

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 17 日現在

機関番号：32689

研究種目：挑戦的萌芽研究

研究期間：2013～2014

課題番号：25540075

研究課題名(和文) 構造モデル学習による一般化性能強化

研究課題名(英文) Enhanced Generalizability by Structural Model Learning

研究代表者

石川 博 (Ishikawa, Hiroshi)

早稲田大学・理工学術院・教授

研究者番号：60381901

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：CNNにおいては、平行移動で移り合うような神経素子は同じ値になるように訓練し、画像認識においては必須である平行移動による普遍性を持つ特徴をデータから学習させることができる。この平行移動のような一般の変換について同様の効果をめざし、構造の代数的表現と、そのデータ空間におけるセマンティクスを一様に定義することにより、生のデータの中にパターンが存在するかどうかという質問に答えることができる理論の応用を目指した理論的研究を行った。また学習アルゴリズムの応用例として、CNNとsupport vector machine (SVM)によるランドサット衛星画像中の地物認識アルゴリズムを開発した比較した。

研究成果の概要(英文)：In the convolutional neural network (CNN), features with invariance under parallel translation, which is essential in image recognition tasks, can be learned by training the neurons that translate to each other by parallel translation together so that they have the same value. Aiming at similar effect in the case of general transformations, we conducted a theoretical research aimed at the application of the theory that can answer the question on the presence of patterns in raw data by uniformly defining algebraic representation of structures and their semantics in the data space. Also, as an example of application of learning algorithm, we compared algorithms for ground object recognition for Landsat images using CNN and support vector machine (SVM).

研究分野：コンピュータビジョン

キーワード：コンピュータビジョン

1. 研究開始当初の背景

現在増え続けている非記号データは通常、多数の数値の集まりとして記述されるが、そのような高次元情報は、各数値は個別の意味を持たず、その値の組み合わせ、しかも特定の組み合わせだけに情報があるという特徴を持つ。その結果、表現可能なデータの空間の巨大さのため、その中に埋め込まれた意味のある信号(パターン)はごく一部となる。このようなデータから有用な情報を引き出すためには、高次元の巨大な空間のごく一部のみ注目する必要がある。実際の信号は未知の拘束条件の元で生成されていて、それをかなりの程度知ることはなしに信号から有用な情報を引き出すことは望めない。この拘束条件は高次情報が生のデータとどう関係するかを表すモデルと考えることができる。

そのような特徴を持つ高次元情報の典型的な例は画像である。画像から有用な高次情報を引き出すことを目指すのがいわゆるパターン認識であり、そこでは、ツマミのついた機械のセッティングを見つけるような「学習」はするが、機械そのもの、つまりモデルは人手で設計され変化しない。そのため一般的なカテゴリに属する物体の認識、例えば犬と猫を識別させるなどということは困難であり、文字認識のような、対象カテゴリをデータレベル(この場合は画像)に近い形で直接的に持つことのできる場合とは対照的である。これは、データレベルから遠くなり認識したい対象が抽象的になるほど、人間の手でモデルを作ることが困難なためである。

一方、ごく最近、画像認識能力の飛躍的向上で注目されている、feature learningあるいはdeep learningと呼ばれる機械学習手法では、膨大な計算量によってある程度モデルを暗黙に学習するようである。しかし、それを明示的に記述することができないため、モデルを一般化することができない。

2. 研究の目的

表現可能なデータは実際に与えられ得る学習データより遙かに多い。少ない学習データから有用なモデルを得るためには、データが含まれる空間の特徴を利用して一般化を行う必要がある。この特徴は空間を特徴付ける写像の組み合わせであたえることができる。本研究の目的は、データの含まれる空間に特徴的な構造を利用してモデルを学習し、機械学習における一般化能力を強化することである。

3. 研究の方法

そのために、(1) 複数の空間、それらの直積、それらの空間を特徴付ける写像を計算機上に表現し、特定の構造を仮定しないように、また任意の構造を平等に扱えるようにするアーキテクチャの設計・実装をする。(2) その上で、与えられたデータのある確率分布からのサンプルと仮定し、そのデータの属する空間に自然な写像の組み合わせで得られる写像でその分布を送り、送り先のエントロピ

ーが減少する場合を探すアルゴリズムを開発するという方針をとった。

より具体的には、データの入っている空間の構造は、その空間を特徴付ける写像の組み合わせによって表現する。例えば画像平面のようなユークリッド空間であれば、2点間の距離を与える写像、2点間のベクトルを与える写像などである。この例からすぐに解くことは、それらの写像の像はまた別の空間に入っているということである。つまり、距離写像なら実数の空間、ベクトルを与える写像ならば、ベクトル空間である。するとそれらの空間を特徴付ける写像も組み合わせ、より多くの構造が存在することがわかる。例えばユークリッド空間に存在しない、ベクトルを実数倍するという写像がベクトル空間には存在する。また、上記の写像は二つとも、実際はユークリッド空間上の写像というよりは、ユークリッド空間2つの直積上の写像である。このような複数の空間、それらの直積、それらの空間を特徴付ける写像を計算機上に表現し、特定の構造を仮定しないように、また写像として与えられる任意の構造を平等に扱えるようにするアーキテクチャの設計・実装をするという方法をとった。

4. 研究成果

多層モデルである convolutional neural network(CNN)においては、通常、平行移動で移り合うような神経素子は同じ値になるように同時に訓練する。CNNによる認識システムでは、これによって画像認識においては必須である平行移動による普遍性を持つ特徴をデータから学習させることができる。平成25年度には、この平行移動にあたる、より一般の変換について同様の効果をめざし、構造の代数的表現、つまり計算と、そのデータ空間における意味(セマンティクス)を広い範囲で一樣に定義することにより、生のデータの中にパターンが存在するかどうかという質問に答えることができる理論の直接的な応用を目指した理論的研究を行った。

一方、CNNの学習性能を飛躍的に向上させたのは Restricted Boltzmann Machine を使ったことによるが、これは2値のマルコフ確率場である。このように多層モデルとマルコフ確率場は関連が深い。特に最近、高階のマルコフ確率場の研究が活発になっている。この高階マルコフ確率場モデルについて、高階2値エネルギーを1階エネルギーに還元する新たなアルゴリズムを開発した。既存手法では高階の2値エネルギーに還元するために変数を付加していたが、開発したアルゴリズムでは変数を加えずに還元することを可能にした。そのため結果として、既存手法より少ないメモリでより高速な還元を可能とするアルゴリズムの開発に成功し、成果は CVPR2014 で発表した。

平成26年度には、与えられたデータのある確率分布からのサンプルと仮定し、そのデータの属する空間に自然な写像の組み合わ

せで得られる写像でその分布を送り、送り先のエントロピーが減少する場合を探すアルゴリズムの開発・実装を行った。与えられたデータのエントロピーを特異的に減少させる写像を見つけるために、各空間で典型的な分布も同じ写像で送って検討した。一方、高階確率場のグラフ構造の設計にこの情報を使うことを検討した。

また学習アルゴリズムの応用例として、convolutional neural network(CNN) と support vector machine (SVM)によるランドサット衛星画像中の地物認識アルゴリズムを開発した比較した。

また、医用画像処理における高階確率場の応用において、データから確率場を学習することの検討を行った。特に医用画像のセグメンテーションにおける高階エネルギー最小化と学習を検討した。多臓器セグメンテーションは多ラベルのラベル付け問題と考えることができるが、高階の多ラベルラベル付け問題はセグメンテーション問題としてはまだあまり応用例がない。エネルギーを正解データから学習することにより作成するために、どのような統計を使うとセグメンテーションに効果があるのかは未知であり、検討を要した。6つの高階エネルギーを提案し、実験を行った。直角三角錐状に並んだクリークを用いた高階ポテンシャル、離れたボクセルを用いたクリークのポテンシャル、高階平滑化項、CT画像の輝度勾配を考慮して全臓器に作用する高階ポテンシャル、エッジ周辺にのみ作用する高階ポテンシャルでは、よいセグメンテーション結果を得られなかった。提案ポテンシャルを加えることで、セグメンテーション精度が悪くなる原因の一つとして、提案ポテンシャルが全体に一律に導入される場合には、確率アトラス項の邪魔になっていることが考えられる。よりよいポテンシャルを設計するには、アトラス項と共存できるポテンシャルを設計する必要がある。他方、CT画像の輝度勾配を考慮した高階ポテンシャルでは、すべての症例でセグメンテーション精度が向上するわけではないが、一部の症例では高階ポテンシャルの導入によりセグメンテーション精度の向上を確認することができた。症例間で高階ポテンシャルの導入によるセグメンテーション精度改善の有無が存在するので、今後は症例間でロバストな高階ポテンシャルを考える必要がある。

また、肺の血管のCT画像を動脈と静脈に分けるセグメンテーションは、動脈・静脈間の見た目の違いがほとんどないため難しいことが知られている。そのため、血管の根元の、動脈と静脈の分類が既知の部分からの連続性によって血管全体を分類する手法を試みた。その際、動脈と静脈が接触していたりすれ違っていたりする部位において、従来の1階エネルギーを使った手法では相互に侵食する問題があった。そこで高階エネルギーを使うことにより、肺血管が比較的まっすぐで

あるという性質をエネルギー中に表現することを可能にした。この際、高階項の重みを正解付きCTデータから学習することを提案した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計8件)

①M. Morita, A. Okagawa, Y. Oyamada, Y. Mochizuki, and H. Ishikawa, Multiple-Organ Segmentation Based on Spatially-Divided Neighboring Data Energy, The 14th IAPR Conference on Machine Vision Applications (MVA 2015), pp.158-161, 2015, 査読有.

②A. Okagawa, Y. Mochizuki, Y. Oyamada, and H. Ishikawa, Multi-Organ Segmentation by Minimization of Higher-Order Energy for CT Boundary, The 14th IAPR Conference on Machine Vision Applications (MVA 2015), pp.547-550, 2015, 査読有.

③T. Ishii, Y. Mochizuki, H. Nakada, R. Nakamura, and H. Ishikawa, Surface Object Recognition with CNN and SVM in Landsat 8 Images, The 14th IAPR Conference on Machine Vision Applications (MVA 2015), pp.341-344, 2015, 査読有.

④N. Kobayashi, Y. Oyamada, Y. Mochizuki, and H. Ishikawa, Three-DoF Pose Estimation of Asteroids by Appearance-based Linear Regression with Divided Parameter Space, The 14th IAPR Conference on Machine Vision Applications (MVA 2015), pp.551-554, 2015, 査読有.

⑤L. Prasuhn, Y. Oyamada, Y. Mochizuki, and H. Ishikawa, A HOG-Based Hand Gesture Recognition System on a Mobile Device, The IEEE International Conference on Image Processing (ICIP2014), pp. 3973-3977, 2014, 査読有.
DOI: 10.1109/ICIP.2014.7025807

⑥Y. Kitamura, Y. Li, W. Ito, H. Ishikawa, Coronary Lumen and Plaque Segmentation from CTA Using Higher-Order Shape Prior, The 17th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI2014), pp. 339-347, 2014, 査読有.
DOI: 10.1007/978-3-319-10404-1_43

⑦ H. Ishikawa, Higher-Order Clique Reduction Without Auxiliary Variables, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2014), pp. 1362-1369, 2014, 査読有.
DOI: 10.1109/CVPR.2014.177

⑧ Y. Kitamura, Y. Li, W. Ito, and H.

Ishikawa, Adaptive higher-order submodular potentials for pulmonary artery-vein segmentation, Fifth International Workshop on Pulmonary Image Analysis (PIA2013). pp.53-61, 2013, 査読有.

<http://www.lungworkshop.org/2013/styled-2/index.html>

〔学会発表〕(計 8 件)

①石川 博、最適化としての視覚と認識、第6回暗号フロンティアセミナー(招待講演)、北陸先端科学技術大学院大学(石川県能美市) 2015年3月18日

②H. Ishikawa, Higher-order Graph Cuts, ACCV2014 Area Chairs Workshop (招待講演), Nanyang Technological University, Singapore, 2014年9月3日.

③石川 博、グラフカット・その後、画像の認識・理解シンポジウム(MIRU2013). (招待講演)、国立情報学研究所(東京都)、2013年7月29日.

〔図書〕(計 0 件)

〔産業財産権〕

○出願状況(計 1 件)

名称：情報処理方法、情報処理装置およびそのプログラム

発明者：石川 博

権利者：早稲田大学

種類：特許

番号：特願 2014-121876

出願年月日：2014 年 6 月 12 日

国内外の別：国内

6. 研究組織

(1) 研究代表者

石川 博 (ISHIKAWA, Hiroshi)

早稲田大学・理工学術院・教授

研究者番号：60381901