

令和元年6月14日現在

機関番号：15501

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16116

研究課題名（和文）自律的な知識の発見と構造化を実現するホロニック構造型識別器の開発とその応用

研究課題名（英文）Building Classifiers with Holonic Structures for Automatic Knowledge Discovery and Structurization and Its Applications

研究代表者

間普 真吾（Mabu, Shingo）

山口大学・大学院創成科学研究科・准教授

研究者番号：70434321

交付決定額（研究期間全体）：（直接経費） 2,700,000円

研究成果の概要（和文）：本研究では、データから識別に有用な特徴や概念を抽出し、それらを適切に統合することで高い識別能力を有する人工知能を開発することを目的とした。まず、データから有用な知識を抽出するデータマイニング手法をベースに、抽出された知識を必要に応じて選択・利用するアンサンブル学習法の開発、少ない教師データでも識別器を構築可能な教師なし、半教師あり学習法の開発を行った。また、深層学習を用いてデータ識別に有用な特徴を抽出する手法を複数提案した。さらに、開発手法を医用画像（胸部CT画像）における陰影識別、肺聴診音における正常音と異常音の識別、人工衛星画像における災害領域検出システムに応用しその有用性を明らかにした。

研究成果の学術的意義や社会的意義

大量のデータが利用可能となった高度情報社会において、データから知識を発見し、人間と同等以上の判断能力を有する人工知能を開発することは今後の社会の発展のために重要である。本研究では、データマイニング手法や深層学習を用い、医用データにおける疾患の識別や、災害発生時に人工衛星画像から迅速に被災地域を発見できるシステムの構築を例に、識別に有用な特徴抽出法や、問題に適した識別器の構成を複数提案し、その性能を明らかにしている。

研究成果の概要（英文）：This research aimed to develop AI systems that can extract and combine useful features for data classification with high accuracy. First, based on data mining methods that extract useful knowledge from data, ensemble learning algorithms that select and use extracted knowledge appropriately according to the necessity were proposed, and unsupervised and semi-supervised learning algorithms that can build classifiers with a small number of labeled data were proposed. Second, some feature extraction methods that extract useful features for data classification using deep learning were proposed. Finally, the proposed methods were applied to opacity classification of medical images (chest CT images), normal and abnormal sounds classification of the lung, and disaster area classification of satellite images, then the effectiveness of the proposed methods were clarified.

研究分野：知能情報学

キーワード：機械学習 ニューラルネットワーク 進化論的計算手法 データマイニング パターン認識

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

大量のデータから知識を発見し、人間と同等以上の判断能力を有する人工知能を開発することは持続可能な社会の実現へ向けて重要である。これが実現できれば、例えば数百億台と言われる多種多様な機器がネットワーク化された世界で起こり得る異常の検知や予測システム、数万枚におよぶ医用画像から病変を発見する医用画像診断システムなどの確立に近づく。

2. 研究の目的

本研究は、大量のデータから高度な情報抽出・判断を行う人工知能の実現を目指し、人間に近い意思決定過程を持つホロニック構造型識別器の構築を目的とした。具体的には、大量のデータに潜む本質的な特徴や概念を獲得し、多くの概念を統合しながら最終判断を行うメカニズムの構築と、医用画像診断システムや人工衛星画像における災害領域検出システムの構築を通じて提案手法の有用性を検証することを目的とした。

3. 研究の方法

(1) ルールベース型識別器のアンサンブル学習法

様々な概念を状況に応じて組合せ、識別を行う手法の研究を行った。まず、遺伝的ネットワークプログラミング (Genetic Network Programming, GNP) と呼ばれる進化論的計算手法を用いて、クラス相関ルールを抽出する。その後、データ識別を行う際、すべてのルールを使用するのではなく、識別すべきデータの特徴に応じて使用するルールを多層パーセプトロンによって効果的に選択する。

(2) 半教師ありデータマイニング法

少数の教師ありデータと大量の教師なしデータを効率的に学習に利用することで、識別率の向上を図る方式である。一般に、教師データが少数であるとき、データマイニングによって信頼性の高いルール抽出を十分に行うことができない。そこで、大量の教師なしデータから候補ルールを抽出し、さらに少数の教師ありデータによって候補ルールの信頼度を計算する方式とした。

(3) 教師なし特徴抽出・画像分類法

様々な画像の分類に必要な特徴量を自動的に獲得し、それを用いてクラスタリングを行う手法を提案した。具体的には、深層自己符号化器および Bag-of-features 法を用いて特徴を抽出し、抽出された特徴に対して K-means クラスタリングを適用する方式である。本方式は医用画像における陰影分類システムに応用した。

(4) 半教師あり画像分類法

(3)の特徴抽出法に加えて、自己学習と能動学習を行うことで、少数の訓練データから効率的な学習が可能な仕組みとした。本方式も医用画像における陰影分類システムに応用した。

(5) 人工衛星画像に対する異常検知システムの構築

自然災害発生時に、人工衛星画像から災害発生領域を検出する方式の開発を行った。通常、複数クラス (正常、異常など) に分類する識別器を構築するには、複数クラスそれぞれについて訓練データが必要である。衛星画像から災害発生地域を検出するには、正常と異常の画像が大量に必要である。しかし、正常画像と比較して異常画像を大量に取得するのは困難であり、取得できる画像数に差がある。したがって、正常画像のみを用いて特徴量の正常領域を求め、正常領域から外れた画像を異常と判断する識別器を構築した。本方式の実現のため、畳込み自己符号化器 (Convolutional Autoencoder, CAE) による特徴量抽出と、One-Class Support Vector Machine (One-Class SVM) による異常検知手法を組合せた。

(6) マルチチャネルデータを用いた識別器の構築

衛星画像中のある領域で災害が発生しているかどうかを判断するには、判断する領域のみならず、周辺の状況も考慮することが重要である。したがって、識別対象領域と周辺領域の画像をまとめてマルチチャネル化し、畳込みニューラルネットワーク (Convolutional Neural Network, CNN) で識別する手法を提案した。複数の情報をまとめて入力することで、CNN が重要な概念を自動抽出し、それらを組み合わせて判断する仕組みとなることが期待できる。また、識別対象領域と周辺領域の画像を組合せる際、画像サイズを一致させる必要があるが、ゼロパディング法とマップ合成法と呼ばれる2つの方式を提案した。

(7) 胸部 CT 画像におけるアンサンブル学習型識別器の構築

胸部 CT 画像におけるびまん性肺疾患の陰影識別器を構築した。(3)で使用した深層自己符号化器は特徴抽出を教師なしで行うことが可能であるものの、陰影クラスごとの違いを明確に表現できず、識別率が向上しないケースも見られた。そこで、畳込み敵対的自己符号化器 (Convolutional Adversarial Autoencoder CAE) と呼ばれる教師なし特徴抽出法を用い、さらに、識別対象の関心領域のみならず、CT 画像全体の特徴も統合し判断を行うアンサンブル

学習型の識別器とした。

(8) 事前学習を用いた肺聴診音識別器の構築

肺聴診音において、正常音と異常音を識別するシステムを構築した。従来の音声認識研究でしばしば用いられるメル周波数ケプストラム係数 (Mel Frequency Cepstrum Coefficient, MFCC) を特徴量として抽出し、さらに畳込みニューラルネットワーク (CNN)、または長短期記憶ネットワーク (Long Short-Term Memory, LSTM) を利用することで、音声認識に重要な特徴をさらに抽出する方式とした。CNN および LSTM には事前学習と呼ばれる教師なし学習を利用することで、少ない音声データ数でも識別に有用な概念を獲得できる方式とした。

(9) 畳込み自己符号化器つき U-net による人工衛星画像分析システム

U-net と呼ばれる深層学習アルゴリズムをベースとし、これを畳込み自己符号化器と融合することで災害領域の検出精度向上を図った。U-net は領域抽出画像を出力するものだが、同時に入力画像の再構成画像も出力するハイブリッドな構造とすることで、多様な特徴を抽出することが可能となった。

4. 研究成果

(1) ルールベース型識別器のアンサンブル学習法 (提案手法) による識別性能を、アンサンブル学習なしの方式、および代表的なアンサンブル学習法であるランダムフォレストと比較し、提案手法が優れていることを明らかにした。評価用データは、UCI Machine Learning Repository から取得した。

(2) 教師データ数が少ない場合において、半教師あり学習で構築されたルールベースに基づく識別器と、教師あり学習で構築されたルールベースに基づく識別器を組み合わせることで識別率が向上することを明らかにした。教師あり学習と半教師あり学習では、異なる観点でルール抽出が行われるため、様々な状況に対応できるルールが獲得できたと考えられる。評価用データは、(1)と同様に UCI Machine Learning Repository から取得した。

(3) 胸部 CT 画像における陰影分類問題において、教師データを必要としない分類システムを構築することができた。従来の特徴抽出法である Histogram of Oriented Gradients (HOG) を用いた識別器と性能を比較し、提案手法の有用性を明らかにした。

(4) 自己学習と能動学習を行うことによって、少数の教師データから効率よく重要な情報を抽出することが可能となり、その結果識別率が向上することがわかった。本手法は、大規模な教師ラベルつきデータベースを構築する場合にも応用が可能である。

(5) 正常画像と異常画像の分類において、正常画像のみで識別器の構築が可能であることを明らかにした。また、畳込み自己符号化器による特徴抽出の有無による、感度・特異度の比較も行い、畳込み自己符号化器の有用性を明らかにした。

(6) 識別対象領域と周辺領域の画像を組合せてマルチチャネル化する方式が有用であること示した。またゼロパディング法とマップ合成法の比較も行い、マップ合成法が優れていることがわかった。

(7) 畳込み自己符号化器 (CAE) と畳込み敵対的自己符号化器 (CAAE) を比較し、CAAE によって抽出された特徴量を用いた陰影識別の方が優れていることがわかった。また、陰影識別対象領域と CT 画像全体の特徴量を統合することで識別率の向上が図れることを明らかにした。

(8) CNN, LSTM の両方について、事前学習が肺聴診音の識別率向上に有用であることがわかった。多層パーセプトロン (Multi-Layer Perceptron, MLP) との比較、事前学習の有無による比較、メル周波数ケプストラム係数 (MFCC) の抽出次元数を変更した場合の比較、MFCC をベースとした動的特徴量使用の有無による比較も行ったところ、CNN において事前学習有り、MFCC40 次元、動的特徴量有りの場合において最も高い識別率を示した。

(9) 畳込み自己符号化器つき U-net は、災害領域を抽出するのみならず、入力画像も再構成できるよう構成することで、U-net と比較して高い検出率を示した。また、災害領域抽出には、災害前後の画像に加え、標高データや土地利用データも有用であることがわかった。

5. 主な発表論文等

[雑誌論文] (計 5 件)

1. Shingo Mabu, Shun Gotoh, Masanao Obayashi, Takashi Kuremoto, A random-forests-based classifier using class association rules and its application to an intrusion detection system, *Artificial Life and Robotics*, Springer, Vol. 21, Issue 3, pp. 371-377, 2016, DOI: 10.1007/s10015-016-0281
2. Shingo Mabu, Masanao Obayashi and Takashi Kuremoto, An Evolutionary Algorithm for Making Decision Graphs for Classification Problems, *Journal of Robotics, Networking and Artificial Life*, Vol. 3, No. 1, pp. 45-49, 2016, DOI: 10.2991/jrnal.2016.3.1.11
3. Shingo Mabu, Masanao Obayashi, Takashi Kuremoto, Noriaki Hashimoto, Yasushi Hirano, Shoji Kido, Unsupervised Class Labeling of Diffuse Lung Diseases Using Frequent Attribute Patterns, *International Journal of Computer Assisted Radiology and Surgery*, Vol. 12, Issue 3, pp. 519-528, 2017, DOI: 10.1007/s11548-016-1476-2
4. Shingo Mabu, Kenzoh Azakami, Masanao Obayashi and Takashi Kuremoto, A Rule-Based Classification System Enhanced by Multi-Objective Genetic Algorithm, *Journal of Robotics, Networking and Artificial Life*, Vol. 4, No. 3, pp. 239-242, 2017
5. Shingo Mabu, Kyoichiro Kobayashi, Masanao Obayashi and Takashi Kuremoto, Unsupervised Image Classification Using Multi-Autoencoder and K-means++, *Journal of Robotics, Networking and Artificial Life*, Vol. 5, No. 1, pp. 75–78, 2018, DOI: 10.2991/jrnal.2018.5.1.17

[学会発表](計 30 件)

1. Shingo Mabu, Masanao Obayashi, Takashi Kuremoto, Noriaki Hashimoto, Yasushi Hirano, Shoji Kido, Unsupervised Class Labeling of Diffuse Lung Diseases Using Frequent Attribute Patterns, *Computer Assisted Radiology and Surgery (CARS2017)*, 2017
2. Shingo Mabu, An Unsupervised Classification Model of Lung Diseases, the Third International Symposium on the Project "Multidisciplinary Computational Anatomy", 2017
3. Kenzoh Azakami, Shingo Mabu, Masanao Obayashi, Takashi Kuremoto, A Rule-Based Classification System Enhanced by Multi-Objective Genetic Algorithm, *The 2017 International Conference on Artificial Life and Robotics*, 2017
4. Shingo Mabu, Masanao Obayashi, Takashi Kuremoto, Noriaki Hashimoto and Yasushi Hirano, Unsupervised Opacity Annotation of Diffuse Lung Diseases Using Deep Autoencoder and Bag-of-Feature, *The 4th World Congress of Thoracic Imaging*, 2017
5. Shingo Mabu, Noriaki Hashimoto, Shoji Kido, Masanao Obayashi, Takashi Kuremoto and Yasushi Hirano, Semi-Supervised Learning for Automatic Annotation of Diffuse Lung Diseases Using Deep Autoencoder and SVM, *Computer Assisted Radiology and Surgery, 32nd International Congress and Exhibition*, 2017
6. Shingo Mabu, Kyoichiro Kobayashi, Masanao Obayashi and Takashi Kuremoto, Unsupervised Image Classification Using Multi-Autoencoder and K-means++, *The 2018 International Conference on Artificial Life and Robotics*, 2018
7. Kohki Fujita, Shingo Mabu and Takashi Kuremoto, Anomaly Detection of Disaster Areas from Satellite Images Using Convolutional Autoencoder and One-class SVM, *The 2018 International Conference on Artificial Life and Robotics*, 2018
8. S. Mabu, S. Kido, N. Hashimoto, Y. Hirano and T. Kuremoto, Opacity annotation of diffuse lung diseases using deep convolutional neural network with multi-channel information, *SPIE Medical Imaging 2018*, 2018
9. Shingo Mabu, Shoji Kido, Yasushi Hirano, Takashi Kuremoto, Unsupervised and Semi-Supervised Learning for Efficient Opacity Annotation of Diffuse Lung Diseases, *International Forum on Medical Imaging in Asia 2019*, 2019
10. Ami Atsumo, Shingo Mabu, Shoji Kido, Yasushi Hirano, Takashi Kuremoto, Analysis of the effects of transfer learning on opacity classification of diffuse lung diseases using convolutional neural network, *International Forum on Medical Imaging in Asia 2019*, 2019

11. Risa Takemiya, Shoji Kido, Yasushi Hirano, Shingo Mabu, Detection of pulmonary nodules on chest X-ray images using R-CNN, International Forum on Medical Imaging in Asia 2019, 2019
12. Unsupervised Class Labeling of Diffuse Lung Diseases Using Evolutionary Data Mining and K-means, 間普真吾, 大林正直, 吳本堯, 橋本典明, 平野靖, 木戸尚治, 第 35 回日本医用画像工学会大会, 2016
13. 中山良彬, 間普真吾, 大林正直, 吳本堯, 多層パーセプトロンを用いたルールベース型識別器のアンサンブル学習, 平成 28 年度 (第 67 回) 電気・情報関連学会中国支部連合大会, 2016
14. 樋口拓郎, 間普真吾, 大林正直, 吳本堯, 半教師ありデータマイニングによる識別器の構築, 平成 28 年度 (第 67 回) 電気・情報関連学会中国支部連合大会, 2016
15. 皆上建三, 間普真吾, 大林正直, 吳本堯, クラス相関ルールのクラスタリングによる識別システムの改良とその性能評価, 平成 28 年度 (第 67 回) 電気・情報関連学会中国支部連合大会, 2016
16. 樋口拓郎, 間普真吾, 大林正直, 吳本堯, 半教師ありデータマイニングを用いたアンサンブル学習型識別器の構築, 平成 29 年度 (第 68 回) 電気・情報関連学会中国支部連合大会, 2017
17. 間普真吾, 機械学習による大規模データ分析に関する応用事例, IEEE SMC Hiroshima Chapter 主催 Special Lecture, 2017 (招待講演)
18. 中山良彬, 間普真吾, 吳本堯, マルチチャネル情報を利用した畳込みニューラルネットワークによる SAR 画像解析, 第 59 回計測自動制御学会システム工学会研究会, 2018
19. 間普真吾, AI 学習の効率化に関する研究事例, 第 4 回呼吸機能イメージング研究会サマーセミナー, 2018 (招待講演)
20. 間普真吾, 木戸尚治, 橋本典明, 平野靖, 吳本堯, 医用画像の機械学習, 第 57 回日本生体医工学会大会, 2018 (招待講演)
21. 間普真吾, Deep learning and radiomics: From the viewpoint of machine learning theories, 第 57 回日本生体医工学会大会, 2018 (招待講演)
22. 植田大介, 間普真吾, 吳本堯, 災害前後 SAR 画像と DEM データを用いた CNN による土砂災害検出, 第 32 回人工知能学会全国大会, 2018
23. 村上佳菜子, 橋本典明, 木戸尚治, 平野靖, 間普真吾, 近藤堅司, 小澤順, CNN, FCN, U-Net を用いたびまん性肺疾患の領域抽出の比較, 第 32 回人工知能学会全国大会, 2018
24. 樋口拓郎, 間普真吾, 橋本典明, 木戸尚治, 平野靖, 近藤堅司, 小澤順, 畳込みニューラルネットワークと SVM を用いたびまん性肺疾患の陰影識別, 第 32 回人工知能学会全国大会, 2018
25. 若本亮佑, 間普真吾, 木戸尚治, 吳本堯, 平野靖, 河村武郎, 畳込みニューラルネットワークを用いた肺聴診音データの識別, 計測自動制御学会システム・情報部門学術講演会 2018, 2018
26. 樋口拓郎, 間普真吾, 大林正直, 吳本堯, 遺伝的ネットワークプログラミングを用いた半教師ありデータマイニングによる識別器の構築, 計測自動制御学会システム・情報部門学術講演会 2018, 2018
27. 平田宗一郎, 間普真吾, 宮腰穂, 原田真悟, 矢野康英, 遺伝的ネットワークプログラミングを用いたハイブリッド車両走行データ分析, 計測自動制御学会システム・情報部門学術講演会 2018, 2018
28. 植田大介, 間普真吾, 吳本堯, SAR 画像および土地情報を用いた深層学習による土砂災害地域検出, 計測自動制御学会システム・情報部門学術講演会 2018, 2018
29. 間普真吾, 若本亮佑, 木戸尚治, 吳本堯, 平野靖, 河村武郎, 深層学習を用いた呼吸音識別器の構築, 第 11 回呼吸機能イメージング研究会学術集会, 2019
30. 植田大介, 間普真吾, 吳本堯, SAR 画像と土地情報を用いた畳込み自己符号化器付 U-net による土砂災害地域検出, 第 61 回計測自動制御学会システム工学会研究会, 2019

〔図書〕(計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号：
出願年：
国内外の別：

○取得状況（計 0 件）

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

(1)研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号（8桁）：

(2)研究協力者

研究協力者氏名：

ローマ字氏名：

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。