

令和 3 年 6 月 2 日現在

機関番号：12601

研究種目：挑戦的研究（萌芽）

研究期間：2019～2020

課題番号：19K22861

研究課題名（和文）モダリティネットワークにおける知識ルーティング法の構築

研究課題名（英文）Knowledge Routing Method for Modality Networks

研究代表者

中山 英樹（Nakayama, Hideki）

東京大学・大学院情報理工学系研究科・准教授

研究者番号：00643305

交付決定額（研究期間全体）：（直接経費） 4,900,000円

研究成果の概要（和文）：本研究では、個別に存在する多様なデータリソースを、それぞれのデータは秘匿したまま総合的に活用するための基盤技術の提案と開発を行った。具体的には、教師なし学習や少ラベル学習によって各ローカルデータが持つ情報を十分に引き出すと共に、オープンセットを考慮した連合学習によってこれらを間接的に連携させることを試みた。本研究の成果は、TAFLやECCV等のトップ国際論文誌・国際会議に多数採択されている。

研究成果の学術的意義や社会的意義

現在成功している機械学習技術の多くは、入出力の事例（ラベル付きデータ）が十分に与えられる教師あり学習の枠組みに則っているが、実際の応用場面では大量のラベル付きデータを各人が用意することは必ずしも容易ではない。一つの解決策として、社会において個別に存在するデータを協調的に活用することが考えられるが、単純にデータを共有することはプライバシー等の観点から難しいことが多い。本研究で得られた技術により、さまざまなデータを各所有者が秘匿したまま総合的に活用することを可能とし、社会における新しいデータ活用の一助になると期待できる。

研究成果の概要（英文）：In this study, we proposed and developed a fundamental technology for collaboratively utilizing various data resources that exist individually, while keeping the respective data secret. Specifically, we attempted to extract enough information from each local data by unsupervised learning and few-label learning, and to indirectly link them by open-set federated learning. The results of this research have been accepted by top international journals and conferences such as TAFL and ECCV.

研究分野：コンピュータビジョン、自然言語処理、機械学習

キーワード：教師なし学習 少ラベル学習 連合学習 マルチモーダル 画像認識 自然言語処理

1. 研究開始当初の背景

深層学習の顕著な成功を背景に、人工知能の基盤技術として機械学習が非常に大きな注目を集めるようになったが、その中心となっているのはいわゆる教師あり機械学習であり、入出力の事例（パラレルデータ）が十分に与えられることが前提となっている。この枠組みにおいては、米国の巨大 Web 企業群のように、データを集約的に保有し入出力の対応付け（アラインメント）が容易にとれる企業が圧倒的に有利であり、我が国は著しく不利な状況にあると言わざるを得ない。一方で、伝統的に製造業やセンサデバイス等の技術は世界をリードしており、各企業がそれぞれ類のない価値あるセンサデータや職人によるアノテーションを保有している例は多い。しかし、一般にそのようなローカルデータを完全にオープンにすることはプライバシーや企業戦略の観点から難しいため、社会全体で統合的に利活用していくことは困難である。このように、IoT・AI 時代にあたり我が国はハードウェア面は先行する立場にあるが、その活用の要となるデータのアラインメントをとるための戦略に欠けていると言える。

本研究では、この問題に対し、社会において局所的に存在する多数のマルチモーダルデータが全体として構成する疎なグラフを考え、各ローカルデータそのものは秘匿したまま、グラフ上の任意のモダリティ間のアラインメントを陰に発見・最適化するための基盤技術を開発することを目的とする。これにより、誰もが自由に社会のさまざまなローカルデータを活用して自己の問題解決に役立てることを可能とし、社会におけるデータ活用の在り方を大きく変える一助になると期待される。

2. 研究の目的

本研究では、さまざまな種類のローカルデータ（モダリティ）が為すグラフ（ネットワーク）上で任意の二点間の関連性を教師なしで学習するための数理的基盤技術を開発する。このためには、まずそれぞれの小規模ローカルデータから、できるだけ豊かな特徴表現の学習を行う必要がある（項目 1, 2）。また、秘匿されたローカルデータ同士の関連性を発見するために、教師なしでクロスモーダルな変換を推定できる技術が必要である（項目 3）。さらに、これらを連携して動作させるための連合学習の枠組みが必要である（項目 4）。以上まとめると、次の 4 項目が本研究で取り組んだ主要な課題である。

- 項目 1：ラベルなしデータからの構造学習法の開発
- 項目 2：少量のラベル付きデータからの学習法の開発
- 項目 3：教師なしクロスモーダル学習法の開発
- 項目 4：オープンセットとパーソナライゼーションを考慮した連合学習法の開発

3. 研究の方法

項目 1：ラベルなしデータからの構造学習法の開発

本項目では、教師なし学習に基づき、ラベルなしデータから階層的な構造学習を行う手法を開発した。具体的には、自然言語処理のタスクの一つである談話構造解析を題材とし、Viterbi EM と呼ばれる技術を用いて談話構造のパarser を訓練する。これは、ある時点での parser によるバージン処理（E ステップ）と、その結果を用いた parser モデルの更新（M ステップ）の繰り返しにより実現される。以下、これを手法 1 と呼ぶ。本手法は汎用的な枠組みを提案するものであり、自然言語処理の他タスクはもちろん、画像や動画など他のモダリティにおける構造学習へも応用可能であると期待できる。

項目 2：少量のラベル付きデータからの学習法の開発

第一に、少量データから汎化性を向上させるために、実用上特に重要な技術であるデータ拡張に注目し、画像認識においてデータ拡張ポリシーを高速に自動探索する手法を開発した。具体的には、データ拡張の各変換（反転、回転、色調変化など）や、必要な変換を選択するプロセスを微分可能な関数によって近似することで、目的の損失を最小化するエンドツーエンドの学習の中でデータ拡張ポリシーを最適化する枠組みを実現した（図 1）。以下、これを手法 2-1 と呼ぶ。本手法は一義的には画像認識を対象としたものであるが、他のドメインにおいても基本的なデータ拡張操作さえ定義できれば、同様の枠組みでそれらの最適な使い方を探索することができる汎用的な方法であることが特徴である。

第二に、純粋に少量データのみから学習を行うことには限界があるため、メタ学習を用いて他の汎用データセットから知識転移を行う手法を開発した。本研究計画では最終的にさまざまなデータやドメインから得られる知識を統合することを考えるため、各ローカルデータで学習を行う際には、訓練データに含まれない未知事象が存在しうることを、すなわちオープンセットを前提とする必要がある。そこで、本項目では未知事象を却下することを可能とする少サンプル 1 クラス認識手法の開発に取り組んだ。具体的には、少量のデータから得られる基本的な統計量（平均、分散）をメタ学習時に活用することで、データ分布に応じて画像クラスの境界を定める 1 クラス識別器を実現する（図 2）。以下、これを手法 2-2 と呼ぶ。

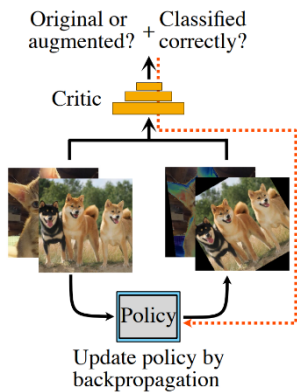


図 1 : 高速なデータ拡張ポリシー探索手法 (手法 2-1)

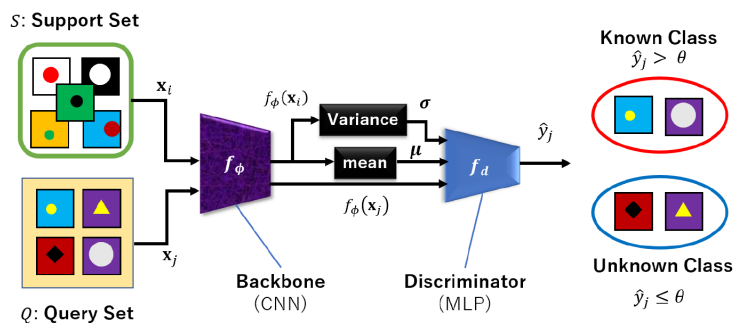


図 2 : メタ学習を用いた少サンプル 1 クラス画像認識手法 (手法 2-2)

項目 3 : 教師なしクロスモーダル学習法の開発

あるローカルデータと別のローカルデータの関連性を発見するためには、教師なしで両者を結びつける手法が必要となる。そこで、物体関連性検出という画像認識タスクを題材とし、教師なし学習で画像情報から言語情報（シーングラフ）を結びつけ、両者を相互に変換する手法の開発を行った。具体的には、画像情報を言語情報へと変換し、これをさらに画像情報へ戻した場合に元の画像情報と一致すべきであるとする循環損失と、変換した言語情報が言語的に妥当であるかを判定する plausibility の二つを損失関数として用いることで、教師なしでありながら両者の妥当な対応関係を推定することができる (図 3)。以下、これを手法 3 と呼ぶ。

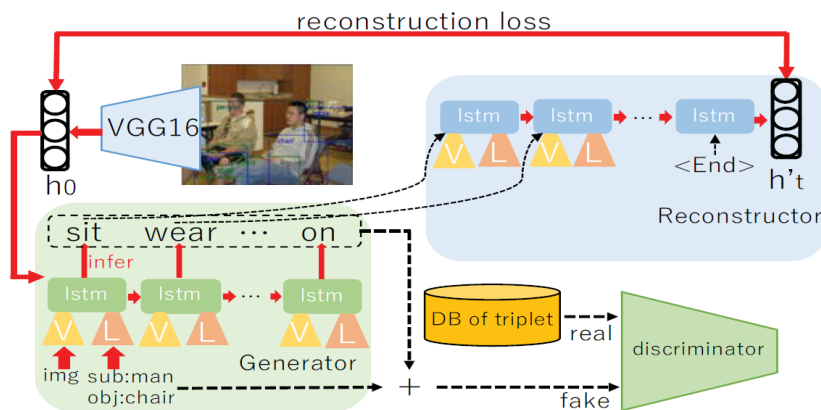


図 3 : 教師なし物体関連性検出手法 (手法 3)

項目 4 : オープンセットとパーソナライゼーションを考慮した連合学習法の開発

最終的に、さまざまなローカルデータをそれぞれ秘匿したまま知識を統合するために、連合学習と呼ばれる枠組みが有力である。連合学習は近年注目を集めている技術であるが、基本的にはローカルデータから得られる知識を集約して一つの平均化されたモデルを作ることが目的であり、本研究計画で目的とするように各ローカル環境での性能の最適化や、オープンセットを考慮したものは少ない (引用文献①)。本項目では、これらの要素を盛り込んだ新しい連合学習の枠組みを考える。より具体的には、手法 3 によって発見される目的タスクと関連性の高いデータを対象とし、手法 2-2 のオープンセットを考慮した評価関数を用いることで全体の学習を行う。

4. 研究成果

項目 1 : ラベルなしデータからの構造学習法の開発

談話構造解析の標準的なベンチマークである RST-PARSEVAL (引用文献②) において手法 1 を適用したところ、既存の教師なし学習手法を大きく上回る性能を得て、教師あり学習に匹敵する良好な精度を得た (表 1)。また、モデルの初期化を工夫することで学習をしなくても非常に高いスコアが得られることが分かり、本タスクにおける先見知識の重要性を示唆する興味深い知見が得られた。本成果は、自然言語処理分野のトップ国際論文誌である Transactions of the Association for Computational Linguistics (ACL) にて発表された。

項目 2 : 少量のラベル付きデータからの学習法の開発

手法 2-1 について、CIFAR-10、SVHN、ILSVRC 等の画像認識における標準的なベンチマークデータセットを用いて評価を行ったところ、提案手法は先行研究と同等の画像認識を得ると共に、既存手法(引用文献③)の数十倍から数百倍高速な学習を実現した(表 2)。本成果は、コンピュータビジョン分野のトップ国際会議である European Conference on Computer Vision (ECCV2020)で発表されると共に、国内会議である MIRU2020 において学生優秀賞を受賞した。

手法 2-2 について、CUB-200、Food-101、ILSVRC 等のベンチマークデータセットを用い、標準的な少サンプル学習の手順に則って評価を行った。この結果、提案手法は少サンプル 1 クラス認識の問題において、既存の全ての手法を上回る精度を得た(表 3)。また、合成データセットを用いた評価により、さまざまなデータ分布を持つクラスが入力される場合にも、提案手法は安定に良好な精度を得られることが示された。本成果は、国内研究会である PRMU において月間ベストプレゼンテーション賞を受賞した。

表 1 : 教師なし談話構造解析手法
(手法 1) の評価結果

Method	UP	UR	Micro F_1
Unsupervised			
Kobayashi et al. (2019)	-	-	- (80.8)
Ours, initialized by (a)	66.2	68.8	67.5 (84.0)
Ours, initialized by (b)	66.8	69.4	68.0 (84.3)
Ours (b) + Aug.	67.3	69.9	68.6 (84.6)
Supervised			
Ours, supervised	68.3	70.9	69.6 (85.1)
Feng and Hirst (2014)*	-	-	- (84.4)
Joty et al. (2015)*	-	-	- (82.5)
Human	-	-	- (88.7)

表 2 : 高速なデータ拡張ポリシー探索手法
(手法 2-1) の学習時間の比較

Dataset	AA	PBA	Fast AA	Faster AA (ours)
CIFAR-10	5,000	5.0	3.5	0.23
SVHN	1,000	1.0	1.5	0.061
ImageNet	15,000	-	450	2.3

表 3 : 少サンプル 1 クラス画像認識手法 (手法 2-2) の評価結果

Method	ILSVRC		CUB-200-2011		Food-101	
	AUROC	F1	AUROC	F1	AUROC	F1
OC-MAML	0.7855	0.7229	0.8848	0.7991	0.6704	0.6699
OOD-MAML	0.8258	0.7596	0.9150	0.8315	0.7457	0.7107
OW-ProtoNet	0.8753	0.7962	0.9104	0.8340	0.7969	0.7406
BeNet (ours)	0.8982	0.8289	0.9244	0.8584	0.8290	0.7654

項目 3 : 教師なしクロスモーダル学習法の開発

手法 3 を VRD や MSCOCO 等のデータセットに適用し、実際に教師なしで物体関連性検出手法の学習およびシーングラフの出力が可能であることを示し、世界初となる成果を得た。さらに、ごく少数の教師ありデータの使用を許す場合、提案手法は半教師あり手法として拡張でき、教師ありデータのみを用いた学習を大きく上回る精度を達成した。これは、異なるローカルデータ間でごく一部のデータを共有・対応づけ可能であるとした場合に相当し、比較的現実的な設定のもとでより精度よくデータ間の関連性を推定することにつながると期待できる。本成果の一部は、IEEE International Conference on Image Processing (ICIP2020)で発表された。

項目 4 : オープンセットとパーソナライゼーションを考慮した連合学習法の開発

MNIST、MNIST-M、USPS 等の文字画像データセットを用い、各ローカルデータでオープンセットを考慮した連合学習のセットアップを構築し、手法 4 の検証を行った。この結果、実際に通常の連合学習よりも、各ローカル環境での目的タスクで精度が向上できることが示された。

<引用文献>

- ① Chen Zhang, Yu Xie, Hang Bai, Bin Yu, Weihong Li, and Yuan Gao, A survey on federated learning, Knowledge-Based Systems, Volume 216, 2021.
- ② Mathieu Morey, Philippe Muller, and Nicholas Asher, A dependency perspective on RST discourse parsing and evaluation, Computational Linguistics, Volume 44, Issue 2, pp.197-235, 2018.
- ③ Sungbin Lim, Ildoo Kim, Taesup Kim, Chiheon Kim, and Sungwoong Kim, Fast AutoAugment, In Proceedings of NeurIPS, 2019.

5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 0件/うちオープンアクセス 4件）

1. 著者名 Noriki Nishida, Hideki Nakayama	4. 巻 8
2. 論文標題 Unsupervised Discourse Constituency Parsing Using Viterbi EM	5. 発行年 2020年
3. 雑誌名 Transactions of the Association for Computational Linguistics	6. 最初と最後の頁 215-230
掲載論文のDOI（デジタルオブジェクト識別子） 10.1162/tacl.a.00312	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 Jan Zdenek, Hideki Nakayama	4. 巻 -
2. 論文標題 Erasing Scene Text with Weak Supervision	5. 発行年 2020年
3. 雑誌名 Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)	6. 最初と最後の頁 2238-2246
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/WACV45572.2020.9093544	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 Takumi Ohkuma, Hideki Nakayama	4. 巻 -
2. 論文標題 Efficient Base Class Selection Algorithms for Few-Shot Classification	5. 発行年 2020年
3. 雑誌名 Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR)	6. 最初と最後の頁 271-275
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3372278.3390724	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 -
1. 著者名 Ryuichiro Hataya, Zdenek Jan, Kazuki Yoshizoe, Hideki Nakayama	4. 巻 LNCS 12370
2. 論文標題 Faster AutoAugment: Learning Augmentation Strategies using Backpropagation	5. 発行年 2020年
3. 雑誌名 Computer Vision - ECCV 2020	6. 最初と最後の頁 1-16
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-58595-2_1	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 Taiga Kashima, Kento Masui, Hideki Nakayama	4. 巻 -
2. 論文標題 Unsupervised Visual Relationship Inference	5. 発行年 2020年
3. 雑誌名 Proceedings of the IEEE International Conference on Image Processing (ICIP)	6. 最初と最後の頁 1476-1480
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/icip40778.2020.9190770	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 -

1. 著者名 Yuxuan Wu, Hideki Nakayama	4. 巻 LNCS 12624
2. 論文標題 Graph-based Heuristic Search for Module Selection Procedure in Neural Module Network	5. 発行年 2020年
3. 雑誌名 Computer Vision - ACCV 2020	6. 最初と最後の頁 560-575
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-69535-4_34	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 Kai Katsumata, Hideki Nakayama	4. 巻 -
2. 論文標題 Semantic Image Synthesis from Inaccurate and Coarse Masks	5. 発行年 2021年
3. 雑誌名 Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)	6. 最初と最後の頁 2285-2289
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICASSP39728.2021.9414521	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17件 (うち招待講演 5件 / うち国際学会 7件)

1. 発表者名 幡谷龍一郎, 中山英樹
2. 発表標題 MADA0: データ拡張最適化のためのメタ的アプローチ
3. 学会等名 第23回 画像の認識・理解シンポジウム (MIRU 2020)
4. 発表年 2020年

1. 発表者名 大熊拓海, 中山英樹
2. 発表標題 Few-shot classificationにおけるベースクラスデータ多様化の有効性
3. 学会等名 第23回 画像の認識・理解シンポジウム (MIRU 2020)
4. 発表年 2020年

1. 発表者名 中山英樹
2. 発表標題 深層学習の発展と応用
3. 学会等名 応用物理学会 第152回微小光学研究会 (招待講演)
4. 発表年 2019年

1. 発表者名 中山英樹
2. 発表標題 機械学習とマルチメディア認識理解
3. 学会等名 第4回 統計・機械学習若手シンポジウム (招待講演)
4. 発表年 2019年

1. 発表者名 中山英樹
2. 発表標題 マルチモーダル情報と機械翻訳
3. 学会等名 アジア太平洋機械翻訳協会年次セミナー AAMT 2019, Tokyo ~ 機械翻訳最前線 ~ (招待講演)
4. 発表年 2019年

1. 発表者名 Takumi Ohkuma, Hideki Nakayama
2. 発表標題 Belonging Network -- Few-shot One-class Image Classification for Classes with Various Distributions --
3. 学会等名 電子情報通信学会 PRMU 2020年度12月研究会
4. 発表年 2020年

1. 発表者名 杉浦昇太, 西田典起, 朱中元, 中山英樹
2. 発表標題 相互情報量最小化による例文に基づく制御可能な言い換え生成
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2021年

1. 発表者名 中山英樹, 松井孝太, Xueting Wang, 幡谷龍一郎
2. 発表標題 続・限られたデータからの深層学習 (オーガナイズドセッション)
3. 学会等名 第27回 画像センシングシンポジウム (招待講演)
4. 発表年 2021年

1. 発表者名 Hideki Nakayama
2. 発表標題 Efficient Training of Neural Module Networks and Applications
3. 学会等名 Fifth International Workshop on Symbolic-Neural Learning (SNL-2021) (招待講演) (国際学会)
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

少ないデータからの機械学習に関する研究内容と成果
<http://www.nlab.ci.i.u-tokyo.ac.jp/projects.html>

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