



研究代表者 神谷之康 (ATR 脳情報研究所・客員室長/京都大学大学院情報学研究所・教授)

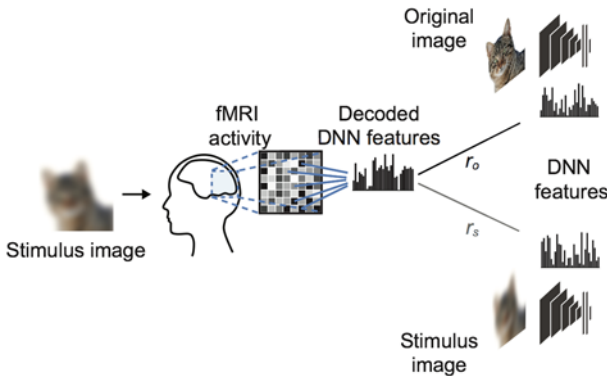


図1：ブラー処理した画像を見ているときの脳活動から深層ニューラルネットワーク特徴量を予測（デコード）すると、元画像を見ているときの特徴に近づいていることがわかった。トップダウン処理により、脳内の特徴表現が「シャープ」になっていることが示唆される (Abdelhack & Kamitani, 2017)。

○研究の背景と目的

従来の質感研究では、少数の物理パラメータに注目して脳活動との相関が調べられてきましたが、われわれが経験する質感は、言語表現とも結びついた複雑な情報構造をもつと考えられます。本課題では、多様な質感情報表現を探索・発見するために、大規模脳・画像・テキストデータのデータマイニングを用いた発見的なアプローチで質感研究を推進しています。

機能的磁気共鳴画像法 (functional magnetic resonance imaging, fMRI) 等により計測されるヒトの脳活動信号を機械学習によるパターン認識で解析することで心の状態を解読する技術は「ブレイン・デコーディング」と呼ばれ、当研究グループが世界に先駆けて開発してきました。一方、深層ニューラルネットワーク (deep neural network, DNN) は、脳の基本素子であるニューロンやシナプスにヒントを得て作られた人工ニューラルネットワークの一種で、近年ではヒトと同レベルの物体認識精度を達成しています。

最近われわれは、画像を見ている時のヒトの脳活動パターンと、同じ画像を入力としたときの DNN の信号パターンの中に相同性を発見し、脳から DNN への信号変換が可能であることを見出しました。DNN の各ユニットの活動は、大規模データを用いてネットワークを訓練する際に自動的に抽出される特徴量を表現していると考えられます。ヒトの脳活動から DNN 信号パターンに変換できるということは、

データ駆動的に得られる膨大な数の特徴の中にヒトの脳活動と密接に結びついた特徴が多数含まれることを示唆しています。この方法により、質感を生み出す多彩な特徴とその脳内表現にアプローチすることができると考えられます。

○これまでに得られた成果

本研究課題の第一段階では、脳から深層ニューラルネットワークへの信号変換技術を完成させるための研究を進めてきました (Horikawa, Kamitani, 2017)。

本研究では、コンピュータ・ビジョンの分野で用いられる「物体概念を視覚特徴のパターンによって表現する」という考え方を採用し、DNN 特徴量のパターンによって物体画像を表現しました。そして、ブレイン・デコーディングを応用して、被験者が物体画像を見ているときの脳活動パターンから DNN 信号を予測 (脳から DNN への変換) するデコーダを構築しました。学習済みのデコーダで予測された特徴パターンを大規模画像データベースとマッチングすることで、画像を見たり想像したりしているときの脳活動から、任意の物体カテゴリーを解読 (検索) する方法を考案しました。

このアプローチの有効性を検証するため、われわれはまず、脳活動から DNN 特徴量を正確に予測できるかどうかを確認しました。脳の視覚野の異なる部位の活動を使ってデコーダの学習を行ったところ、低次視覚野の活動を使って学習したデコーダは、DNN 低次層の特徴量をよく予測でき、高次視覚野の活動を使って学習したデコーダは、DNN 高次層の特徴量をよく予測できることがわかりました。この結果から、脳と DNN の間の相同性 (ホモロジー) を確認することができました。

デコーダを、物体を想像中の脳活動に対して適用すると、DNN 高次層の特徴量を予測した成績のピークのタイミングが、低次層の予測成績のピークのタイミングに先行するという結果が得られました。物体画像を想像しているときに、想像中の物体の視覚特徴が実際に画像を見ているときと共通の脳活動パターンによって表現されていることが明らかになるとともに、脳が高次から低次の順にトップダウン的に階層的特徴表現を活用している様子を可視化することができました。

これらの結果をもとに、提案するアプローチを用いて、ランダムに選択した 1000 個の物体カテゴリーから、知覚・想像している物体カテゴリーを正しく予測 (検索) できるか検証した結果、正解の物体カ

テゴリーが検索ランキングの上位にランクされる結果が得られました。また、予測が外れた場合であっても、正解のカテゴリーと意味的によく似たカテゴリーを予測できていることがわかりました（例：“duck”が正解の場合に他の鳥類のカテゴリーが予測として与えられる）。

本研究により、モデルの学習に用いていない任意の物体に関しても、ヒトが物体を知覚・想像しているときの脳活動から物体のカテゴリー情報を予測することが可能になりました。また、脳の視覚野とDNNの階層的相同性が示されるとともに、物体を想像しているときには、階層的情報表現が高次から低次に向けて段階的に活用されていることが明らかになりました。

本研究では、ブレイン・デコーディング、DNN、大規模画像データベースを組み合わせることで、脳活動パターンから、知覚・想起している任意の物体を解読する方法を開発しました。人工ニューラルネットワークは脳の構造にヒントを得て作られた数理モデルですが、近年では脳のモデルとしてよりも汎用的な機械学習手法として利用されてきました。この成果によって、人工ニューラルネットワークが再び実際の脳と対応づけられることとなり、ブレイン・デコーディングへの応用だけでなく、脳型人工知能の開発にも貢献することが期待されます。

上記の成果をもとに、今年度われわれは、トップダウンの視覚情報処理が脳内の階層的特徴表現に与える影響について調べました（図 1; Mohamed, Kamitani, 2017）。

この研究では、低周波フィルタによりブラー処理した画像（ブラー画像）を刺激として脳活動を計測し、そこから、上記と同じアルゴリズムで DNN 特徴量をデコードしました。デコードされた DNN 特徴量を、ブラー画像自体の DNN 特徴量とブラー処理する前の元画像の DNN 特徴量それぞれと比較しました。

その結果、脳からデコードされた DNN 特徴量は、刺激画像と比較して元画像の DNN 特徴量により近くなることを見出されました。これは、脳がブラー処理で失われた特徴を補完して表現していることを示しています。ぼやけた画像を知覚する際には、トップダウン処理により補完された質感が意識に上っている可能性が示唆されます。

この研究では、ブラー処理によりぼやけた画像を刺激とし用いましたが、一般に視覚入力においては、物理的な遮蔽や光学的制約により、多くの情報が失われています。ベイズ推論の考え方を利用してトップダウン処理を説明する理論がこれまで提案されてきましたが、実際の脳活動データと対応づけられた具体的な情報表現を調べることは困難でした。脳活動パターンから DNN 特徴量をデコードすることにより、トップダウン処理による具体的な情報表現の

計画研究 B01-3

変調を捉えることが初めて可能になりました。

今後はこれらのアプローチを画像生成や言語情報にも拡張し、質感認知に関わる階層的情報表現を明らかにしたいと考えています。

○関連する研究発表

論文

1. [Horikawa, T.](#), [Kamitani, Y.](#), Generic Decoding of Seen and Imagined Objects using Hierarchical Visual Features. *Nature Communications* 8: 15037, 2017.
2. [Abdelhack, M.](#), [Kamitani, Y.](#), Top-down sharpening of hierarchical visual feature representations. *Affiliations*. *bioRxiv* <https://doi.org/10.1101/230078>, 2017.
3. [Satake, E.](#), [Majima, K.](#), [Aoki, SC.](#), [Kamitani, Y.](#), Sparse ordinal logistic regression and its application to brain decoding. *bioRxiv* <https://doi.org/10.1101/238758>, 2017.