標準的な画像認識データセットを用いた識別実験により提案手法により共変性を考慮しない時よりも精度よく識別が行えること、また実際にテスト画像に変換をかけてから識別した時に通常の手法は大きく精度を落とすのに対して提案手法では変換をかけた画像も上手く識別できる

ことを確かめた。これにより提案手法が実際に不変性を持っていること、不変性を活用することにより高性能な識別を行えることが分かった。また提案手法を end-to-end な Neural Networks の学習に組み込み、同様に高性能な識別が行えること、また可視化により提案手法により学習された Equivariant Neural Networks の着目している点が入力データに対して共変に近く変化している事を確かめた。

(2) Equivariant Neural Networks と self-supervised learning の統合に関して、既存手法である context prediction, solving jigsaw puzzle, Momentum Contrast, SwAV, SimSiam といったタスクに対する共変な拡張を提案した。大規模画像認識データセットである ImageNet を用いた実験により既存のタスクを共変なモデルに適用した時よりも提案する共変なタスクを共変なモデルに適用した方が高い認識性能を示すことを確かめた。また種々の不変性に対して提案手法を用いて評価した。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 2件 / うち国際学会 2件）

1. 発表者名 Yusuke Mukuta
2. 発表標題 Feature coding using invariance and kernel approximation
3. 学会等名 Workshop on Functional Inference and Machine Intelligence (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Yusuke Mukuta
2. 発表標題 On training and application of equivariant neural networks
3. 学会等名 International Workshop on Symbolic-Neural Learning (招待講演) (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>Yusuke Mukuta, Tatsuya Harada. Invariant Feature Coding using Tensor Product Representation. arXiv:1906.01857, 2019. Yusuke Mukuta, Tatsuya Harada. Self-Supervised Learning for Group Equivariant Neural Networks. arXiv:2303.04427, 2023.</p>

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

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
---------	---------