

令和 3 年 6 月 11 日現在

機関番号：11301

研究種目：新学術領域研究(研究領域提案型)

研究期間：2015～2019

課題番号：15H05919

研究課題名(和文)画像と言語を用いた質感情報表現のディープラーニング

研究課題名(英文)Deep Representation Learning of Shitsukan Using Vision and Language

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) 45,700,000円

研究成果の概要(和文):画像1枚からそこに写る物の質感を認識するAIの実現という目標に一歩ずつ近づく段階的な成果を挙げた。具体的には、CNNを用いた画像ペアのランキング学習による質感認識手法、ウェブ上の画像とその説明文のペアから視覚的に認識可能な質感概念をマイニングする方法、質感認識の文脈依存性を捉えつつ商品の画像に適切なタイトルを生成する方法などである。さらに研究の過程で、質感認識は「包括的な画像認識である」と結論するに至り、今日の画像理解のためのマルチモーダル深層ニューラルネットワークの端緒となる、注意機構を用いたネットワークの開発に成功した。

研究成果の学術的意義や社会的意義

人と同じように物の質感を視覚的に認識するAIの実現は、最近のAI技術の進展を踏まえても難しかった。質感は人にも言語化が難しく、AIが学習すべき教師データを作るのが難しかったことが大きな理由である。本研究は段階的にこの困難さを解決する方法を与えた。AIによる質感認識が難しいことは現在のAIの人とのギャップの1つであると言え、本研究はこのギャップを狭めることに成功し、これはAIそのものの発展に貢献する学術的な意義である。社会的な意義には、以前は人にしかできなかった質感認識に関わる作業を代替することを初めとする工学的応用がある。

研究成果の概要(英文):We have obtained step-by-step achievements towards realizing an AI that can recognize the Shitsukan of an object from its single image. We have developed the following methods: i) a CNN-based method for Shitsukan recognition based on learning to rank image pairs; ii) a method for mining visually recognizable Shitsukan concepts from pairs of images and their descriptions on the Web; iii) a method for generating appropriate titles for images of products while capturing the context-aware Shitsukan concepts. The developments of the above lead to the conclusion that Shitsukan recognition is "comprehensive image recognition," and we successfully developed a network with an attention mechanism, which is an offshoot of today's multimodal deep neural networks for image understanding.

研究分野:画像認識,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 質感 画像と言語

1. 研究開始当初の背景

本計画の主要な目標は、物の質感を画像から認識するシステムを構築することである。研究開始時、それ以前の検討を踏まえてわれわれは、その鍵は質感の概念についての画像と言語の対応を明らかにすることにあると判断していた。その点で、画像記述(=画像の内容を言語で説明する技術)に関する当時の研究——人工知能への期待を背景に画像記述の研究が活発化していた——と関連が深いと考えられた。それらは自然言語処理の研究が主導しており、主要な関心は、単語や語句の「特徴ベクトル」をコーパス内の共起性に基づいて定義し、画像と関係付ける方法に向かっていた。しかしながら質感とは、人でもしばしば言語化が困難な、言葉(意味)を超えた心理的な表象である。コーパスから得られる情報は極めて限られており、むしろ画像を基盤とする方法に可能性があると考えられた。さらに質感の場合、大規模な学習データを作ることが本質的に難しいという問題があった。計画開始時期には、CNNによる物体カテゴリ認識が関連分野で盛んに研究されており、百万枚オーダーの(質感がアノテートされた)画像があれば、単純な教師あり学習が使えることは明らかとなりつつあった。しかしながら質感は、言語化しづらく個人差や主観の排除も難しいことから、現実のデータセットは小規模かつ不完全なものとならざるを得ない。この困難さの解決が必要となっていた。

2. 研究の目的

画像からそこに写る物の質感をコンピュータ上で認識するシステムを実現することを目的とする。大規模データを用いた畳み込みニューラルネット(以下CNN)の学習により、質感情報を表現する空間(以下質感表現空間)を獲得する方法を研究した。本計画で言う質感表現空間とは、画像の写像先となる空間に言語(形容詞などの単語)が埋め込まれたものを指す。質感は言語化しづらく、それゆえに人と人、あるいは人と機械の間での共有・伝達が難しいという性質がある。例えば「漆のような黒」や「ダイヤモンドの透明感」は、様々に解釈が成り立ち得る。これを踏まえ、画像に写る物の質感を「数値化」し、万人が共有できるような「質感の標準化」を可能にすることを目標とした。このようなことが実現できれば、工学的な利益は大きい。質感を数値化・可視化することで、工業デザインの効率化(=製品の試作や実物目視の必要性を排除)を図ることができる。また、ウェブの検索サイト等での情報取得や、商品推薦の精度を向上させることもできるだろう。人が感じとる質感因子の成分が明らかにされることで、画像(に写る物)の質感を制御することや、映像伝送時の情報圧縮への応用も期待できる。

3. 研究の方法

以下に述べるような方法で研究を実施した。

大規模学習システムの構築

研究開始時にCNNの大規模学習に使用する計算機システムを構築し、このシステム上で深層学習専用のソフトウェアプラットフォームを運用可能とした。

質感データセットの構築

質感語がアノテートされた画像データからなる質感データセット——質感語の表現空間への埋め込みに用いるもの——を構築した。

多層畳み込みニューラルネット(CNN)の内部表現分析

物体認識を大量の学習サンプルを用いて訓練した多層CNNが、どういった特徴を取り出し、どのように表現しているかを分析した。具体的には本計画では、物体認識の他、テキスト認識、マテリアル認識等の異なるタスクを対象に、特徴空間のクラスタリング等を行うことで、その構造を分析した。

質感形容詞のランキング学習の高精度化と評価方法の開発

質感の知覚には個人差が存在する。例えば2枚の画像ペアを与え、どちらがより「滑らか」に感じるかと複数の被験者に尋ねた場合、その答えは被験者によって分かれることがある。このこれを踏まえて人と比べたとき、CNNの性能を正しく評価する方法を研究した。

多層CNNとサル視覚皮質の比較分析

サル視覚野の脳皮質電位図と多層CNN間の関係を調べた。大量の材質および物体の画像をサルに提示し、記録したデータが手元にあり、これと多層CNNの中間層出力間の関係を、回帰や相関分析によって調べ、特に時間情報の使われ方について、新しい知見を得ることを目指した。以上の研究を進めることで、言葉で言い表せないがゆえに人と人、あるいは人と機械の間で伝達が難しかった質感を「数値化」する方法の実現を狙った。サル視覚野の脳皮質電位図と多層CNNの関係を調べた。大量の材質および物体の画像をサルに提示し、脳活動を記録するとともにその

分析を行なった。時間・脳皮質表面の空間それぞれについての脳活動の因果性に注目し、それと CNN 内での情報処理の関係を分析した。

深層学習を用いた画質改善手法

質感認識に限らないが、CNN と人の視覚の 1 つの違いとして、画像のノイズに対する耐性がある。綺麗な画像のみで訓練された CNN は、ノイズが画像に混入すると精度を大きく落とすが、人ではそういったことはない。CNN を未知のノイズに対してもロバストにする方法を研究し、人との性能差を埋めるとともに、質感認識の本質を探求した。

画像と言語のよりよい統合

質感認識は最終的に言語で表現されるため、画像を言葉で表現する方法が本質である。前年度において、画像に対する質問に答えるタスクを高精度で解く方法を開発した。当該年度はこれを発展させ、一層の性能向上を目指した。具体的には、複数の異なるタスクを一つのネットワークで学習する、マルチタスク学習を画像と言語の融合問題に適用する方法を検討する。タスクには、入力画像に関する質問に答える visual question answering (VQA) に加えて、画像を説明する説明文をあらかじめ用意された集合から検索する image caption retrieval (ICR) と、説明文内の語句と画像内の物体やその他の概念を対応づける visual grounding (VR) を検討した。

4. 研究成果

(1) 画像ペアのランキングに基づく質感認識手法並びに認識性能評価手法

画像 1 枚からそこに写る物体の（あるいは画像全体から受ける）質感に関わる形容詞の強弱を出力するシステムを開発した。人の知覚の内容を知るために、2 枚の画像の順序付け (ranking) を人に行ってもらい。具体的には 2 枚の画像を被験者（クラウドソーシングのワーカー）に提示し、13 形容詞それぞれ、どちらがより強く感じるか（例えば、より aged か）と問うて、その答えを記録する。多数の画像ペアを対象に、形容詞それぞれについて、順序の情報を得る。この結果を可能な限り忠実に再現するよう、CNN の学習を行う。個人差を学習に反映させるため、投票数から推定される選択確率の分布を予測する問題として定式化する。そのため、Siamese 型の構造を持つ CNN を使って、画像ペアを 2 つの入力層に入力し、出力層では 13 の形容詞それぞれ、ソフトマックス関数にて順序結果を得るようにした。図 1 に我々の CNN の認識を可視化したものを示す。各形容詞の強弱の値を計算した場合に、画像のどこからどれだけの貢献を得たかを表しており、赤が正、青が負の貢献を与えていることを示す。

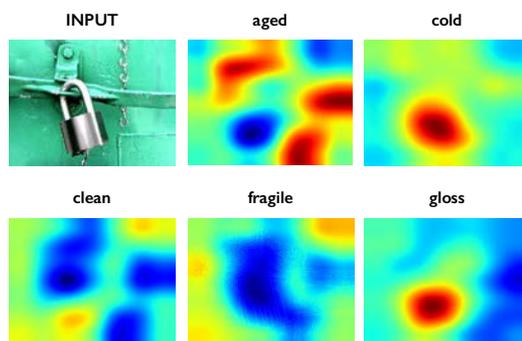


図 1. 質感形容詞の CNN による認識の可視化

(2) ウェブデータを用いた質感概念のマイニング

上のシステムでは、事前に選んだ形容詞につき、クラウドソーシングによって学習データを作り出していた。理屈の上では、この目的で辞書にある質感形容詞全てに対し、同様に学習データを作ることでも可能だが、視覚的に判別不可能な形容詞が多数あることもあって、現実的でない。そこで、ウェブ上にあるデータを用いて、質感形容詞あるいは類語を自動的に発見する方法を検討した。データとして、主として手作りの商品を販売する Etsy (etsy.com) の商品データを用いた。商品画像とその説明文のペアを分析することで、人手を使ったアノテーションを特に行うことなく、視覚的に認識可能な概念を、自動的に獲得するようにした。発見された語と、その認識を可視化したものを図 2 に示す。多様な概念を自動的に取り出せていることが分かる。



図 2. 視覚的に認識可能な概念を表す語彙の発見

(3) 文脈を反映した質感認識

形容詞一語のみで表せる質感は限られる。中には、文脈によって意味が変化するような質感概念もある。読み取った質感を表現するには、形容詞＋名詞でできた句や、さらに柔軟な言語表現を用いる必要がある。これを可能にするため、画像記述 (image captioning) をこの問題に適用した。上と同じ Etsy のデータを使い、商品の画像からその商品の内容を的確に表現したタイトル文を生成することを目標に、Etsy の商品のタイトル文から、半自動でここでの目的に不適當な

サンプルを取り除き、教師データとする。これを使って CNN-LSTM ハイブリッドを学習し、画像からタイトル文を生成できるようにした。概ね、より自然な商品タイトルを生成できた。

(4) 深層学習を用いた画像理解

1 枚の画像と質問文のペアを入力に、質問文に答える問題、すなわち Visual Question Answering (VQA) のための深層学習を用いた手法を開発した。与えられる質問文は、そこに写るシーンに関するもので、それに的確に答えるには画像の理解が必要になる。

われわれの研究成果は、VQA というタスクの適切なネットワーク構造がどのようなものかを考えることにある。注目したのは、注意機構のデザインである。VQA の研究では、最初に与えられた質問文から、画像のどこに注意を向けるべきかを計算する注意機構が考え出され、次に、これと同時に画像から質問文中で重視すべき単語の位置を計算する注意機構 (Co-Attention) が考え出された。ただしそれら先行研究では、質問文全体から画像の各部分領域を相手に注意 (重み分布) を計算し、また画像全体から質問文の各単語を相手とする注意を計算していた。これに対しわれわれは、画像の各領域から質問文の各単語への注意と、その逆を計算できるようなネットワーク構造を考え出した。言い換えれば、画像の各領域と質問文の各単語のより密で細かいインタラクションが生まれるような構造である。実験を通じて、この構造が高い正答率をもたらすことが確かめられている。図 3 に出力のサンプルを示す。左上から右下に順番に正答例、誤答例、正答例、誤答例となっている。いずれにせよ、人の正答率とはまだギャップがあり、おそらく学習していないデータ、具体的にはシーンや質問のパタンに対して、脆弱であることを考えると性能向上の余地が残されている。

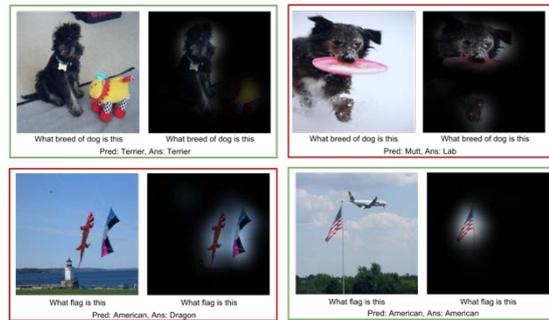


図 3. VQA の結果の例。各欄内、左が入力画像、右が生成された注意のマップ。

(5) 画像の画質向上

画像は、撮影時の条件や伝送時の圧縮等の原因で品質が劣化することがある。画像の画質は、質感認識に様々な角度から影響を与える。そのように劣化した画像から元の画像を推定する画像復元の問題は、品質の高い映像を得る目的の他、様々な応用において画像認識の性能を確保する上でも重要である。そこで、多様な要因により品質が劣化した画像を対象とする画像復元の方法を研究し、画像復元のための畳み込みニューラルネットワーク (以下 CNN) の基本構造「二重残差結合 (dual residual connection)」を開発した。ノイズ、モーションブラー、降雨、雨滴、霧・霞という 5 つの画質劣化要因それぞれに適用し、9 つのベンチマークテストにて世界最高精度を達成した。

(6) ファッション認識への応用

質感の画像認識技術を応用し、ユーザが所有するアイテム、例えば衣服、アクセサリ、靴をどのように組み合わせるべきかを推薦 (レコメンデーション) するシステムを提案した。その中核に、服装を構成するアイテムの画像を入力すると、その組み合わせの良し悪しを評価する深層ニューラルネットワークがある。ファッションの情報共有サイトにおいて、同サイトのユーザが登録した 40 万件以上の組み合わせ事例を用いて同ニューラルネットワークを訓練している。実験を通じて、以上の方法によって高精度な推薦が可能であることを示した (図 4 左)。また、上で構築した推薦システムの中核にある深層ニューラルネットワークを対象に、その判断の理由を、人間が理解できる形で説明する方法を提案した。同ニューラルネットワークが出力する服装の良さを表すスコアに、各アイテムとその色や形などの特徴が、どの程度の大きさの影響を与えているかを計算することで、例えば、悪いと判断された組み合わせのうち、どのアイテムのどの特徴が原因となってその悪い評価につながったかを提示できる (図 4 右)。



図 4. 左：着合わせの良し悪しを答える推薦システム。右：良し悪しの判断を説明する AI による実行例。

(7) 霊長類の皮質脳波の符号化・復号化

質感認識のメカニズムを考える上で動物、特に霊長類の脳は良い手本である。われわれは、霊長類の視覚に関わる脳内情報処理の仕組みを解明することを目的に、マカクザルの頭蓋内に留置した電極で測定される皮質脳波(以下 ECoG)を使い、脳波による視覚の符号化と復号化の方法を研究し、複数の成果を得た。

まず、マカクザルの視覚皮質(下側頭葉皮質)から得た ECoG の信号を使って、その各周波数帯域においてどのような情報が表現されているのかを調べている。静止自然画像1枚をサルに提示したときに測定される脳波と、同一の画像を深層畳み込みニューラルネットワーク(以下 CNN)に入力したときに得られる内部の層の出力との関係を、回帰分析によって調べた(図5)。その結果、脳波のシータ帯が CNN の上位層の特徴と、ガンマ帯が CNN の下位層の特徴と、それぞれ密接に対応付けられることが分かった。この知見は、視覚皮質における脳波のシータ帯とガンマ帯が、それぞれ高次と低次の画像特徴を表現していることが示唆される。

また、マカクザルの視覚皮質の ECoG 信号を元に、サルが見ている画像を再構成する問題を研究した。信号の適切な前処理を検討し、再構成に適した構造の深層ニューラルネットワークをデザインし、敵対的学習の枠組みによってその学習を行うことによって、画像の内容を部分的に再構成することに成功している(図6)。皮質脳波(ECoG)を扱ったものとしては、関連分野で最初の成功事例である。

さらに CNN での質感内部表現を解明し、質感の表現空間構築へのヒントを得るため、質感認知標準刺激に対する皮質脳波応答を下側頭葉皮質から記録した。その結果、下側頭葉皮質で各素材次元が空間マップとして表現されることが明らかになった。また決定木回帰法により、質感 CG の画像基本統計量による皮質脳波事象関連スペクトラム摂動の回帰分析を行なった結果、質感 CG の特定の画像基本統計量が、特定の時間・周波数領域で動的に表現されることが明らかになった(図7)。

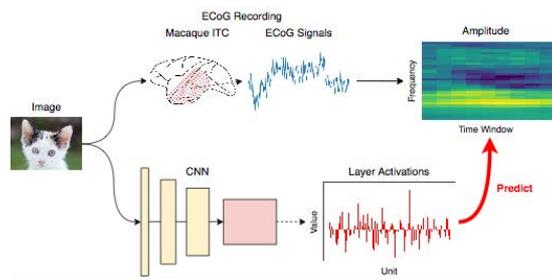


図5. 動物が画像を見たときの皮質脳波を、同じ画像を CNN に入力したときの層出力から予測。



図6. 画像(最上段)を見たときのサルの皮質脳波から同画像を再構成した例。

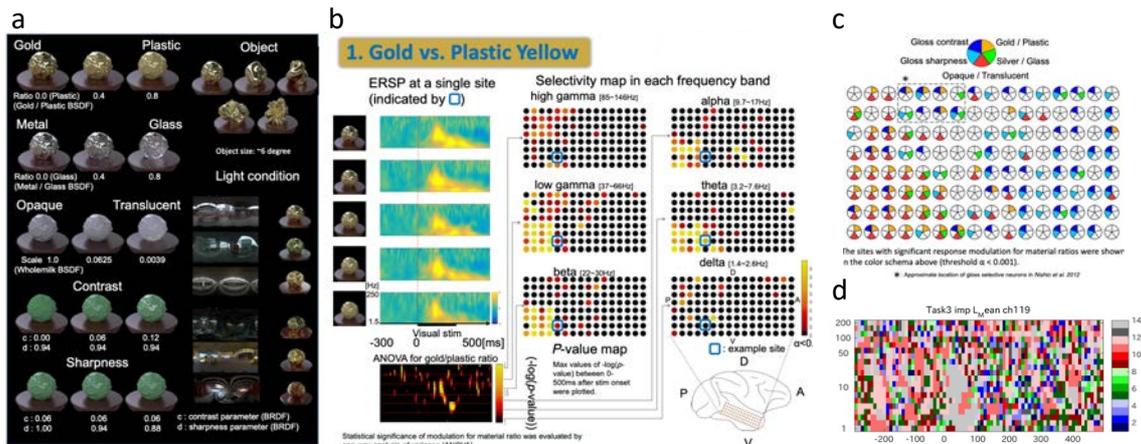


図7. a. 視覚刺激: 質感認知標準課題画像データベース(sawayama et al., 2019)を用いた。物理ベースレンダリングコンピュータグラフィックス(CG)によって、3次元物体を生成し、物体素材を金とプラスチック、金属とガラス、不透明物と半透明物の間で段階的に変化させた。また表面光沢のコントラスト、鋭さを段階的に変化させた。b. サル下側頭葉皮質から記録された質感CG刺激に対する皮質脳波応答: 左列、金とプラスチックの割合を段階的に変化させた時の事象関連スペクトラム摂動。右2列、有意な事象関連スペクトラム摂動がみられた電極位置を周波数帯域毎に示した。b最下段右: 大脳右半球での電極留置位置。各周波数帯ごとに示した。c. 各素材変化に対して有意な変調が見られた電極位置。d. 質感CGの画像基本統計量(1. L*平均2. L*標準偏差3. L*歪度4. L*尖度5. a*平均6. a*標準偏差7. a*歪度8. a*尖度9. b*平均10. b*標準偏差11. b*歪度12. b*尖度13. 色相平均14. 彩度平均)による皮質脳波事象関連スペクトラム摂動の回帰。

5. 主な発表論文等

〔雑誌論文〕 計24件（うち査読付論文 23件 / うち国際共著 0件 / うちオープンアクセス 17件）

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 Electrocardiography Signals	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE International Conference on Bioinformatics and Biomedicine	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 -
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 Hiroto Date, Keisuke Kawasaki, Isao Hasegawa, Takayuki Okatani	4. 巻 -
2. 論文標題 Deep Learning for Channel-Agnostic Brain Decoding across Multiple Subjects	5. 発行年 2020年
3. 雑誌名 Proceedings of International Winter Conference on Brain-Computer Interface	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） 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 -
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 Nguyen Duy-Kien, Okatani Takayuki	4. 巻 -
2. 論文標題 Improved Fusion of Visual and Language Representations by Dense Symmetric Co-attention for Visual Question Answering	5. 発行年 2018年
3. 雑誌名 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition	6. 最初と最後の頁 6087-6096
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR.2018.00637	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 Sun Zhun, Ozay Mete, Zhang Yan, Liu Xing, Okatani Takayuki	4. 巻 -
2. 論文標題 Feature Quantization for Defending Against Distortion of Images	5. 発行年 2018年
3. 雑誌名 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition	6. 最初と最後の頁 7957-7966
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPR.2018.00830	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 Masanori Suganuma, Mete Ozay, Takayuki Okatani	4. 巻 80
2. 論文標題 Exploiting the Potential of Standard Convolutional Autoencoders for Image Restoration by Evolutionary Search	5. 発行年 2018年
3. 雑誌名 Proceedings of Machine Learning Research	6. 最初と最後の頁 47714780
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 Tangseng Pongsate, Yamaguchi Kota, Okatani Takayuki	4. 巻 -
2. 論文標題 Recommending Outfits from Personal Closet	5. 発行年 2018年
3. 雑誌名 Proceedings of IEEE Winter Conference on Applications of Computer Vision	6. 最初と最後の頁 269-277
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV.2018.00036	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 Hu Junjie, Ozay Mete, Zhang Yan, Okatani Takayuki	4. 巻 -
2. 論文標題 Revisiting Single Image Depth Estimation: Toward Higher Resolution Maps With Accurate Object Boundaries	5. 発行年 2019年
3. 雑誌名 Proceedings of IEEE Winter Conference on Applications of Computer Vision	6. 最初と最後の頁 1043-1051
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/WACV.2019.00116	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 Okatani Takayuki, Liu Xing, Sukanuma Masanori	4. 巻 -
2. 論文標題 Improving Generalization Ability of Deep Neural Networks for Visual Recognition Tasks	5. 発行年 2019年
3. 雑誌名 Proceedings of International Workshop on Computational Color Imaging	6. 最初と最後の頁 3~13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-13940-7_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 Xing Liu, Masanori Sukanuma, Zhun Sun, Takayuki Okatani	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 Conference on Computer Vision and Pattern Recognition	6. 最初と最後の頁 印刷中
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 Masanori Sukanuma, Xing Liu, Takayuki Okatani	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 Conference on Computer Vision and Pattern Recognition	6. 最初と最後の頁 印刷中
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 Conference on Computer Vision and Pattern Recognition	6. 最初と最後の頁 印刷中
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 Yan Zhang, Mete Ozay, Shuhao Li, and Takayuki Okatani	4. 巻 -
2. 論文標題 Truncating Wide Networks Using Binary Tree Architectures	5. 発行年 2017年
3. 雑誌名 Proceedings of International Conference on Computer Vision	6. 最初と最後の頁 2097-2105
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICCV.2017.231	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 Shuhao Li, Kota Yamaguchi, and Takayuki Okatani	4. 巻 -
2. 論文標題 Attention to Describe Products with Attributes	5. 発行年 2017年
3. 雑誌名 Proceedings of Machine Vision Applications	6. 最初と最後の頁 215-218
掲載論文のDOI (デジタルオブジェクト識別子) 10.23919/MVA.2017.7986839	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 Sun Z, Ozay M, Okatani T	4. 巻 -
2. 論文標題 Design of Kernels in Convolutional Neural Networks for Image Classification	5. 発行年 2016年
3. 雑誌名 Proc. European Conference on Computer Vision 2016	6. 最初と最後の頁 51-66
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-319-46478-7_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 Vittayakorn S, Umeda T, Murasaki K, Sudo K, Okatani T, Yamaguchi K	4. 巻 -
2. 論文標題 Automatic Attribute Discovery with Neural Activations	5. 発行年 2016年
3. 雑誌名 Proc. European Conference on Computer Vision 2016	6. 最初と最後の頁 252-268
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-319-46493-0_16	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 Yasushi Akashi, Takayuki Okatani	4. 巻 146
2. 論文標題 Separation of reflection components by sparse non-negative matrix factorization	5. 発行年 2016年
3. 雑誌名 Computer Vision and Image Understanding	6. 最初と最後の頁 77-85
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.cviu.2015.09.001	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 Takuya Yashima, Naoaki Okazaki, Kentaro Inui, Kota Yamaguchi, Takayuki Okatani	4. 巻 -
2. 論文標題 Learning to Describe E-Commerce Images from Noisy Online Data	5. 発行年 2016年
3. 雑誌名 Proc. Asian Conference on Computer Vision 2016	6. 最初と最後の頁 85-100
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-319-54193-8_6	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 Y, Ozay M, Liu X, Okatani T	4. 巻 -
2. 論文標題 Integrating Deep Features for Material Recognition	5. 発行年 2016年
3. 雑誌名 Proc. International Conference on Pattern Recognition	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) 10.1109/ICPR.2016.7900209	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 Shuohao Li, Kota Yamaguchi, and Takayuki Okatani	4. 巻 -
2. 論文標題 Attention to Describe Products with Attributes	5. 発行年 2017年
3. 雑誌名 Proc. International Confernece on Machine Vision Applications	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) 10.23919/MVA.2017.7986839	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 Muraoka, M., Maharjan, S., Saito, M., Yamaguchi, K., Okazaki, N., Okatani, T., & Inui, K.	4. 巻 -
2. 論文標題 Recognizing Open-Vocabulary Relations between Objects in Images	5. 発行年 2016年
3. 雑誌名 Proc. Pacific Asia Conference on Language, Information and Computation	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 Tatsuro Orikasa, Takayuki Okatani	4. 巻 99-D(3)
2. 論文標題 A Gaze-reactive Display for Simulating Depth-of-field of Eyes When Viewing Scenes with Multiple Depths	5. 発行年 2016年
3. 雑誌名 IEICE TRANSACTIONS on Information and Systems	6. 最初と最後の頁 739-746
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2015EDP7110	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 岡谷貴之	4. 巻 31(2)
2. 論文標題 画像認識のための深層学習の研究動向：畳込みニューラルネットワークとその利用法の発展	5. 発行年 2016年
3. 雑誌名 人工知能	6. 最初と最後の頁 169-179
掲載論文のDOI (デジタルオブジェクト識別子) 10.11517/jjsai.31.2_169	査読の有無 無
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52件（うち招待講演 39件 / うち国際学会 7件）

1. 発表者名 三木 治憲、安齋 健太、澤山 正貴、松尾 健、鈴木 隆文、岡谷 貴之、長谷川 功、川寄 圭祐
2. 発表標題 サル下側頭葉皮質における質感属性の時空間周波数表現
3. 学会等名 Neuro 2019
4. 発表年 2019年

1. 発表者名 K. KAWASAKI, H. MIKI, K. ANZAI, M. SAWAYAMA, T. MATSUO, T. SUZUKI, I. HASEGAWA, T. OKATANI
2. 発表標題 Spatial and time-frequency representations of glossy material properties in the monkey inferior temporal cortex
3. 学会等名 Neuroscience 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 三木治憲、安齋健太、澤山正貴、松尾 健、鈴木隆文、岡谷貴之、飯島淳彦、長谷川功、川寄圭祐
2. 発表標題 物体素材に対するサル下側頭葉皮質における皮質脳波応答
3. 学会等名 第66回 中部日本生理学会
4. 発表年 2019年

1. 発表者名 三木治憲、澤山正貴、岡谷貴之、川寄圭祐
2. 発表標題 物体素材に対するサル下側頭葉皮質における皮質脳波応答
3. 学会等名 質感のつどい 第5回公開フォーラム
4. 発表年 2019年

1. 発表者名 Xing LIU, Masataka SAWAYAMA, Ryusuke HAYASHI, Mete OZAY, Takayuki Okatani, Shin'ya NISHIDA
2. 発表標題 Perturbation Tolerance of Deep Neural Networks and Humans in Material Recognition
3. 学会等名 Vision Sciences Society Annual Meeting (国際学会)
4. 発表年 2018年

1. 発表者名 Takayuki Okatani
2. 発表標題 Toward Computer Vision Capable of Understanding Objects Beyond Their Categories Like Humans Do
3. 学会等名 Hekksagon Data Science Workshop (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 画像認識・処理のための深層学習の研究の最新動向と今後, 第21回 けいはんな「エジソンの会」
3. 学会等名 第21回 けいはんな「エジソンの会」(招待講演)
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 AIとディープラーニングの現在と今後
3. 学会等名 みやぎ組込み産業振興協議会第12回通常総会(招待講演)
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 ディープラーニング：基礎から最近のトレンドまで
3. 学会等名 第149回微小光学研究会（招待講演）
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習のこれまで，現在と今後---画像応用を中心とした議論
3. 学会等名 精密工学会第397講習会（招待講演）
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 画像分野でのAI特に深層学習の現在と今後
3. 学会等名 日本光学会年次学術講演会・AI Opticsキックオフシンポジウム（招待講演）
4. 発表年 2018年

1. 発表者名 Takayuki Okatani
2. 発表標題 Toward Computer Vision for Understanding Objects Beyond Their Categories
3. 学会等名 Bilateral Workshop between Tohoku University and National Tsing Hua University（国際学会）
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 AIの中核技術としての深層学習の限界と可能性
3. 学会等名 AIビジネス推進コンソーシアムAIBPC Meetup winter 2019 (招待講演)
4. 発表年 2019年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習のフロンティア：画像応用を中心に
3. 学会等名 JSPS産学協力研究委員会「生体ひかりイメージング技術と応用第185委員会」2月期185委員会研究会 (招待講演)
4. 発表年 2019年

1. 発表者名 岡谷貴之
2. 発表標題 AI(主に深層学習)の現状
3. 学会等名 土木学会・構造工学でのAI活用に関する研究小委員会 (招待講演)
4. 発表年 2019年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の基礎～最先端
3. 学会等名 AIOptics第1回研究会 (招待講演)
4. 発表年 2019年

1. 発表者名 Takayuki Okatani
2. 発表標題 Improving Generalization Ability of Deep Neural Networks for Visual Recognition Tasks
3. 学会等名 Computational Color Imaging Workshop 2019 (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 岡谷貴之
2. 発表標題 ディープラーニングによる画像認識・処理の発展と今後
3. 学会等名 日本音響学会2018年春季研究発表会 (招待講演)
4. 発表年 2018年

1. 発表者名 岡谷貴之
2. 発表標題 質感の画像認識について
3. 学会等名 ビジョン技術の実利用ワークショップ (ViEW) 2017 (招待講演)
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の画像認識・処理への応用の現在と今後
3. 学会等名 仙台高専産学連携振興会 産学官交流技術フォーラム (招待講演)
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 画像認識を中心とする深層学習の研究状況とわれわれの取り組みについて
3. 学会等名 画像学会・フリートーカー ImagingToday (招待講演)
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 画像認識を中心とした 深層学習の学術研究の最前線と産業応用について
3. 学会等名 計測・制御・システム工学部会シンポジウム (制御技術部会共催) 「センシングデータにおけるデータサイエンスの最前線」 (招待講演)
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 ビジョンへの深層学習の応用：向上する性能と拡大する応用範囲
3. 学会等名 日本ロボット学会・ロボット工学セミナー (招待講演)
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 人の視覚と機械の視覚：深層学習から多視点幾何まで
3. 学会等名 計測自動制御学会東北支部研究集会 (招待講演)
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 画像認識を中心とする深層学習の研究状況とわれわれの取り組みについて
3. 学会等名 精密工学会・画像応用技術専門委員会総会（招待講演）
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 人の視覚と機械の視覚・深層学習から多視点幾何まで
3. 学会等名 東京大学情報理工学系研究科コンピュータ科学専攻講演会（招待講演）
4. 発表年 2017年

1. 発表者名 岡谷貴之
2. 発表標題 多様体上の最適化に基づく深層畳込みニューラルネットワークの学習と理解
3. 学会等名 横浜国立大学研究集会「新たな数理科学の可能性に向けて- IoT・人工知能・量子ウォークとその周辺 -」（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 主に画像認識に関するディープラーニングの研究動向と、われわれの取り組みについて
3. 学会等名 東北大学 通研共プロジェクト研究会（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 画像認識・空間計測技術の最新動向とバイオリギングへの応用の可能性
3. 学会等名 第12回日本バイオリギング研究会シンポジウム（招待講演）
4. 発表年 2016年

1. 発表者名 伊達裕人, 川崎圭佑, Mete Ozay, 本郷拓実, 長谷川功, 岡谷貴之
2. 発表標題 Modeling electrocorticography signals on the macaque inferior temporal cortex in space, time and frequency domains using hierarchical visual features of convolutional neural networks
3. 学会等名 Neuroscience 2016
4. 発表年 2016年

1. 発表者名 Takayuki Okatani
2. 発表標題 Deep Learning and its Application to Computer Vision: Current Status and Future Challenges
3. 学会等名 SICE Annual Conference 2016（招待講演）（国際学会）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 材質・質感の画像認識に関する研究について
3. 学会等名 ニューロコンピュータビジョン研究会（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習 (Deep Learning) の基礎と応用
3. 学会等名 自動車技術会特別講演会 (招待講演)
4. 発表年 2016年

1. 発表者名 Yashima T, Okazaki N, Inui K, Yamaguchi K, Okatani T
2. 発表標題 Learning to Describe Product Images from Noisy Online Data
3. 学会等名 画像の認識・理解シンポジウム2016
4. 発表年 2016年

1. 発表者名 Liu X, Ozay M, Zhang Y, Okatani T
2. 発表標題 Understanding Deep Representations Learned in CNNs for Material Recognition
3. 学会等名 画像の認識・理解シンポジウム2016
4. 発表年 2016年

1. 発表者名 Ozay M, Okatani T
2. 発表標題 Transformation Invariant Learning in Convolutional Neural Networks
3. 学会等名 画像の認識・理解シンポジウム2016
4. 発表年 2016年

1. 発表者名 *伊達裕人, 川崎圭佑, Mete Ozay, 本郷拓実, 長谷川功, 岡谷貴之
2. 発表標題 Correspondence between the internal representation of convolutional neural networks and the time-space-frequency domains of electrocortigraphy in inferior temporal cortex
3. 学会等名 第39回日本神経科学大会
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 畳込みニューラルネットの理解へ向けて： 学習方法の効率化に関する議論
3. 学会等名 パーティクルフィルタ研究会（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習と材質・質感認識への応用
3. 学会等名 視覚情報基礎研究会（招待講演）
4. 発表年 2016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最新動向： 画像応用を中心とした話題
3. 学会等名 日本学術振興会第171委員会58研究会（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 畳込みニューラルネットワークのより良い学習へ向けて：多様体上の最適化手法
3. 学会等名 視覚情報基礎研究会（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習による材質・質感の画像認識
3. 学会等名 人工知能学会全国大会（招待講演）
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の概要と画像応用の最前線
3. 学会等名 第60回システム制御情報学会研究発表講演会（招待講演）
4. 発表年 2016年

1. 発表者名 *Liu X, Ozay M, Zhang Y, Okatani T
2. 発表標題 Learning Deep Representations of Objects and Materials for Material Recognition
3. 学会等名 Vision Science Society 16th Annual Meeting（招待講演）
4. 発表年 2016年

1. 発表者名 伊達裕人, 川崎圭祐, 岡谷貴之
2. 発表標題 素材カテゴリの脳内表現
3. 学会等名 質感のつどい 第1回公開フォーラム
4. 発表年 2015年

1. 発表者名 岡谷貴之
2. 発表標題 画像認識の様々なタスクに対する畳み込みニューラルネットワークの有効性
3. 学会等名 日本神経科学会視覚科学フォーラム第19回研究会 (招待講演)
4. 発表年 2015年

1. 発表者名 岡谷貴之
2. 発表標題 ディープラーニングと大規模計算
3. 学会等名 VLSI夏の学校 (招待講演)
4. 発表年 2015年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習入門
3. 学会等名 情報論的学習理論と機械学習 (IBISML) (招待講演)
4. 発表年 2015年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の理論的理解と画像認識および周辺の問題への展開について
3. 学会等名 STARCセミナー (招待講演)
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習の現状と今後 ~ コンピュータビジョンへの応用を中心に ~
3. 学会等名 電子情報通信学会RCS-IT-SIP研究会 (招待講演)
4. 発表年 2016年

1. 発表者名 岡谷貴之
2. 発表標題 深層学習とその画像認識への応用に関する最近の研究動向
3. 学会等名 Annual Meeting on Advanced Computing System and Infrastructure (招待講演)
4. 発表年 2016年

1. 発表者名 Xing Liu, Ozay Mete, Yang Zhang, Takayuki Okatani
2. 発表標題 Learning Deep Representations of Objects and Materials for Material Recognition
3. 学会等名 Vision Science Society 2016 (国際学会)
4. 発表年 2016年

〔図書〕 計1件

1. 著者名 麻生 英樹, 安田 宗樹, 前田 新一, 岡野原 大輔, 岡谷 貴之, 久保 陽太郎, ボレガラ ダヌシカ	4. 発行年 2016年
2. 出版社 近代科学社	5. 総ページ数 288
3. 書名 深層学習 Deep Learning	

〔産業財産権〕

〔その他〕

質感の画像認識 http://www.vision.is.tohoku.ac.jp/jp/research/ 東北大学大学院情報科学研究科コンピュータビジョン研究室・研究内容 http://www.vision.is.tohoku.ac.jp/

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	川崎 圭祐 (Kawasaki Keisuke) (60511178)	新潟大学・医歯学系・准教授 (13101)	
研究分担者	山口 光太 (Yamaguchi Kota) (10742596)	東北大学・情報科学研究科・助教 (11301)	2016年度一杯で研究分担者から外れた。

	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
研究協力者	劉 星 (Liu Xing) (60870095)	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
---------	---------