

令和 4 年 6 月 15 日現在

機関番号：25406

研究種目：研究活動スタート支援

研究期間：2019～2021

課題番号：19K24365

研究課題名（和文）動画ビッグデータの構造適応型マルチモーダル深層学習の研究

研究課題名（英文）A study of multi-modal adaptive structural deep learning for big data

研究代表者

鎌田 真（Kamada, Shin）

県立広島大学・公立大学の部局等（広島キャンパス）・講師

研究者番号：30845178

交付決定額（研究期間全体）：（直接経費） 2,200,000円

研究成果の概要（和文）：入力データに対して、最適と考えられる隠れニューロン数や層の数を学習しながら自己組織的に求める構造適応型深層学習法を開発した。本手法は画像分類等の複数のベンチマークテストに対し、他の深層学習手法よりも高い精度を示した。本研究では、複数のモダリティからの情報を統合的に処理し、それらの関連づけを実現するマルチモーダル構造適応型深層学習法について探求した。複数のモダリティを含む処理を統合するため、Teacher-Studentモデルの概念に基づいたCo-learningモデルを探索し、複数のモダリティ間の知識の融合モデルを開発した。提案手法を動画等のビッグデータに適用し評価し、高い性能が得られた。

研究成果の学術的意義や社会的意義

深層学習が登場して以降、画像認識に対する性能は従来に比べて飛躍的に向上した。しかしながら、医療データや人の主観が含まれるようなデータに対しては誤判定のケースが多く、これは、単一のデータのみでは判定が困難であるからである。例えば、医療データには画像の他に血液検査や問診等の結果があり、人間の医師はこれらの複数の情報を統合的に処理し、関連付けを行い、思考した上で最終的な判定をされると考えられる。このため、本研究のように、複数のモダリティを統合的に扱い、関係性を考慮した上で予測を行う学習システムは必要であり、人工知能の研究がさらなる進化を遂げると考えた。

研究成果の概要（英文）：The adaptive structural deep learning was developed, which can self-organize the suitable network structure by the generation algorithm of hidden neurons and layers for given input data. The method achieved higher classification accuracy than the existing deep learning models for several image benchmark test. In this research, the multi-modal adaptive structural deep learning was studied, which represents the multiple features of input data by several models and aligns them. The co-learning with Teacher-Student(TS) model was developed to transform the trained knowledge in the model to another model. The proposed model showed higher performance than the previous models for some big data sets including video data.

研究分野：計算知能

キーワード：深層学習 マルチモーダル 構造適応型学習 動画ビッグデータ

1. 研究開始当初の背景

近時、深層学習が人工知能の分野で注目を集めている。VGG16, ResNet 等は畳み込みニューラルネットワーク(CNN: Convolutional Neural Network)と呼ばれる深層学習法として知られ、特に画像認識で高い性能を持つ。しかし、高性能な CNN モデルの構築は、最適なネットワーク構造や様々なパラメタ設定について試行錯誤的な探索が必要であり、新たに構築することは容易ではない。このため、研究代表者らは入力データの特徴に対して最適なネットワーク構造(ニューロンや層の数)を学習中に自動で求める構造適応型深層学習法を開発した(図1)。具体的には、RBM(Restricted Boltzmann Machine)を階層的に積み上げたDBN(Deep Belief Network)において、RBMの最適なニューロン数を学習中に決定する構造適応型RBM(Adaptive RBM)と、事前学習した Adaptive RBM を積み上げることで、自動で階層構造を構築する構造適応型 DBN(Adaptive DBN)を開発した[1]。ベンチマーク等で既存の CNN 手法よりも高い性能を示し、画像分類、時系列データ予測、物体検出の面で有効性を示した。

構造適応型学習法は、画像や数値で表現される計測データ等に対しては高い性能を示したが、曖昧な情報を含む医療画像や顔感情画像の一部のクラスでは、80%の精度であった。これらは、全ての特徴や文脈を表現できないケースが多く含まれていることが原因であった。一般的に、人間は医師の診断や感情推定のように、複数のモダリティから得られた情報を統合的に処理し、関連付けを行い、思考した上で最終的な行動をすと考えられる。このため、人間が行うように、複数のモダリティ(5感など)を統合的に扱い、関係性を考慮した上で予測を行う学習システムが必要である。それぞれのモダリティに関する研究は独自に進化しているが、深層学習によってマルチモーダルな統合モデルの構築が可能となり、本研究によって、人工知能の研究がさらなる進化を遂げると考えた。

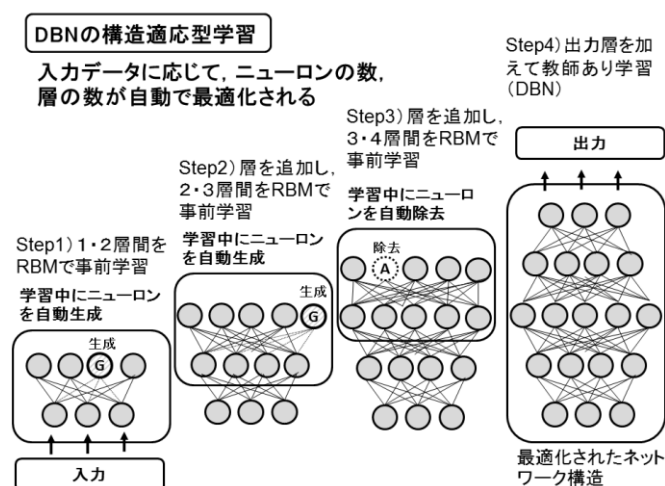


図1 : Adaptive DBN

2. 研究の目的

本研究では、開発した構造適応型深層学習法において、複数のモダリティからの情報を統合的に処理し、それらの関連づけを行う「マルチモーダル構造適応型深層学習法」について探求する。複数のモダリティを含む処理のため、Teacher-Student (T/S) モデルの概念に基づいた Co-learning モデルを提案し、複数のモダリティ間の知識の融合モデルを研究する。提案手法を複数のビッグデータに適用し評価する。

3. 研究の方法

マルチモーダル深層学習には、文献[2]によれば、複数のモダリティの表現方法の Representation, 変換方法の Translation, 関連付けの Alignment, 融合方法の Fusion, 知識の転移の Co-learning が必要とされる。本研究では、構造適応型深層学習において、ネットワークの内部の入出力パターンに関する知識を木構造の IF-THEN ルールとして抽出する手法や、画像と数値が混在したデータを一度に取り扱い、分類精度を下げることなく高速に学習する構造適応型学習法を既に提案している[3, 4]。

本研究では、これらの手法を改良し、また複数のモダリティを含む処理を統合するため、Teacher-Student (TS) モデルの概念に基づいた Co-learning モデルを探求する。ここで、TS モデルとは、構造適応型深層学習において、アンサンブル学習のように、元の親モデルで学習できなかったデータに対して、複数の子モデルを生成し、データの多様な特徴を表現可能なモデルである。文献[5]では、8種類の人間の感情と顔画像の組が与えられている顔感情画像データベース Affect Net に対して、TS モデルが適用されている。ここでは、感情のラベルが複数のアノテーターの主観によって定めることから誤分類が生じていることに着目し、TS モデルを用いて、複数のアノテーターの基準や事前知識を、複数のモデルを用いて表現している。

本研究では、TS モデルにおいて、各子モデルが学習した特徴を元の親モデルに組み込む知識の転移の実現のために、まず、各子モデルと親モデルの確率分布の違いを KL 情報量で測定した。次に、モデルを木構造とみなし、入力層から出力層に至るまでの各層のニューロンの発火パタン

(パス)を分析することで、各モデルの知識を表現し、類似性や生起頻度により各モデル間の関連性や変換を定義した。さらに、これらの特徴に基づき、子モデルのニューロンを知識として親モデルに転移させる方法について研究し、親モデルの精度向上を試みた。開発した手法を医療データや動画ビッグデータに適用し評価した。

4. 研究成果

本研究では、医療データとして、軽度認知症(MCI)判定のためのオープンデータである ADNI[6]を用いて、手法の評価を行った。本論文では、MRI の T1 強調画像で、Axial 軸(軸位断)の画像を CN, MCI, AD の 3 種類のカテゴリに分類するモデルを構築した。図 2 は画像のサンプルを示している。表 1 は学習後の分類性能を示している。従来手法(通常の Adaptive DBN)では、分類精度が 87.6%であったが、提案手法により、92.3%に向上した。

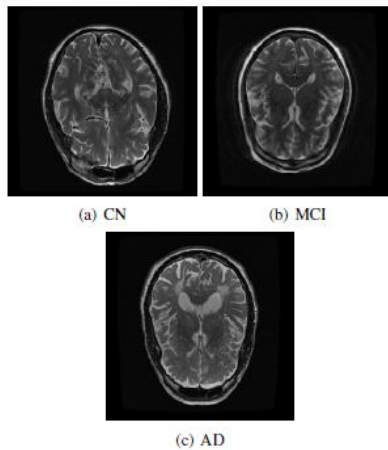


図 2: MRI 画像のサンプル

表 1: 分類性能

	分類精度
従来手法	87.6%
提案手法	92.3%

また、画像以外の多様なデータに対しても深層学習の適用が期待されていることから、Moving MNIST[7]と呼ばれるベンチマークデータを用いて、動画ビッグデータへの適用を試みた。Moving MNIST は、図 3 のように、一定の長さの間、2つの数字が画像内をランダムに動く動画ベンチマークで、与えられた系列から、次の数字の動きを予測する学習法を開発した。学習の結果、次の動きを 90%以上の精度で予測が可能となり、既存の手法よりも高い予測性能を示した。

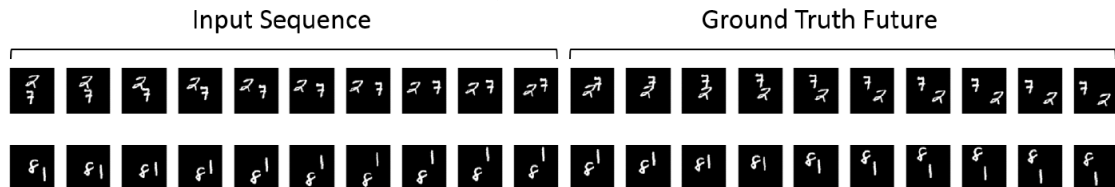


図 3 Moving MNIST

(参考文献)

- [1] S.Kamada, T.Ichimura, A.Hara, and K.J.Mackin, "Adaptive Structure Learning Method of Deep Belief Network using Neuron Generation-Annihilation and Layer Generation", Neural Computing and Applications, doi.org/10.1007/s00521-018-3622-y, pp.1-15 (2018)
- [2] T.Baltrušaitis et al, "Multimodal Machine Learning: A Survey and Taxonomy", https://arxiv.org/abs/1705.09406v1(2017)
- [3] S.Kamada, T.Ichimura, and T.Harada, "Knowledge Extraction of Adaptive Structural Learning of Deep Belief Network for Medical Examination Data", International Journal of Semantic Computing (IJSC), Vol.13, No.01, pp.67-86 (2019)
- [4] S.Kamada and T.Ichimura, "Fast Training of Adaptive Structural Learning Method of Deep Learning for Multi Modal Data", International Journal of Computational Intelligence Studies, Vol.7, No.3/4, pp.169-191 (2018)
- [5] T.Ichimura, S.Kamada, "Re-learning of Child Model for Misclassified data by using KL Divergence in AffectNet: A Database for Facial Expression", Proc. of 2019 IEEE 11th International Workshop on Computational Intelligence and Applications (IWCIA 2019), pp.15-20 (2019)
- [6] Alzheimer's Disease Neuroimaging Initiative, http://adni.loni.usc.edu/ (2022/6/10)
- [7] Moving MNIST, http://www.cs.toronto.edu/~nitish/unsupervised_video/ (2022/6/10)

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 Takumi Ichimura, Shin Kamada	4. 巻 Vol.6, No.1
2. 論文標題 An Ensemble Learning Method of Adaptive Structural Deep Belief Network for AffectNet	5. 発行年 2022年
3. 雑誌名 IIAI International Journal Series, International Journal of Smart Computing and Artificial Intelligence	6. 最初と最後の頁 pp.1-17
掲載論文のDOI（デジタルオブジェクト識別子） 10.52731/ijscai.v6.i1.640	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Shin Kamada, Takumi Ichimura	4. 巻 Vol.10, No.2/3
2. 論文標題 A Video Prediction Method by using Long Short Term Memory based Adaptive Structural Learning of Deep Belief Network and its Investigation of Input Sequence Length for Data Structure	5. 発行年 2021年
3. 雑誌名 International Journal of Computational Intelligence Studies	6. 最初と最後の頁 pp.198-215
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計24件（うち招待講演 0件／うち国際学会 8件）

1. 発表者名 Takumi Ichimura, Shin Kamada
2. 発表標題 A Distillation Learning Model of Adaptive Structural Deep Belief Network for AffectNet: Facial Expression Image Database
3. 学会等名 Proc. of 2020 9th International Congress on Advanced Applied Informatics (IIAI-AAI), pp.454-459 (2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Shin Kamada and Takumi Ichimura
2. 発表標題 An Embedded System for Image-based Crack Detection by using Fine-Tuning model of Adaptive Structural Learning of Deep Belief Network
3. 学会等名 Proc. of IEEE TENCON 2020, pp.1203-1208 (2020) (国際学会)
4. 発表年 2020年

1. 発表者名 寒竹夏水, 鎌田真, 市村匠
2. 発表標題 航空写真を用いた深層学習による道路網認識手法RoadTracerの動作パラメタ実験
3. 学会等名 2020 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.24-28 (2020)
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 組み込みPCを用いたコンクリート構造物のひび割れ検出構造適応型深層学習システムの高速化
3. 学会等名 2020 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.47-50 (2020)
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 構造適応型深層学習の再学習法を用いたAffectNet顔表情画像の分析
3. 学会等名 2020 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.51-55 (2020)
4. 発表年 2020年

1. 発表者名 市村匠, 鎌田真
2. 発表標題 構造適応型深層学習による親族間の顔画像データベースの学習
3. 学会等名 2020 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.56-59 (2020)
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠, 原田俊英
2. 発表標題 認知症早期判定のための構造適応型深層学習によるMRI/PET画像の分類
3. 学会等名 計測自動制御学会システム・情報部門学術講演会2020講演論文集 (SSI2020), pp.119-123 (2020)
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠, 岩崎貴志
2. 発表標題 コンクリート構造物のひび割れ検出構造適応型深層学習システムの開発～広島県橋梁・砂防ダム・湾岸施設の検査システム～
3. 学会等名 計測自動制御学会システム・情報部門学術講演会2020講演論文集 (SSI2020), pp.150-155 (2020)
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 構造適応型深層学習を用いた道路網認識手法RoadTracerへの適用
3. 学会等名 計測自動制御学会第17回コンピューテーショナル・インテリジェンス研究会, pp.68-72 (2021)
4. 発表年 2021年

1. 発表者名 Shin Kamada, Takumi Ichimura
2. 発表標題 A Video Recognition Method by using Adaptive Structural Learning of Long Short Term Memory based Deep Belief Network
3. 学会等名 Proc. of 2019 IEEE 11th International Workshop on Computational Intelligence and Applications (IWCIA 2019), pp.21-26 (国際学会)
4. 発表年 2019年

1. 発表者名 Takumi Ichimura, Shin Kamada
2. 発表標題 Re-learning of Child Model for Misclassified data by using KL Divergence in AffectNet: A Database for Facial Expression
3. 学会等名 Proc. of 2019 IEEE 11th International Workshop on Computational Intelligence and Applications (IWCIA 2019), pp.15-20 (国際学会)
4. 発表年 2019年

1. 発表者名 Shin Kamada, Takumi Ichimura, Takashi Iwasaki
2. 発表標題 An Adaptive Structural Learning of Deep Belief Network for Image-based Crack Detection in Concrete Structures Using SDNET2018
3. 学会等名 Proc. of 2020 International Conference on Image Processing and Robotics (ICIPRoB 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 SDNET2018: コンクリート構造物画像データセットを用いたひび割れ検出構造適応型深層学習システム
3. 学会等名 計測自動制御学会システム・情報部門学術講演会2019講演論文集 (SSI2019), pp.315-320
4. 発表年 2019年

1. 発表者名 市村匠, 鎌田真, 多羅尾直, 山口亮
2. 発表標題 構造適応型DBNの物体検出法におけるパラメタ自動調整と建築設備図面における図面記号への適用
3. 学会等名 計測自動制御学会第16回コンピューテーショナル・インテリジェンス研究会, pp.7-12
4. 発表年 2019年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 KL情報量に基づく学習後の構造適応型DBNにおけるニューロン付加手法の提案と顔表情画像への適用
3. 学会等名 計測自動制御学会第16回コンピュータショナル・インテリジェンス研究会, pp.1-6
4. 発表年 2019年

1. 発表者名 市村匠, 鎌田真, 石川ますみ, 小川諒, 河野利治
2. 発表標題 構造適応型DBNによる眼底画像データセットOptosの分類
3. 学会等名 計測自動制御学会第47回知能システムシンポジウム, C-1-1, pp.1-4
4. 発表年 2020年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 組み込みPCによるコンクリートひび割れ検出構造適応型深層学習システムの開発
3. 学会等名 計測自動制御学会第47回知能システムシンポジウム, C-3-2, pp.1-4
4. 発表年 2020年

1. 発表者名 Shin Kamada, Takumi Ichimura, Toshihide Harada
2. 発表標題 Image based Early Detection of Alzheimer's Disease by using Adaptive Structural Deep Learning
3. 学会等名 Proc. of KES-IDT21, Intelligent Decision Technologies, Smart Innovation, Systems and Technologies, vol 238, pp.595-605 (国際学会)
4. 発表年 2021年

1. 発表者名 Takumi Ichimura, Shin Kamada
2. 発表標題 Adaptive Structural Deep Learning to Recognize Kinship using Families in Wild Multimedia
3. 学会等名 Proc. of KES-IDT21, Intelligent Decision Technologies, Smart Innovation, Systems and Technologies, vol 238, pp.559-568 (国際学会)
4. 発表年 2021年

1. 発表者名 Shin Kamada, Takumi Ichimura
2. 発表標題 Automatic Extraction of Road Networks from Satellite Images by using Adaptive Structural Deep Belief Network
3. 学会等名 Proc. of 10th International Conference on Informatics, Electronics & Vision (ICIEV 2021), paper 37 (国際学会)
4. 発表年 2021年

1. 発表者名 市村匠, 鎌田真
2. 発表標題 構造適応型深層学習法によるRoadTracerの道路検出に対する一考察
3. 学会等名 2021 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.16-23
4. 発表年 2021年

1. 発表者名 鎌田真, 市村匠, 原田俊英
2. 発表標題 Teacher-Student構造適応型深層学習によるアンサンブル学習と認知症MRI画像診断への応用
3. 学会等名 2021 IEEE SMC Hiroshima Chapter Young Researchers WorkShop, pp.30-35
4. 発表年 2021年

1. 発表者名 鎌田真, 市村匠
2. 発表標題 構造適応型深層学習を用いた道路網認識手法RoadTracerによる土砂検出の試み
3. 学会等名 インテリジェント・システム・シンポジウム2021 (FAN2021), pp.181-186
4. 発表年 2021年

1. 発表者名 市村匠, 鎌田真, 原田俊英, 井上健
2. 発表標題 Teacher-Student構造適応型深層学習によるアンサンブル学習と認知症MRI画像診断への応用
3. 学会等名 第33回日本老年医学会 中国地方会, No.23
4. 発表年 2021年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 検出パラメタ生成装置、検出パラメタ生成方法、検出パラメタ生成プログラム、オブジェクト検出装置、オブジェクト検出方法、およびオブジェクト検出プログラム	発明者 市村匠, 鎌田真, 多羅尾直, 山口亮	権利者 同左
産業財産権の種類、番号 特許、特願 2 0 1 9 - 1 4 3 9 6 4	出願年 2019年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

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