

令和 5 年 6 月 16 日現在

機関番号：12601

研究種目：基盤研究(B)（一般）

研究期間：2019～2022

課題番号：19H04166

研究課題名（和文）言語概念のマルチメディアへのグラウンディングに基づく世界知識構築

研究課題名（英文）Building World Knowledge by Grounding Language and Multimedia

研究代表者

中山 英樹（Hideki, Nakayama）

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） 13,200,000円

研究成果の概要（和文）：本研究では、画像・動画・テキスト等のマルチメディアデータからさまざまな概念に関する常識的知識をマイニングし、これをグラフ構造を有するデータベース（マルチメディア世界知識と名付ける）として整え、コンピュータビジョンや自然言語処理などの応用タスクへ活用する一連の手法を開発した。これらの成果は合計27件の国際会議・国際論文誌で発表され（うち9件は最難関国際会議で採択）、国際的に高い評価を得たといえる。

研究成果の学術的意義や社会的意義

画像やテキストの認識・生成を行う人工知能は近年驚異的な進歩を遂げ、社会的にも大きなインパクトを与えている。その原動力となったのは、大量のデータから学習した大規模言語モデルであるが、本研究では知識グラフという異なる形式の知識基盤の獲得方法とその有効性を示した点に独創性がある。知識グラフは、大規模言語モデルと併用することで人工知能の性能をさらに向上できる可能性があると同時に、解釈性や可制御性に優れた表現であるため、今後の社会でより信頼性の高い人工知能技術を実装・展開するための重要なアプローチになると考えられる。

研究成果の概要（英文）：In this research, we developed a series of methods for mining commonsense knowledge about various concepts from multimedia data such as images, videos, and text, preparing it as a database with a graph structure (named Multimedia World Knowledge), and utilizing it for various artificial intelligence tasks. These results have been presented at a total of 27 international conferences and international journals (9 of which were accepted by top conferences), and have been highly acclaimed internationally.

研究分野：画像認識・自然言語処理・機械学習・人工知能

キーワード：画像認識 自然言語処理 知識獲得 知識グラフ 世界知識 マルチメディア マルチモーダル 人工知能

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、研究者個人に帰属します。

## 1. 研究開始当初の背景

深層学習の発展により自然言語処理は大きな進歩を遂げたが、未だ本質的には陽に観測されるシンボル列(単語列)の統計処理に留まっているといえる。すなわち、表層的な入出力関係のモデル化で解ける従来タスクの枠組みの中で精度を向上させたにすぎない。しかしながら、自然言語には表層から推し量れないさまざまなレベルの曖昧性や隠れた文脈が存在し、我々人間は社会的に共有するさまざまな常識的知識を活用してこれを補いながら言語理解を行っている。このような「常識」もしくは「世界知識」が必要となる場面では従来の自然言語処理技術はほぼ無力であり、これをどのように獲得し、計算機へ実装するかが自然言語処理における最大の問題の一つであり、古くから議論が続いている。

今世紀に入ると、Web から大規模なデータが容易に入手可能になり、計算機の能力も劇的に向上したことを背景として、大規模なテキストデータからの知識獲得が注目されるようになった。しかしながら、そもそもテキストとは他の人間へ情報を伝えるためにアウトプットされるものが大半であるため、大多数の人間にとって価値がない、あるいは興味を持たれない情報は明示的に表れにくい。このためテキストマイニングは、例えば IBM の QA システム Watson のように高度なクイズに回答できるほどの知識を得ることに成功する一方で、誰にとっても自明な知識ほど得難いというジレンマがあり、日常的な推論を可能とするための知識を得るには本質的な限界がある。

このような背景を踏まえ、本研究では画像・動画像等のマルチメディアをマイニングし、世界知識を構築するアプローチを提案する。画像や動画像は実世界の観察そのものであり、陽にテキスト化されにくい日常的な物事の時空間的な共起関係や因果関係に関する豊富な情報を含んでいるため、テキストマイニングで得られる世界知識を強力に補完するものになりうると期待できる。

## 2. 研究の目的

本研究では、画像・動画像等のマルチメディアに言語概念をグラウンディング(記号接地)させ、時空間的な共起関係を手掛かりにさまざまな概念間の関連性を推定し、グラフ構造を有するデータベース(知識グラフ)として整える。これをマルチメディア世界知識と呼ぶことにする。さらに、マルチメディア世界知識を従来の知識 DB と統合し、各種応用タスクの推論機構と接続することで精度向上が行えることを示す。

本研究の目的を達成するための核となる研究項目をより具体的に述べる。まず、画像からグラフ形式による情報抽出を行うシーングラフ生成技術が、マルチメディア世界知識を構築するための入り口となる(項目1)。次に、多数のテキストや画像などのデータから、抽象化された知識グラフを構築する技術が必要である(項目2)。さらに、マルチメディア知識グラフを外部の知識データベースと接続することが、知識の可用性・汎用性を高めるために重要である(項目3)。

以上まとめると、次の3項目が本研究で取り組んだ主要な課題である。なお、本プロジェクトの研究実施期間中に、大規模言語モデルが世界知識の基盤として極めて驚異的な進歩を遂げたため、既存の知識グラフのみならず大規模言語モデルとの比較評価や接続についても柔軟に取り組んだ。

項目1: 画像シーングラフ生成手法の開発

項目2: マルチメディアデータからの知識グラフ構築法の開発

項目3: マルチメディア世界知識と外部知識基盤の接続法の開発

## 3. 研究の方法

### 項目1: 画像シーングラフ生成手法の開発

本項目では、データ効率性の高い画像シーングラフ生成手法の開発に取り組んだ。シーングラフとは、一枚の画像中の複数物体の名称や属性、互いの関連性などをグラフ構造で表した表現であり、マルチメディアからの知識グラフ構築において最小単位をなす重要な要素であると考えられる。しかしながらシーングラフ生成は、一枚の画像に対する詳細なアノテーションが必要であることからアノテーターへの負担が大きく、訓練データの作成が高コストとなることや、ラベルが単純なイベントへ偏りやすいデータバイアスの問題が課題となっている。これらに対処するため、本項目では以下に述べる二つの手法を開発した。

まず、教師なし学習によるシーングラフ生成手法を開発した。具体的には、画像情報をトリプレット(subject-predicate-object)の形式で表されるシーングラフの構成単位へと変換し、これをさらに画像情報へ戻した場合に元の画像情報とできるだけ一致することを求める循環一致損失と、変換したトリプレットの言語的妥当性を求める損失の二つを損失関数として用いることで、パラレルデータなしにシーングラフ生成の学習を行うことを実現した。以下、この手法を手法1-1とよぶ。

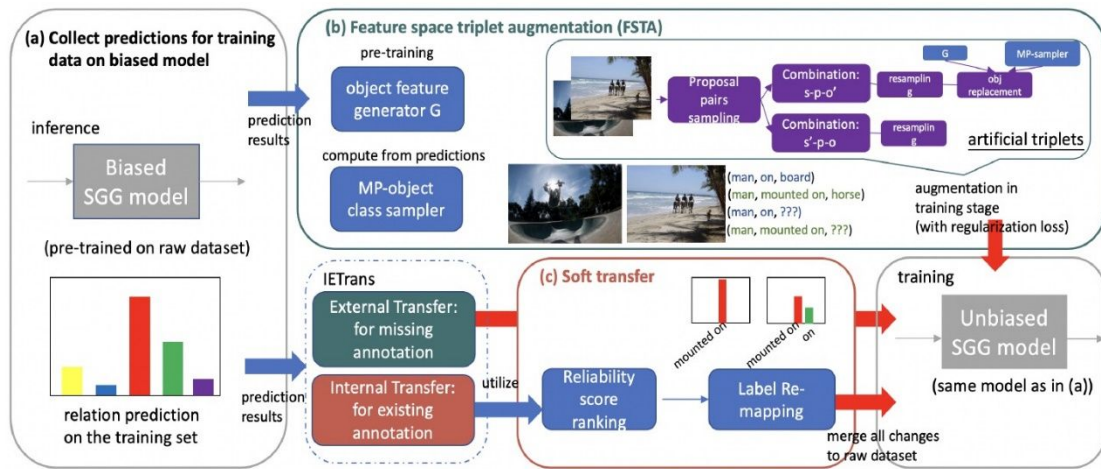


図 1 : ロングテールデータからのシーングラフ生成の学習 (手法 1 - 2)

次に、データ拡張によりロングテールな訓練データからバランスのとれた学習を行う手法を開発した。本手法の中心となるアイデアは、同一バッチ中の画像の特徴量を組み合わせ、頻度の少ないトリプレット（テールデータ）のサンプルを人工的に合成することにある（図 1）。この方法により、訓練データの偏りを低減しながら学習することで、テールデータの領域を含めた多様な出力が可能となる。以下、この手法を手法 1 - 2 とよぶ。

## 項目 2 : マルチメディアデータからの知識グラフ構築法の開発

まず、テキストデータのみから知識グラフの構築を行うベースライン手法を開発した。ここでは自然言語処理の物語生成タスクを題材とし、多数の物語文章のコーパスから、物語における物事の自然な遷移関係を表す知識グラフを構築する（図 2）。まず、コーパス中のテキストデータのクラスタリングを行い、代表的な物語トピックの抽出を行う。次に、それぞれのトピックについて時系列に沿ったイベントの抽出を行い、一定の頻度を超えるイベントおよびその遷移関係をそれぞれノード・エッジとして表したグラフを構築する。このようにして得られる知識グラフを物語の骨子の計画に用いることで、説明性や可制御性を備えた物語生成を行うことができる。以下、この手法を手法 2 - 1 と呼ぶ。

このメカニズムに加え、テキストと画像の特徴表現のグラウンディングを行うことで、テキストのみならず画像を用いた知識グラフの構築も可能になると期待できる。これを目指し、BERT により辞書データ (Wiktionary) から特徴抽出を行った単語概念と画像領域特徴をアラインメントし、共通の埋め込み空間を学習する手法を提案した（図 3）。この空間では、Transformer の注意機構により、画像領域の視覚的特徴のみならず複数領域の共起関係や位置情報なども考慮された埋め込みが行われていることが重要であり、この空間上での距離を基準として概念が為すグラフを構築することができる。その具体的な応用として、画像中の未知物体に対しグラフ上の最近傍の単語を検索することで、ゼロショットの画像キャプションを高い精度で実現できることを示した。以下、この手法を手法 2 - 2 と呼ぶ。

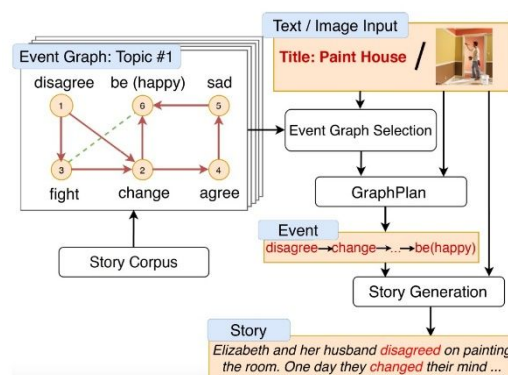


図 2 : テキストコーパスからの知識グラフ構築と物語生成手法 (手法 2 - 1)

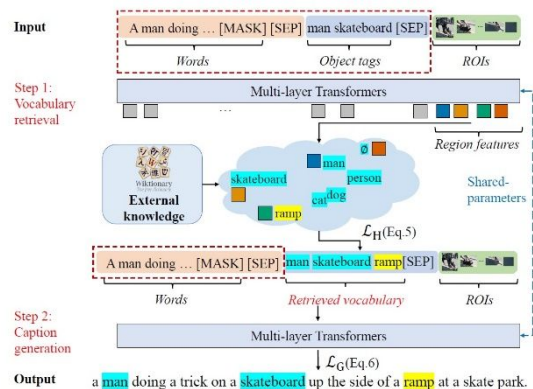


図 3 : 画像・テキストのアラインメントと外部辞書を用いたゼロショット画像キャプション (手法 2 - 2)

### 項目 3：マルチメディア世界知識と外部知識基盤の接続法の開発

手法 2 - 1 の考え方を拡張し、マルチメディアから時間的なダイナミクスを内包する知識グラフを抽出する手法を開発した。まず、画像時系列から物体・イベントなどのコンセプトを抽出し、その時空間的な共起や遷移関係をグラフ構造の形で表現する。さらに、ここへ外部の大規模知識リソース (ConceptNet) を接続し、トップダウンな常識的知識を加えた広範な知識グラフを構築することができる。

これを活用するダウンストリームタスクとして、与えられた画像時系列を題材として物語を生成する画像ストーリー生成 (visual story telling) の手法開発を行った (図 4)。本手法は、画像情報を足掛かりに ConceptNet 等の知識グラフ上の探索を行うことにより、単なる画像の記述にとどまらない多様なストーリーを生成することが可能である。以下、この手法を手法 3 - 1 と呼ぶ。さらに、与えられた画像時系列の未来の状況を予測して文章で表す予知キャプションングを新規に提案し、これを実現する手法を開発した (図 5)。これを手法 3 - 2 と呼ぶ。

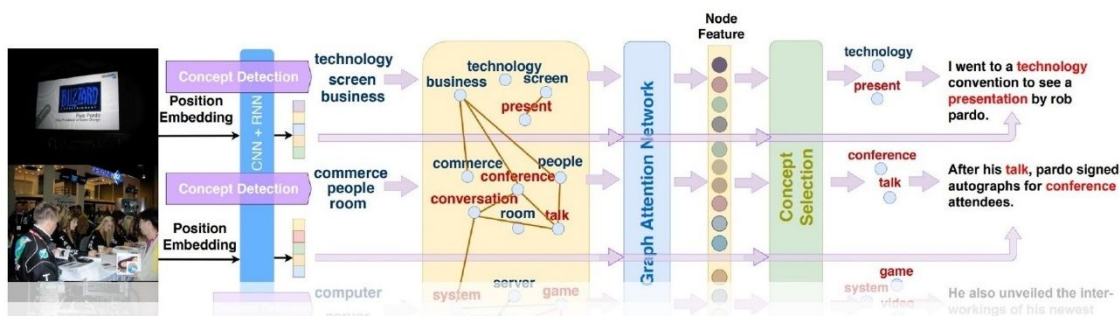


図 4：外部知識を接続した知識グラフに基づく画像ストーリー生成 (手法 3 - 1)

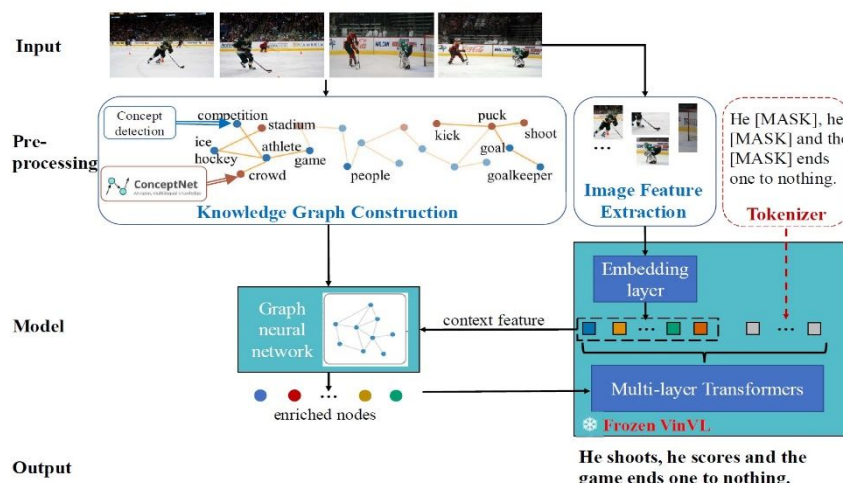


図 5：外部リソースを接続した知識グラフによる予知キャプションング (手法 3 - 2)

## 4. 研究成果

### 項目 1：画像シーングラフ生成手法の開発

手法 1 - 1 について、VRD や MSCOCO 等の標準的なデータセットを用いて実装を行い、教師なしで物体関連性検出の学習およびシーングラフの出力が実際に可能であることを示した。教師なしによるシーングラフ生成は当時世界初となる成果である。また、提案手法を半教師あり手法として拡張し、ラベル付き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。

手法 1 - 2 について、Visual Genome データセットを用いた評価を行った。提案手法はデータ拡張によって訓練データの増強を行うものであるから、基本的にはモデルアーキテクチャに依存しない貢献をなすため、既存のさまざまなシーングラフ生成手法と併用することが可能である。実際に、提案手法を実装することにより、既存のシーングラフ生成手法の精度を改善でき、特にテールデータに相当するトリプレットの生成精度を大きく向上させることが示された。また、他のロングテール学習手法と比較し、全体的に優れた精度を得る結果となった。

### 項目 2：マルチメディアデータからの知識グラフ構築法の開発

手法 2 - 1 について、ROGStories コーパスを用い、入力トピックに対し 5 文からなる短い物語を生成するタスクにおいて、自動評価と人手評価の両面から既存手法と性能比較を行った。この結果、提案手法は生成する物語が多様性・質 (与えたトピックとの関連性および論理性) の両

面で優れた性能を得ることが示された。特に、研究当時に最先端の言語モデルであった GPT-2 では、冗長な表現の繰り返しや前後での矛盾などの問題を含む出力を行うことが多かったが、提案手法では知識グラフを用いた計画によりこれらの問題を低減でき、より優れた物語生成が実現できることが示された。

手法 2 - 2 について、held-out COCO および Nocaps データセットを用い、ゼロショット画像キャプションを評価タスクとして性能調査を行った。この結果、提案手法は既存手法を全て上回る良好な精度を得た。また、マルチモーダル特徴空間の可視化を行ったところ、画像の領域特徴と単語特徴が未知語を含め妥当にアラインメントされていることが示唆された ( 図 6 )。本手法は、コンピュータビジョンの最難関国際会議である CVPR 2022 へ採択された。

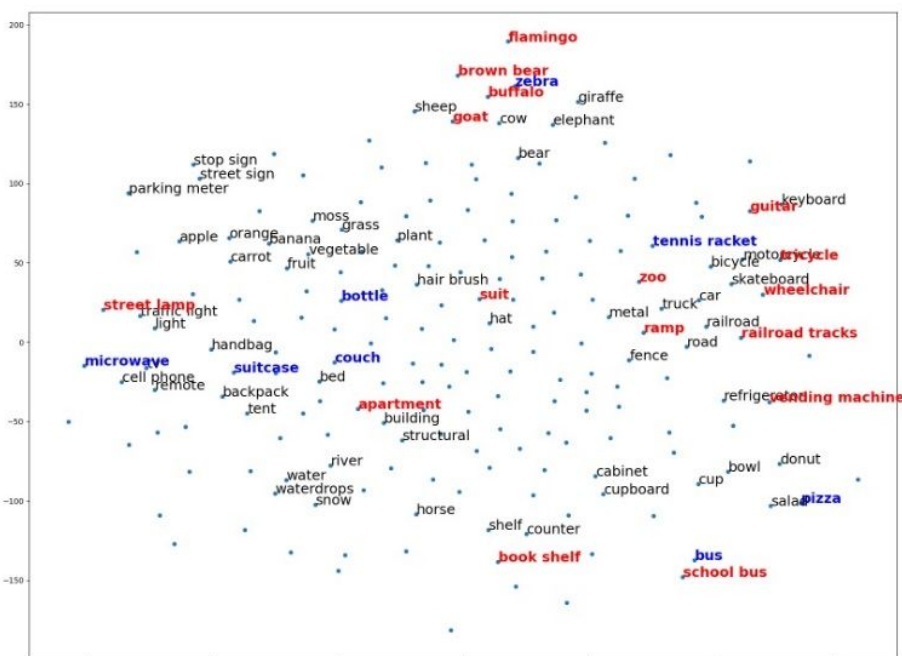


図 6 : 手法 2 - 2 における画像領域特徴と単語特徴のアラインメントの例

### 項目 3 : マルチメディア世界知識と外部知識基盤の接続法の開発

手法 3 - 1 について、画像ストーリー生成の標準的なデータセットである VIST を用い、主観評価実験を行った。この結果、外部知識グラフの接続により、画像情報のみから概念抽出を行った場合と比べて生成するストーリーの質を大きく向上できることが示された。また、当時の最先端の大規模言語モデルである BART をテキストデコーダとして利用したベースライン手法は、非常に流暢なストーリーを生成できるが、これに提案手法による知識グラフを加えるとストーリーの品質がさらに大きく向上することが分かった。この結果は、大規模言語モデルと知識グラフが外部の知識リソースとして相補的な関係にあることを示唆している。本手法は、人工知能の最難関国際会議である AAAI 2021 へ採択された。

手法 3 - 2 についても同様に VIST データセットを用いた評価を行った。予知キャプションは我々が提案した新しいタスクであるため、まずタスクのセットアップや評価方法などの全体の枠組みを整えた。具体的には、VIST データセットの各サンプルが持つ画像 5 枚のうち最初の 4 枚を入力として与え、最後の 1 枚 ( システムへは与えない ) に紐づいた説明文の内容を予測するタスクとして定義した。ベースラインとして、入力画像のみから画像ストーリー生成を行い、その結果を大規模言語モデルへ入れてその先の予測を行うアプローチが考えられる。しかしながら、このベースラインを実装したところ十分な性能は出せず、外部知識グラフを用いた提案手法が大きく上回る結果となった。この結果も、大規模言語モデルに対するマルチメディア世界知識の独自の貢献を示すものであるといえる。本手法は、コンピュータビジョンの最難関国際会議である CVPR 2023 へ採択された。

**総括 :** 以上まとめると、本研究では画像・動画 ( 画像時系列 ) およびテキストデータからのマルチメディア世界知識構築に加え、外部知識グラフの接続による知識拡張や、ダウンストリームタスクにおける有効性の検証に至るまで順調に実施され、当初の目的が達成されたと考える。また、研究期間を通じて合計 27 件の国際会議・国際論文誌で発表を行い、うち 9 件は最難関国際会議に採択されており ( CVPR 3 件、EMNLP 2 件、AAAI 2 件、ECCV、ACM Multimedia )、国際的に高い評価を伴う十分な成果を得たといえる。

## 5. 主な発表論文等

〔雑誌論文〕 計27件（うち査読付論文 27件 / うち国際共著 4件 / うちオープンアクセス 16件）

1. 著者名 Duc Minh Vo, Quoc-An Luong, Akihiro Sugimoto, Hideki Nakayama	4. 巻 -
2. 論文標題 A-CAP: Anticipation Captioning with Commonsense Knowledge	5. 発行年 2023年
3. 雑誌名 Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	6. 最初と最後の頁 -
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 オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Yi-Pei Chen, An-Zi Yen, Hen-Hsen Huang, Hideki Nakayama, Hsin-Hsi Chen	4. 巻 -
2. 論文標題 LED: A Dataset for Life Event Extraction from Dialogs	5. 発行年 2023年
3. 雑誌名 Findings of the Association for Computational Linguistics: EACL 2023	6. 最初と最後の頁 384-398
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 オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Rui Yang, Duc Minh Vo, Hideki Nakayama	4. 巻 -
2. 論文標題 Indirect Adversarial Losses via an Intermediate Distribution for Training GANs	5. 発行年 2023年
3. 雑誌名 Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	6. 最初と最後の頁 4641-4650
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/WACV56688.2023.00463	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 Hong Chen, Rujun Han, Te-Lin Wu, Hideki Nakayama and Nanyun Peng	4. 巻 -
2. 論文標題 Character-Centric Story Visualization via Visual Planning and Token Alignment	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)	6. 最初と最後の頁 8259-8272
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 オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Hong Chen, Duc Minh Vo, Hiroya Takamura, Yusuke Miyao, Hideki Nakayama	4. 巻 -
2. 論文標題 StoryER: Automatic Story Evaluation via Ranking, Rating and Reasoning	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)	6. 最初と最後の頁 1739-1753
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 オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Yuxuan Wu, Hideki Nakayama	4. 巻 -
2. 論文標題 Weakly Supervised Formula Learner for Solving Mathematical Problems	5. 発行年 2022年
3. 雑誌名 Proceedings of the 29th International Conference on Computational Linguistics (COLING)	6. 最初と最後の頁 1743-1752
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 オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Jun Takeuchi, Noriki Nishida, Hideki Nakayama	4. 巻 -
2. 論文標題 Neural Networks in a Product of Hyperbolic Spaces	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop	6. 最初と最後の頁 211-221
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 オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Rui Yang, Raphael Shu, Hideki Nakayama	4. 巻 E105-D
2. 論文標題 Improving Noised Gradient Penalty with Synchronized Activation Function for Generative Adversarial Networks	5. 発行年 2022年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 1537-1545
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2022EDP7019	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 該当する

1. 著者名 Ryuichiro Hataya, Hideki Nakayama	4. 巻 -
2. 論文標題 DJMix: Unsupervised Task-agnostic Image Augmentation for Improving Robustness of Convolutional Neural Networks"	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 International Joint Conference on Neural Networks (IJCNN)	6. 最初と最後の頁 1-8
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/IJCNN55064.2022.9892068	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 Ikki Kishida, Hideki Nakayama	4. 巻 -
2. 論文標題 Pixel to Binary Embedding Towards Robustness for CNNs	5. 発行年 2022年
3. 雑誌名 Proceedings of the 26th International Conference on Pattern Recognition (ICPR)	6. 最初と最後の頁 2279-2285
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICPR56361.2022.9956572	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 Duc Minh Vo, Hong Chen, Akihiro Sugimoto, Hideki Nakayama	4. 巻 -
2. 論文標題 NOC-REK: Novel Object Captioning with Retrieved Vocabulary from External Knowledge	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	6. 最初と最後の頁 17979-17987
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR52688.2022.01747	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, Duc Minh Vo, Hideki Nakayama	4. 巻 -
2. 論文標題 OSSGAN: Open-Set Semi-Supervised Image Generation	5. 発行年 2022年
3. 雑誌名 Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	6. 最初と最後の頁 11175-11183
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR52688.2022.01090	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 Yuhao Xu and Hideki Nakayama	4. 巻 -
2. 論文標題 DCT-based Fast Spectral Convolution for Deep Convolutional Neural Networks	5. 発行年 2021年
3. 雑誌名 Proceedings of the International Joint Conference on Neural Networks (IJCNN)	6. 最初と最後の頁 1-8
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/IJCNN52387.2021.9534135	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 Hong Chen, Raphael Shu, Hiroya Takamura, Hideki Nakayama	4. 巻 -
2. 論文標題 GraphPlan: Story Generation by Planning with Event Graph	5. 発行年 2021年
3. 雑誌名 Proceedings of the 14th International Conference on Natural Language Generation (INLG)	6. 最初と最後の頁 377-386
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 オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Jan Zdenek and Hideki Nakayama	4. 巻 -
2. 論文標題 JokerGAN: Memory-Efficient Model for Handwritten Text Generation with Text Line Awareness	5. 発行年 2021年
3. 雑誌名 Proceedings of the 29th ACM International Conference on Multimedia (ACMMM)	6. 最初と最後の頁 5655-5663
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3474085.3475713	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 Hong Chen, Hiroya Takamura, Hideki Nakayama	4. 巻 -
2. 論文標題 SciXGen: A Scientific Paper Dataset for Context-Aware Text Generation	5. 発行年 2021年
3. 雑誌名 Findings of the Association for Computational Linguistics: EMNLP 2021	6. 最初と最後の頁 1483-1492
掲載論文のDOI (デジタルオブジェクト識別子) 10.18653/v1/2021.findings-emnlp.128	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 Duc Minh Vo, Akihiro Sugimoto, Hideki Nakayama	4. 巻 -
2. 論文標題 PPCD-GAN: Progressive Pruning and Class-Aware Distillation for Large-Scale Conditional GANs Compression	5. 発行年 2022年
3. 雑誌名 Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	6. 最初と最後の頁 1422-1430
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV51458.2022.00149	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 Ryuichiro Hataya, Jan Zdenek, Kazuki Yoshizoe, Hideki Nakayama	4. 巻 -
2. 論文標題 Meta Approach to Data Augmentation Optimization	5. 発行年 2022年
3. 雑誌名 Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	6. 最初と最後の頁 3535-3544
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV51458.2022.00359	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 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 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 Hong Chen, Yifei Huang, Hiroya Takamura, and Hideki Nakayama	4. 巻 -
2. 論文標題 Commonsense Knowledge Aware Concept Selection for Diverse and Informative Visual Storytelling	5. 発行年 2021年
3. 雑誌名 Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)	6. 最初と最後の頁 999-1008
掲載論文のDOI (デジタルオブジェクト識別子) 10.1609/aaai.v35i2.16184	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 Raphael Shu, Jason Lee, Hideki Nakayama, Kyunghyun Cho	4. 巻 -
2. 論文標題 Latent-Variable Non-Autoregressive Neural Machine Translation with Deterministic Inference Using a Delta Posterior	5. 発行年 2020年
3. 雑誌名 Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)	6. 最初と最後の頁 8846-8853
掲載論文のDOI (デジタルオブジェクト識別子) 10.1609/aaai.v34i05.6413	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 該当する

1. 著者名 Ryuichiro Hataya, Hideki Nakayama	4. 巻 -
2. 論文標題 LOL: Learning To Optimize Loss Switching Under Label Noise	5. 発行年 2019年
3. 雑誌名 Proceedings of the 26th IEEE International Conference on Image Processing (ICIP)	6. 最初と最後の頁 3621-3625
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICIP.2019.8803512	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 Yuhao Xu, Hideki Nakayama	4. 巻 -
2. 論文標題 DCT Based Information-Preserving Pooling for Deep Neural Networks	5. 発行年 2019年
3. 雑誌名 Proceedings of the 26th IEEE International Conference on Image Processing (ICIP)	6. 最初と最後の頁 894-898
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICIP.2019.8802962	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 Rui Yang, Hideki Nakayama	4. 巻 -
2. 論文標題 Bipolar Gan: Double Check the Solution Space and Lighten False Positive Errors in Generative Adversarial Nets	5. 発行年 2019年
3. 雑誌名 Proceedings of the 26th IEEE International Conference on Image Processing (ICIP)	6. 最初と最後の頁 4260-4264
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICIP.2019.8803613	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 Ikki Kishida, Hideki Nakayama	4. 巻 1142
2. 論文標題 Empirical Study of Easy and Hard Examples in CNN Training	5. 発行年 2019年
3. 雑誌名 Proceedings of the 26th International Conference on Neural Information Processing (ICONIP 2019), Communications in Computer and Information Science	6. 最初と最後の頁 179-188
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-36808-1_20	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 Yuhao Xu, Hideki Nakayama	4. 巻 -
2. 論文標題 Shifted Spatial-Spectral Convolution for Deep Neural Networks	5. 発行年 2019年
3. 雑誌名 Proceedings of the 1st ACM International Conference on Multimedia in Asia (MMAsia)	6. 最初と最後の頁 1-6
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3338533.3366575	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37件 (うち招待講演 6件 / うち国際学会 28件)

1. 発表者名 Duc Minh Vo, Quoc-An Luong, Akihiro Sugimoto, Hideki Nakayama
2. 発表標題 A-CAP: Anticipation Captioning with Commonsense Knowledge
3. 学会等名 The 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (国際学会)
4. 発表年 2023年

1. 発表者名 Rui Yang, Duc Minh Vo, Hideki Nakayama
2. 発表標題 Indirect Adversarial Losses via an Intermediate Distribution for Training GANs
3. 学会等名 The 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (国際学会)
4. 発表年 2023年

1. 発表者名 Hong Chen, Rujun Han, Te-Lin Wu, Hideki Nakayama and Nanyun Peng
2. 発表標題 Character-Centric Story Visualization via Visual Planning and Token Alignment
3. 学会等名 The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP) (国際学会)
4. 発表年 2022年

1. 発表者名 Hong Chen, Duc Minh Vo, Hiroya Takamura, Yusuke Miyao, Hideki Nakayama
2. 発表標題 StoryER: Automatic Story Evaluation via Ranking, Rating and Reasoning
3. 学会等名 The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP) (国際学会)
4. 発表年 2022年

1. 発表者名 Yuxuan Wu, Hideki Nakayama
2. 発表標題 Weakly Supervised Formula Learner for Solving Mathematical Problems
3. 学会等名 The 29th International Conference on Computational Linguistics (COLING) (国際学会)
4. 発表年 2022年

1. 発表者名 Jun Takeuchi, Noriki Nishida, Hideki Nakayama
2. 発表標題 Neural Networks in a Product of Hyperbolic Spaces
3. 学会等名 The 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop ( 国際学会 )
4. 発表年 2022年

1. 発表者名 Ryuichiro Hataya, Hideki Nakayama
2. 発表標題 DJMix: Unsupervised Task-agnostic Image Augmentation for Improving Robustness of Convolutional Neural Networks"
3. 学会等名 The 2022 International Joint Conference on Neural Networks (IJCNN) ( 国際学会 )
4. 発表年 2022年

1. 発表者名 Ikki Kishida, Hideki Nakayama
2. 発表標題 Pixel to Binary Embedding Towards Robustness for CNNs
3. 学会等名 The 26th International Conference on Pattern Recognition (ICPR) ( 国際学会 )
4. 発表年 2022年

1. 発表者名 Duc Minh Vo, Hong Chen, Akihiro Sugimoto, Hideki Nakayama
2. 発表標題 NOC-REK: Novel Object Captioning with Retrieved Vocabulary from External Knowledge
3. 学会等名 The 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) ( 国際学会 )
4. 発表年 2022年

1. 発表者名 Kai Katsumata, Duc Minh Vo, Hideki Nakayama
2. 発表標題 OSSGAN: Open-Set Semi-Supervised Image Generation
3. 学会等名 The 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (国際学会)
4. 発表年 2022年

1. 発表者名 勝又 海, Duc Minh Vo, 原田 達也, 中山 英樹
2. 発表標題 ラベルノイズ付きオープンセット半教師あり画像生成
3. 学会等名 第25回 画像の認識・理解シンポジウム
4. 発表年 2022年

1. 発表者名 Hideki Nakayama
2. 発表標題 Incorporating External Knowledge for Vision and Language Systems
3. 学会等名 2nd Workshop on Trends and Advances in Machine Learning and Automated Reasoning for Intelligent Robots and Systems (in conjunction with IROS 2022) (招待講演) (国際学会)
4. 発表年 2022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

1. 発表者名 Yuhao Xu and Hideki Nakayama
2. 発表標題 DCT-based Fast Spectral Convolution for Deep Convolutional Neural Networks
3. 学会等名 International Joint Conference on Neural Networks (IJCNN) (国際学会)
4. 発表年 2021年

1. 発表者名 Hong Chen, Raphael Shu, Hiroya Takamura, Hideki Nakayama
2. 発表標題 GraphPlan: Story Generation by Planning with Event Graph
3. 学会等名 International Conference on Natural Language Generation (INLG) (国際学会)
4. 発表年 2021年

1. 発表者名 Jan Zdenek and Hideki Nakayama
2. 発表標題 JokerGAN: Memory-Efficient Model for Handwritten Text Generation with Text Line Awareness
3. 学会等名 ACM International Conference on Multimedia (ACMMM) (国際学会)
4. 発表年 2021年

1. 発表者名 Hong Chen, Hiroya Takamura, Hideki Nakayama
2. 発表標題 SciXGen: A Scientific Paper Dataset for Context-Aware Text Generation
3. 学会等名 Conference on Empirical Methods in Natural Language Processing (EMNLP) (国際学会)
4. 発表年 2021年



1. 発表者名 Duc Minh Vo, Akihiro Sugimoto, Hideki Nakayama
2. 発表標題 PPCD-GAN: Progressive Pruning and Class-Aware Distillation for Large-Scale Conditional GANs Compression
3. 学会等名 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (国際学会)
4. 発表年 2022年

1. 発表者名 Ryuichiro Hataya, Jan Zdenek, Kazuki Yoshizoe, Hideki Nakayama
2. 発表標題 Meta Approach to Data Augmentation Optimization
3. 学会等名 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (国際学会)
4. 発表年 2022年

1. 発表者名 中山 英樹, 幡谷龍一郎
2. 発表標題 深層学習におけるデータ拡張の原理と最新動向
3. 学会等名 第27回画像センシングシンポジウム OS2: 続・限られたデータからの深層学習 (招待講演)
4. 発表年 2021年

1. 発表者名 Ryuichiro Hataya, Zdenek Jan, Kazuki Yoshizoe, Hideki Nakayama
2. 発表標題 Faster AutoAugment: Learning Augmentation Strategies using Backpropagation
3. 学会等名 European Conference on Computer Vision (ECCV) (国際学会)
4. 発表年 2020年

1. 発表者名 Taiga Kashima, Kento Masui, Hideki Nakayama
2. 発表標題 Unsupervised Visual Relationship Inference
3. 学会等名 IEEE International Conference on Image Processing (ICIP) (国際学会)
4. 発表年 2020年

1. 発表者名 Hong Chen, Yifei Huang, Hiroya Takamura, and Hideki Nakayama
2. 発表標題 Commonsense Knowledge Aware Concept Selection for Diverse and Informative Visual Storytelling
3. 学会等名 AAAI Conference on Artificial Intelligence (国際学会)
4. 発表年 2021年

1. 発表者名 Hong Chen, Yifei Huang, Hiroya Takamura, and Hideki Nakayama
2. 発表標題 Commonsense Knowledge Aware Concept Selection For Diverse and Informative Visual Storytelling
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 Ikki Kishida, Hideki Nakayama
2. 発表標題 Incorporating Horizontal Connections in Convolution by Spatial Shuffling
3. 学会等名 第23回 画像の認識・理解シンポジウム (MIRU)
4. 発表年 2020年

1. 発表者名 朱中元, Jason Lee, Kyunghyun Cho, 中山英樹
2. 発表標題 潜在変数の精緻化による非自己回帰型ニューラル機械翻訳
3. 学会等名 言語処理学会 第26回年次大会
4. 発表年 2020年

1. 発表者名 Ryuichiro Hataya, Hideki Nakayama
2. 発表標題 Unifying semi-supervised and robust learning by mixup
3. 学会等名 第22回 画像の認識・理解シンポジウム (MIRU2019)
4. 発表年 2019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 Raphael Shu, Jason Lee, Hideki Nakayama, Kyunghyun Cho
2. 発表標題 Latent-Variable Non-Autoregressive Neural Machine Translation with Deterministic Inference Using a Delta Posterior
3. 学会等名 The 34th AAAI Conference on Artificial Intelligence (AAAI 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Ryuichiro Hataya, Hideki Nakayama
2. 発表標題 LOL: Learning To Optimize Loss Switching Under Label Noise
3. 学会等名 The 26th IEEE International Conference on Image Processing (ICIP 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Yuhao Xu, Hideki Nakayama
2. 発表標題 DCT Based Information-Preserving Pooling for Deep Neural Networks
3. 学会等名 The 26th IEEE International Conference on Image Processing (ICIP 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Rui Yang, Hideki Nakayama
2. 発表標題 Bipolar Gan: Double Check the Solution Space and Lighten False Positive Errors in Generative Adversarial Nets
3. 学会等名 The 26th IEEE International Conference on Image Processing (ICIP 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Ikki Kishida, Hideki Nakayama
2. 発表標題 Empirical Study of Easy and Hard Examples in CNN Training
3. 学会等名 The 26th International Conference on Neural Information Processing (ICONIP 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Yuhao Xu, Hideki Nakayama
2. 発表標題 Shifted Spatial-Spectral Convolution for Deep Neural Networks
3. 学会等名 The 1st ACM International Conference on Multimedia in Asia (MMAAsia 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Ryuichiro Hataya, Hideki Nakayama
2. 発表標題 Unifying Semi-Supervised and Robust Learning by Mixup
3. 学会等名 The 2nd Learning from Limited Labeled Data (LLD) Workshop, ICLR 2019 (国際学会)
4. 発表年 2019年

〔図書〕 計1件

1. 著者名 中山 英樹、二反田 篤史、田村 晃裕、井上 中順、牛久 祥孝	4. 発行年 2022年
2. 出版社 サイエンス社	5. 総ページ数 248
3. 書名 深層学習からマルチモーダル情報処理へ	

〔産業財産権〕

〔その他〕

画像認識・自然言語処理に関する研究内容と成果 <a href="http://www.nlab.ci.i.u-tokyo.ac.jp/projects.html">http://www.nlab.ci.i.u-tokyo.ac.jp/projects.html</a>
---

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		
米国	University of California, Los Angeles	New York University	Amazon
その他の国・地域	National Yang Ming Chiao Tung Univ. (台湾)	Academia Sinica (台湾)	National Taiwan University (台湾)