

令和 5 年 6 月 9 日現在

機関番号：11301

研究種目：基盤研究(A)（一般）

研究期間：2019～2022

課題番号：19H01110

研究課題名（和文）深層学習がもたらすパラダイムシフトを加速するネットワーク設計理論の構築

研究課題名（英文）Developing a Design Methodology of Deep Neural Networks to Accelerate Paradigm Shift Brought by Deep Learning

研究代表者

岡谷 貴之（Okatani, Takayuki）

東北大学・情報科学研究科・教授

研究者番号：00312637

交付決定額（研究期間全体）：（直接経費） 32,800,000円

研究成果の概要（和文）：深層学習では一般に、解きたいタスクに応じてネットワークのアーキテクチャをうまく設計する必要があるが、設計の方法論や指針が確立されていなかった。様々なタスクに対し、それぞれにおいて高い性能を達成するアーキテクチャの設計に取り組むことを通じて、この課題の解消を目指した。画像復元、画像理解、3次元幾何学推定、推論の不確かさ推定、自己教師特徴学習などの諸問題に対し、最高精度（論文発表当時）を達成するネットワークの開発に成功した。それらは、対象としたタスクそれぞれにおいて大きな成果となるとともに、すべてを合わせることで、ネットワーク設計の方法論の基盤となる知見を得ることができた。

研究成果の学術的意義や社会的意義

深層学習は、近年発展著しい人工知能の中核技術であるとともに、その他の工学やサイエンスにも大きな影響を与えつつある。その一方で、深層ニューラルネットワークの構造設計に確たる方法論がないという課題があった。本研究は、様々な応用ごとに優れた性能を発揮するネットワーク構造の研究を通じて、それぞれの応用問題の解決に貢献するとともに、ネットワーク構造に関する新たな知見を多く生み出した。これらの成果は、構造設計の方法論の基盤を与えている。

研究成果の概要（英文）：Deep learning requires developers to design a network architecture that is appropriate for each individual task that they want to solve, but no methodology or guidelines have been established for designing good architectures. We aimed to solve this issue by designing architectures that achieve high performance in various tasks. We have successfully developed networks that achieve the highest accuracy (at the time of publication) for various tasks, including image restoration, image understanding, 3D geometry estimation, uncertainty estimation, and self-supervised feature learning. While the results have significant impact for each of the targeted tasks, their integration provide a foundation toward establishing the methodology of neural architectural design.

研究分野：コンピュータビジョン

キーワード：深層学習 コンピュータビジョン

1. 研究開始当初の背景

研究構想時、深層学習の発展によって人工知能(AI)は大きな発展を遂げていた。画像認識、音声認識、機械翻訳などで部分的に人に匹敵する性能を達成し、実用化も進んでいた。さらにその成功は、AIの範囲を大きく超え、画像・音声の合成、幾何学計算(後述)、あらゆる信号処理、化合物反応予測や流体シミュレーションなどに広がりつつあった。最初の成功例である音声・画像認識の場合には、DNNが内部で行う計算は従来の意味の「パターン認識」として解釈可能であったが、応用範囲の拡大に伴い、このような捉え方では説明ができない事例が次々に登場していた。「数理・物理モデリング+逆問題を解くアルゴリズム構築」という従来の方法論が、深層学習に凌駕されつつあった。広く工学的な問題解決の方法論を一変させるパラダイムシフトが起こりつつあった。深層学習の強みは、従来、手で設計していた入力データからの特徴抽出過程を、データを用いた学習により獲得できることであると説明されてきたが、現実には、ネットワーク構造をどのようにデザインするかも同じくらい重要であった。深層学習の発展の本質は、より良いネットワーク構造の発見にあると言っても過言でなかった。それぞれの問題に適した構造があり、構造によって性能が大きく左右される。申請者らは、深層学習が成功した理由の一つは、このようなネットワーク構造のデザインを通じて、対象とする問題・データが持つ固有の構造をうまく捉えられることにあると考えた。しかし、この新しいパラダイムには次の2つの課題があり、さらなる発展の障害となっていると考えた。

課題1: ネットワークの構造デザインに現状で系統的な方法論がない。対象とする問題ごとに膨大な試行錯誤が必須であり、効率が悪い。

課題2: 一般に大量のデータを必要とする。多くの問題が、学習データを十分用意できない現実に阻まれ、深層学習の適用に至っていない。このパラダイムの最大のボトルネックであると言え、社会により大きな変革をもたらすためには、その解消が必須である。

2. 研究の目的

研究開始時点で、上述の通り深層学習は、AIの範囲を大きく超えて多様な分野で活用され始め、工学の問題解決の方法論を一変させるパラダイムシフトをもたらしつつあった。新たな方法論では、問題解決は、学習データを用意することと、問題に適したネットワーク構造をうまくデザインすることで達成される。課題となっていたのは、ネットワーク構造のデザインに系統的な方法論がなく、試行錯誤に頼らざるを得ないことであった。本計画では、ネットワーク内部の計算内容を解釈する方法の実現や、特定のタスクに対し学習を行わずデザインのみでネットワークを完成させる問題に取り組むことで、深層ネットワーク設計理論を構築する。本計画では次の2つのゴールを設定した:

ゴール1: ネットワーク構造デザインの方法論の確立

ゴール2: データを用いた学習とネットワーク構造の相互関係の明確化

具体的に目標としたのは: ア)対象とする問題に対し最適な構造のネットワークを現在よりも簡単に構築でき、イ)より少ないデータで学習が可能となり、ウ)既存のネットワークの計算アルゴリズムの解釈ができるようになること、であった。

3. 研究の方法

以下に述べる方法によって研究を実施した。

複数の色々な種類のタスクを選定し、それぞれに合わせたDNNのアーキテクチャ及びコンポーネントを設計し、その働きを理解することを目指した。そこでは既存手法を上回る性能を追求することを優先的に考え、性能の最大化を通じて、タスクごとにどういったアーキテクチャーが必要かを理解することを目指した。分野での研究動向、実社会での応用の広がりも合わせて考えて、次のタスク群を選定した。

- (1) 画像復元タスク
- (2) 画像理解タスク
- (3) 3次元幾何推論タスク
- (4) 画像特徴の学習

以上の取り組みを通じて、特定タスクに対するDNNの完全設計、及び学習済みDNNの計算論的解釈を狙った。

4. 研究成果

(1) 画質復元のためのネットワーク詳細構造

画質改善の多様なタスクを対象に、最適なDNNの構造を研究し、それを通じて構造設計に関する知見を蓄積することを目指した。画質を劣化させる要因は多様であり、画質を効果的に改善可能なネットワークの構造は、画質劣化要因ごとに異なると考えられることが背景にあった。以下、われわれの取り組みと成果を述べる。

画像復元のタスク集合、具体的にはノイズ除去、モーションブラー除去、降雨除去、ヘイズ除

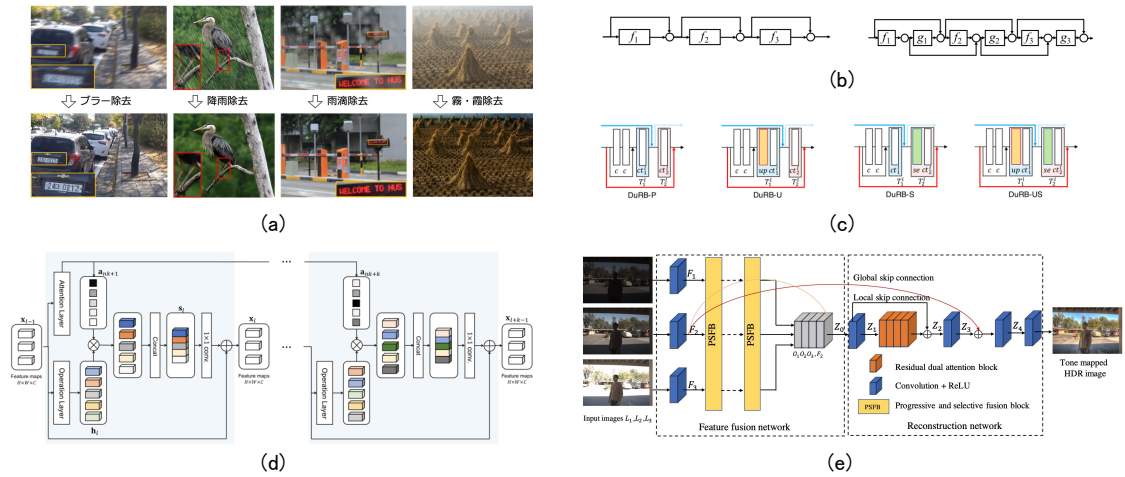


図 1. 画像復元に関する研究成果の例. (a) 異なる要因で劣化した画像の画質を改善. (b) 二重残差ネットワークの構造. 左は通常の残差ネットワーク. 右が提案構造. (c) 劣化要因ごとに最適化された二重残差ネットワーク内部のコンポーネントの構造. (d) 複合的な要因で劣化した画像の画質を改善するネットワーク. (e) 3 枚の露出を変えて撮影した動的シーンの画像から 1 枚のハイダイナミックレンジ画像(HDR)を得るネットワーク.

去などを対象に、高い画像復元精度を達成するネットワーク構造は、それぞれの画質劣化要因に対し、どのようなものであるべきかを研究した。その結果、画像復元用ネットワークの新たな基本構造「二重残差ネットワーク (Dual Residual Network)」を開発し (図 1b, 1c)。多くの画質改善タスクにおいて、当時の最高性能を達成した (図 1a, Liu et al. 2019)。

さらに、複数の画質劣化要因が混ざり合い、しかもその割合が不明な場合の画質改善問題に取り組んだ。ネットワーク内に複数の演算を並置し、入力画像に応じて必要なものを注意機構によって選択できる構造を開発した (図 1d, Suganuma et al. 2019)。このタスクにおける当時の最高性能を達成した。この研究は、後の研究ではスタンダードの一つとなった。

さらに、High-Dynamic Range Imaging (HDR), すなわち露出を変化させた複数画像を入力に、高ダイナミックレンジ画像を生成する問題に対し、ネットワーク構造および学習方法を開発した (図 1e, Ye et al. 2021)。当時の最高精度を達成した。

以上の研究を通じて、ネットワークの新しい構造要素を開発し、それらをタスクの性質に応じてどう使い分けるべきかについて、数多くの知見を得た。

(2) ネットワークの学習・推論内容の解釈・可視化手法

単眼深度推定, すなわちシーンの画像 1 枚からその奥行きを推定するタスクについて、これを学習した DNN を対象に、その各部と全体で行われている計算アルゴリズムの解釈を目標に、入力画像のどこが推論に用いられているかを可視化する方法を開発した (Hu et al. 2019)。入力画像から推定に影響を与えない領域をマスクする方法で、しかもそのマスクを入力画像そのものから別のネットワークで予測する方法である (図 2a)。人の視覚の研究では、人が単眼深度推定を行う際、シーンの画像に含まれる複数の手がかりを効率的に組み合わせていることが知られている。ネットワークにおいてもそれらに近い手がかりを用いているらしいことが示唆される結果を得た (図 2b, c)。

また、動画から物体を検出するタスクを対象に、ネットワークの内部の計算を解釈する問題に取り組み、瞬時的に検出に失敗する現象がどのようなメカニズムで発生するのか、またそれを改善するにはどうすべきかを明らかにした (Hosoya et al. 2020)。

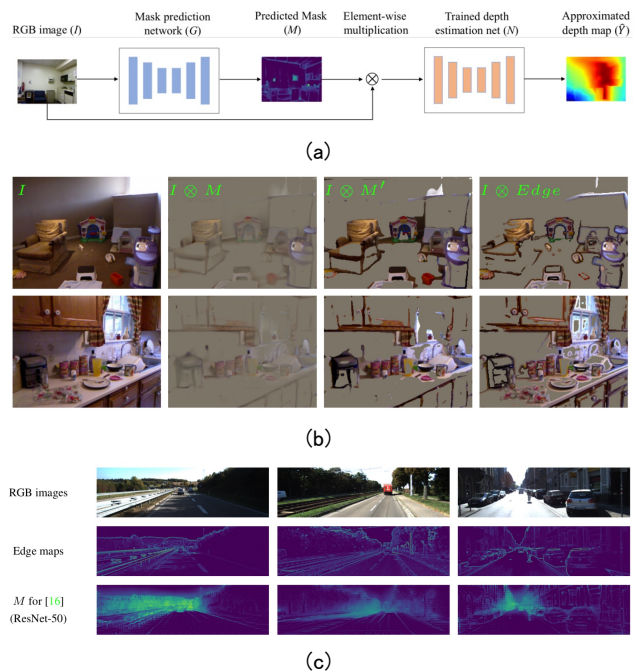


図 2. ネットワークが画像のどこに注目して奥行きを推定しているかを可視化する手法. (a) ネットワーク構造. (b) 室内画像の例. (c) 車載画像の例.

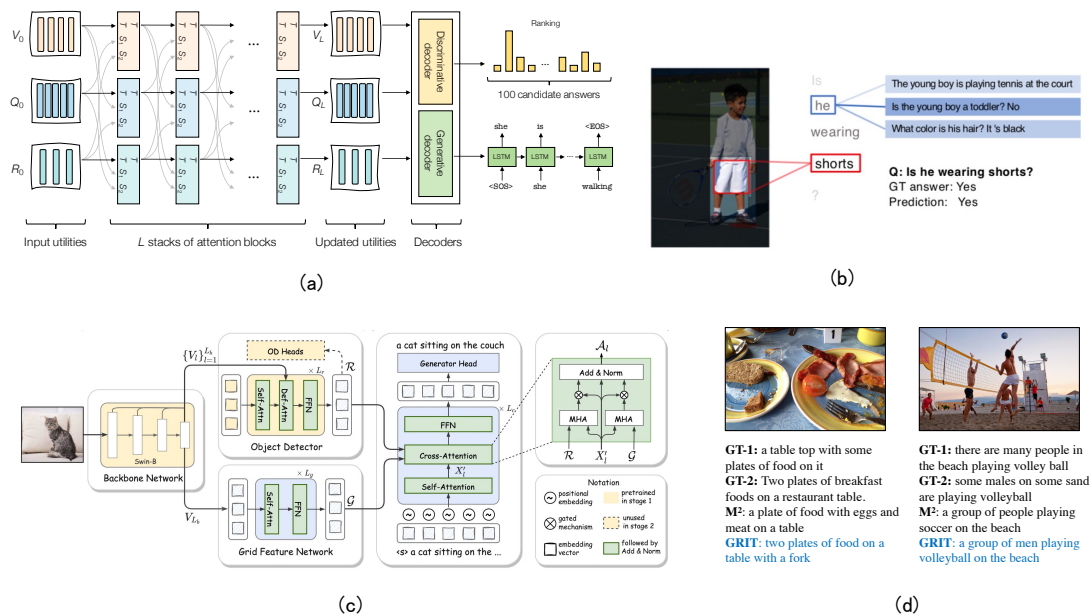


図 3. 画像理解 (画像と言語のマルチモーダルタスク). (a) 視覚対話のために開発した DNN の構造. (b) 視覚対話の途中経過の例. (c) 画像記述のために開発した DNN の構造. (d) 記述例 (GT: 人による記述, GRIT: 提案手法).

(3) 画像と言語のマルチモーダルタスクに対するネットワーク構造設計

深層学習及び AI の発展に伴い、注目を集めている問題に、画像理解がある。これは画像と言語の両方を扱えるマルチモーダル AI であり、その DNN アーキテクチャの最適な設計については、多くの議論があった。

われわれは、人と AI (DNN) が 1 枚の画像に写るシーンについて、自然言語を用いた対話する「視覚対話」を対象に、高い性能を達成する DNN アーキテクチャの構造を研究した。トランスフォーマーをベースとしつつ、その非効率的な部分を効率化した構造を新規に開発し、それを元にした DNN の全体構造を設計した (図 3a)。この DNN は当時の最高性能を達成した (図 3b)。

さらに画像記述のための DNN アーキテクチャを新たに設計し、GRIT と命名した。GRIT の技術的な核は 2 つあり、1 つは画像特徴の抽出方法である。GRIT は入力画像から、物体検出を行なって得た物体の各領域から取り出した画像特徴と、画像全体から取り出した特徴を融合して用いる。さらに従来手法では画像特徴抽出は CNN を用いることが多かったが、全てをトランスフォーマーで設計し、記述生成部分とあわせて全てがトランスフォーマーからなるデザインとした。もう 1 つの核は、その学習方法にある。上のデザインにより、推論に要する計算量が小さく抑えられ、また入力画像から生成する記述までを一気通貫で学習できるようになったことで、画像特徴の抽出性能が向上し、記述の精度が向上した。標準ベンチマークテスト COCO captions において、世界トップクラスの記述精度を、従来手法の 20 分の 1 の学習データしか用いることなく達成しつつ、推論に要する計算時間を 5 分の 1 程度まで短縮した (Nyugen et al. 2022)。

また、理解に基づいて行動までを可能な AI として、自然言語の指示によって作業を行う AI エージェントを実現する DNN を新たに開発した。エージェントが空間を見る視野をなるべく広く取るとともに、作業の指示を与える言語情報を 2 回解釈するアーキテクチャとなっており、これによって高い作業成功率の達成が可能となった (Nguyen et al. 2021)。

(4) 推論の不確かさの予測方法

DNN は様々な推論を実行できるが、その推論が正しいのか間違いなのか、すなわち「推論の不確かさ (uncertainty)」を、推論結果そのものから予測することは、深層学習の根本に関わる問題であり、応用上の価値が大きい。この問題を研究し、クラス分類にコサイン類似度を用いたロジスティック回帰を行う構造をネットワークに持たせ、学習を行うことで、学習外分布 (out-of-distribution) の入力を高精度に検出できることを実験的に示した (Techupanurak et al. 2020)。当時の最高性能を達成した。

(5) 3次元幾何学推定のためのネットワーク構造・学習方法

視覚 SLAM (Simultaneous Localization And Mapping)、ロボットや車両搭載カメラからの自己位置推定・地図生成の問題) を対象に、DNN の最適構造やその学習方法の研究に取り組み、以下に述べる成果を挙げた。

まず、視覚 SLAM で必要となる、多視点画像に対する幾何学的な最適化計算 (バンドル調整) を、グラフネットワークを用いた深層学習によって実行する方法を開発した (Tanaka et al. 2021)。バンドル調整は通常、推定すべき未知数である、多視点カメラの姿勢と空間の点群の 3 次元座標を、画像上の観測と整合するようにニュートン法によって最適化する方法である (図 4a)。このニュートン法の 1 ステップのパラメータの増分を、学習によってグラフニューラルネ

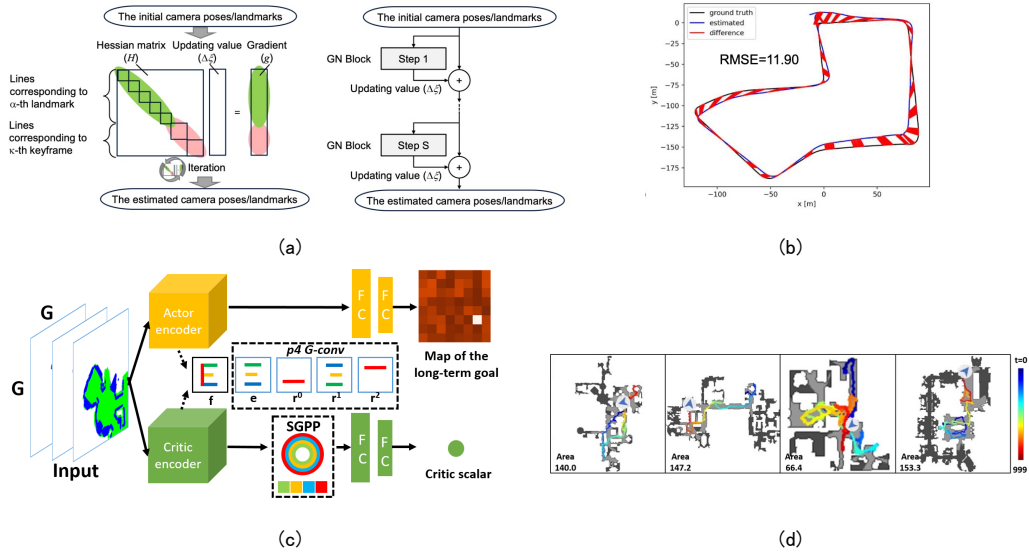


図 4. (a) バンドル調整を深層学習で実行. (b) 視覚 SLAM に組み込んだ時の移動軌跡の推定結果. (c) 視覚等センサ情報を入力に屋内空間を探索するタスクのためのネットワークアーキテクチャ. (d) 屋内空間の探索結果.

ネットワークによって予測する手法である。これは世界初の成果であり、最適化問題へのグラフネットワークの設計・適用方法の新たな可能性を示すものとなっている (図 4b)。

同じく視覚 SLAM や Structure-from-motion の適用領域を、映像センサの能力の下限付近にある低照度シーンでも利用可能にする枠組みを新たに提案した (Song et al. 2021)。同分野の研究者が今後の研究に活用できるように、データセットを構築し公表した。

さらに、屋内空間の探索問題を対象に、入力の幾何学的変換に対する不変性に着目したネットワークのデザイン手法を開発した (Liu et al. 2022)。タスクに必要な変換を見極めた上で、それらへの不変性を構造のみで達成するネットワークを設計する方法で、屋内環境を自律的に探索するタスクに適用し、有効性を示した (図 4c, d)。研究テーマの 1 つである「ネットワークの構造デザインとデータ量の関係」に新たな知見をもたらした。

(6) 自己教師内部表現学習に関する理論研究

教師学習は正解ラベルを必要とし、それは人が通常与えるものであるため、学習を大規模化するのに問題があった。これに対し、自己教師学習は、様々な推論に適した入力 (例えば画像) の特徴を、人が作った正解ラベルを必要とせず、画像のみから学習できる新たなアプローチである。近年、大きく研究が進展した。自己教師学習手法には、正例と負例を用いる対照学習と、負例を用いない非対照学習がある。後者の手法は、明示的な制約がなければ無意味な解に縮退し得るはずだが、BYOL (Grill et al. 2020, 図 5a) や SimSiam (Chen et al. 2020, 図 5a) などの既存手法は、縮退せず、有効な特徴空間を学習できる。その理由が解明されていなかった。われわれは、既に理解が進んでいた Barlow-Twins (Zbontar et al. 2021, 図 5b) や VICReg (Bardes et al. 2021, 図 5c) などの「特徴量が無相関にする制約付きの手法」とこれらの既存手法の間に、これまで知られていなかった関係があることを見出し、それを通じて BYOL/SimSiam がなぜ良い特徴空間を学習できるかが説明できることを示した (図 5d, Liu et al. 2022)。掲げた目標の 1 つ「データを用いた学習とネットワーク構造の相互関係の明確化」に対する一つの答えでもある。

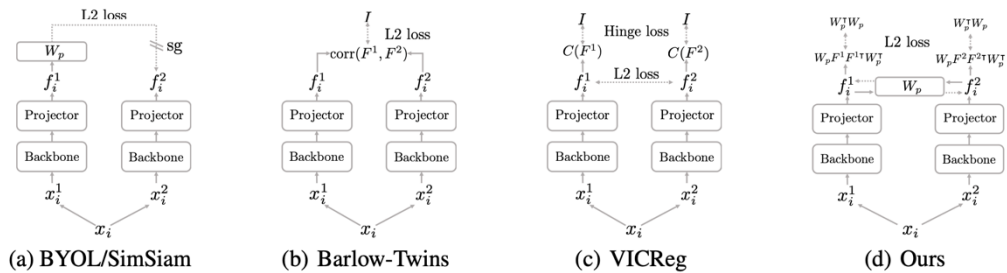


図 5. 非対照自己教師学習のネットワークアーキテクチャ. (a) BYOL/SimSiam. (b) Barlow-Twins. (c) VICReg. (d) 提案手法.

5. 主な発表論文等

〔雑誌論文〕 計27件（うち査読付論文 26件 / うち国際共著 5件 / うちオープンアクセス 23件）

1. 著者名 Zhijie Wang, Masanori Suganuma, Takayuki Okatani	4. 巻 -
2. 論文標題 Unsupervised Domain Adaptation for Semantic Segmentation via Cross-Region Alignment	5. 発行年 2023年
3. 雑誌名 Computer Vision and Image Understanding	6. 最初と最後の頁 採録済
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 Liu Shuang, Okatani Takayuki	4. 巻 -
2. 論文標題 Symmetry-aware Neural Architecture for Embodied Visual Exploration	5. 発行年 2022年
3. 雑誌名 Proceedings of Computer Vision and Pattern Recognition	6. 最初と最後の頁 17221-17230
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR52688.2022.01673	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 Kang-Jun Liu, Masanori Suganuma, Takayuki Okatani	4. 巻 -
2. 論文標題 Bridging the Gap from Asymmetry Tricks to Decorrelation Principles in Non-contrastive Self-supervised Learning	5. 発行年 2022年
3. 雑誌名 Advances in Neural Information Processing Systems 35 (NeurIPS 2022)	6. 最初と最後の頁 19824-19835
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 Zhu Yanjie, Sekiya Hidehiko, Okatani Takayuki, Yoshida Ikumasa, Hirano Shuichi	4. 巻 12
2. 論文標題 Real-time vehicle identification using two-step LSTM method for acceleration-based bridge weigh-in-motion system	5. 発行年 2022年
3. 雑誌名 Journal of Civil Structural Health Monitoring	6. 最初と最後の頁 689 ~ 703
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s13349-022-00576-2	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 Sharma Parmanand, Ninomiya Takahiro, Omodaka Kazuko, Takahashi Naoki, Miya Takehiro, Himori Noriko, Okatani Takayuki, Nakazawa Toru	4. 巻 12
2. 論文標題 A lightweight deep learning model for automatic segmentation and analysis of ophthalmic images	5. 発行年 2022年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 1-18
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-022-12486-w	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 Suzuki-Ohno Yukari, Westfechtel Thomas, Yokoyama Jun, Ohno Kazunori, Nakashizuka Tohru, Kawata Masakado, Okatani Takayuki	4. 巻 12
2. 論文標題 Deep learning increases the availability of organism photographs taken by citizens in citizen science programs	5. 発行年 2022年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 1-10
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-022-05163-5	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 Wenzheng Song, Masanori Suganuma, Xing Liu, Noriyuki Shimobayashi, Daisuke Maruta, Takayuki Okatani	4. 巻 -
2. 論文標題 Matching in the Dark: A Dataset for Matching Image Pairs of Low-light Scenes	5. 発行年 2021年
3. 雑誌名 Proceedings of International Conference on Computer Vision 2021	6. 最初と最後の頁 6009-6018
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICCV48922.2021.00597	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 Qian Ye, Jun Xiao, Kin-Man Lam, Takayuki Okatani	4. 巻 -
2. 論文標題 Progressive and Selective Fusion Network for High Dynamic Range Imaging	5. 発行年 2021年
3. 雑誌名 Proceedings of ACM Multimedia 2021	6. 最初と最後の頁 5290-5297
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3474085.3475651	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 Tetsuya Tanaka, Yukihiro Sasagawa, Takayuki Okatani	4. 巻 -
2. 論文標題 Learning to Bundle-adjust: A Graph Network Approach to Faster Optimization of Bundle Adjustment for Vehicular SLAM	5. 発行年 2021年
3. 雑誌名 Proceedings of International Conference on Computer Vision?2021	6. 最初と最後の頁 6320-6329
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICCV48922.2021.00619	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 Van-Quang Nguyen, Masanori Suganuma, Takayuki Okatan	4. 巻 -
2. 論文標題 Look Wide and Interpret Twice: Improving Performance on Interactive Instruction-following Tasks	5. 発行年 2021年
3. 雑誌名 Proceedings of 30th International Joint Conference on Artificial Intelligence (IJCAI-21)	6. 最初と最後の頁 923-930
掲載論文のDOI (デジタルオブジェクト識別子) 10.24963/ijcai.2021/128	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 該当する

1. 著者名 Engkarat Techapanurak, Masanori Suganuma, Takayuki Okatani:	4. 巻 -
2. 論文標題 Hyperparameter-Free Out-of-Distribution Detection Using Cosine Similarity	5. 発行年 2020年
3. 雑誌名 Proceedings of Asian Conference on Computer Vision	6. 最初と最後の頁 53-69
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-69538-5_4	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 Yoshihiro Hirohashi, Kenichi Narioka, Masanori Suganuma, Xing Liu, Yukimasa Tamatsu, Takayuki Okatani	4. 巻 -
2. 論文標題 Removal of Image Obstacles for Vehicle-mounted Surrounding Monitoring Cameras by Real-time Video Inpainting	5. 発行年 2020年
3. 雑誌名 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops	6. 最初と最後の頁 857-866
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPRW50498.2020.00115	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 Van-Quang Nguyen, Masanori Suganuma, Takayuki Okatani	4. 巻 -
2. 論文標題 Efficient Attention Mechanism for Visual Dialog that Can Handle All the Interactions Between Multiple Inputs	5. 発行年 2020年
3. 雑誌名 Proceedings of European Conference on Computer Vision	6. 最初と最後の頁 223-240
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-58586-0_14	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 Yan Zhang, Junjie Hu, Takayuki Okatani	4. 巻 399
2. 論文標題 Extending information maximization from a rate-distortion perspective	5. 発行年 2020年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 285-295
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neucom.2020.02.061	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 Yusuke Hosoya, Masanori Suganuma, Takayuki Okatani	4. 巻 -
2. 論文標題 Analysis and a Solution of Momentarily Missed Detection for Anchor-based Object Detectors	5. 発行年 2020年
3. 雑誌名 Proceedings of IEEE Winter Conference on Applications of Computer Vision	6. 最初と最後の頁 1399-1407
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV45572.2020.9093553	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 Pongsate Tangseng, Takayuki Okatani	4. 巻 -
2. 論文標題 Toward Explainable Fashion Recommendation	5. 発行年 2020年
3. 雑誌名 Proceedings of IEEE Winter Conference on Applications of Computer Vision	6. 最初と最後の頁 2142-2151
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV45572.2020.9093367	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 Xing Liu, Masanori Suganuma, Takayuki Okatani	4. 巻 -
2. 論文標題 Joint Learning of Multiple Image Restoration Tasks	5. 発行年 2020年
3. 雑誌名 arXiv.org	6. 最初と最後の頁 1-10
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 無
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 Date Hiroto, Kawasaki Keisuke, Hasegawa Isao, Okatani Takayuki	4. 巻 -
2. 論文標題 Deep Learning for Natural Image Reconstruction from Electrocorticography Signals	5. 発行年 2019年
3. 雑誌名 Proceedings of BIBM	6. 最初と最後の頁 2331 ~ 2336
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/BIBM47256.2019.8983029	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 該当する

1. 著者名 Liu Xing, Suganuma Masanori, Sun Zhun, Okatani Takayuki	4. 巻 -
2. 論文標題 Dual Residual Networks Leveraging the Potential of Paired Operations for Image Restoration	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE Computer Vision and Pattern Recognition	6. 最初と最後の頁 7007 ~ 7016
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR.2019.00717	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 Suganuma Masanori, Liu Xing, Okatani Takayuki	4. 巻 -
2. 論文標題 Attention-Based Adaptive Selection of Operations for Image Restoration in the Presence of Unknown Combined Distortions	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE Computer Vision and Pattern Recognition	6. 最初と最後の頁 9039 ~ 9048
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR.2019.00925	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 Nguyen Duy-Kien, Okatani Takayuki	4. 巻 -
2. 論文標題 Multi-Task Learning of Hierarchical Vision-Language Representation	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE Computer Vision and Pattern Recognition	6. 最初と最後の頁 10492 ~ 10501
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR.2019.01074	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 Shao Mingzhen, Sun Zhun, Ozay Mete, Okatani Takayuki	4. 巻 -
2. 論文標題 Improving Head Pose Estimation with a Combined Loss and Bounding Box Margin Adjustment	5. 発行年 2019年
3. 雑誌名 Proceedings of FG	6. 最初と最後の頁 1 ~ 5
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/FG.2019.8756605	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 Hu Junjie, Zhang Yan, Okatani Takayuki	4. 巻 -
2. 論文標題 Visualization of Convolutional Neural Networks for Monocular Depth Estimation	5. 発行年 2019年
3. 雑誌名 Proceedings of International Conference on Computer Vision	6. 最初と最後の頁 3868~3877
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICCV.2019.00397	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 Liu Shuang, Ozay Mete, Xu Hongli, Lin Yang, Okatani Takayuki	4. 巻 -
2. 論文標題 A Generative Model of Underwater Images for Active Landmark Detection and Docking	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems	6. 最初と最後の頁 8034 ~ 8039
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/IRoS40897.2019.8968146	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 Hosoya Yusuke, Suganuma Masanori, Okatani Takayuki	4. 巻 -
2. 論文標題 Analysis and a Solution of Momentarily Missed Detection for Anchor-based Object Detectors	5. 発行年 2020年
3. 雑誌名 Proceedings of IEEE Winter Conference on Applications of Computer Vision	6. 最初と最後の頁 1399 ~ 1407
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV45572.2020.9093553	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 Tangseng Pongsate, Okatani Takayuki	4. 巻 -
2. 論文標題 Toward Explainable Fashion Recommendation	5. 発行年 2020年
3. 雑誌名 Proceedings of IEEE Winter Conference on Applications of Computer Vision	6. 最初と最後の頁 2142 ~ 2151
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV45572.2020.9093367	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 Zhang Yan, Hu Junjie, Okatani Takayuki	4. 巻 399
2. 論文標題 Extending information maximization from a rate-distortion perspective	5. 発行年 2020年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 285 ~ 295
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neucom.2020.02.061	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 -

〔学会発表〕 計7件 (うち招待講演 6件 / うち国際学会 0件)

1. 発表者名 岡谷貴之
2. 発表標題 画像を理解し、その内容を人と共有できるAIの実現へ向けて 3. 学会等名 土木学会講演会 (招待講演) 4. 発表年 2022年 1. 発表者名 岡谷貴之 提出
3. 学会等名 土木学会講演会 (招待講演)
4. 発表年 2022年

1. 発表者名 岡谷貴之
2. 発表標題 画像を扱うAI(深層学習)に関する研究
3. 学会等名 仙台X-Tech (招待講演)
4. 発表年 2022年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の現在と近未来：深奥質感からAIの今後を考える
3. 学会等名 新道路成果報告会AI活用 (招待講演)
4. 発表年 2021年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の現在:問題解決の方法論として
3. 学会等名 日本天文学会 (招待講演)
4. 発表年 2021年

1. 発表者名 岡谷貴之
2. 発表標題 ディープラーニングの課題: 現場からフロンティアまで
3. 学会等名 日本機会学会 (招待講演)
4. 発表年 2021年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習(AI)の現在と近い将来
3. 学会等名 精密工学会(招待講演)
4. 発表年 2021年

1. 発表者名 岡谷貴之
2. 発表標題 言語による画像理解とそれに基づく行動の生成まで
3. 学会等名 CRESTセミナー
4. 発表年 2021年

〔図書〕 計1件

1. 著者名 岡谷 貴之	4. 発行年 2022年
2. 出版社 講談社	5. 総ページ数 384
3. 書名 深層学習 改訂第2版	

〔産業財産権〕

〔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
研究 分担者	菅沼 雅徳 (Suganuma Masanori) (00815813)	東北大学・情報科学研究科・助教 (11301)	

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			
中国	Hong Kong Polytechnic University			