

[研究助成 (A)]

人間の感性を理解する脳融合型 AI の開発

Development of brain-integrated AI recognizing human feeling

2191025



研究代表者

情報通信研究機構
脳情報通信融合研究センター

主任研究員

西田 知 史

[研究の目的]

近年の人工知能 (AI) の発展は、客観的な視覚カテゴリ判別などにおいて機械の大幅な能力向上をもたらした。しかし、「美しさ」の評価や好き嫌いの判断のような人間の主観が強く影響する感性情報の推定において、AI の能力はまだ低い。人間と機械の調和を促進するためには、機械による感性の理解が極めて重要である。本研究は、計測脳活動から定量化した脳情報を既存 AI 技術と融合する、革新的な技術を開発することで、人間の感性とその個人差を推定可能な脳融合型 AI の実装を試みた。この研究で得られた成果は、人間の感性を理解してインタラクションを行う機械の実現につながり、人間と機械が協働する未来社会を実現するための重要な基盤技術としての価値を持つといえる。

[研究の内容, 成果]

■脳融合型 AI 技術

本技術では、最新 AI 技術の一つである深層学習を用い、任意の視聴覚情報を入力としたときの深層学習ネットワーク (DNN) の内部情報を、いったん脳情報 (脳活動) に変換してから、視聴覚入力と結びついた多種の感性情報の推定へ利用する。脳情報に変換することで、DNN の情報表現を脳の情報表現に近づけ、人間の感性情報の推定に効果的な形に変えることが

できると考える。

DNN の内部情報から脳情報への変換のために、機能的磁気共鳴画像法 (fMRI) で計測した脳活動を用いて、脳活動予測モデルと自己回帰モデルをあらかじめ構築する。これらのモデルの概要図を図 1 に示す。

脳活動予測モデルは、視聴覚情報を DNN に入力したときの DNN の中間層活性化パターンと、同視聴覚情報を視聴中に fMRI で計測した脳活動の対応関係を統計的に学習することで訓練する。また、自己回帰モデルは計測脳活動の時間的依存関係 (現在の脳活動と先行する脳活動の対応関係) を統計的に学習することで訓練する。いったんモデルの訓練が完了すれば、任意の視聴覚入力に対して、新たな脳計測を行わずに、脳活動予測モデルにより脳活動を計算し、自己回帰モデルによりその予測脳活動を更新することで脳活動の予測を行う。

本研究では、DNN として視覚情報を扱う

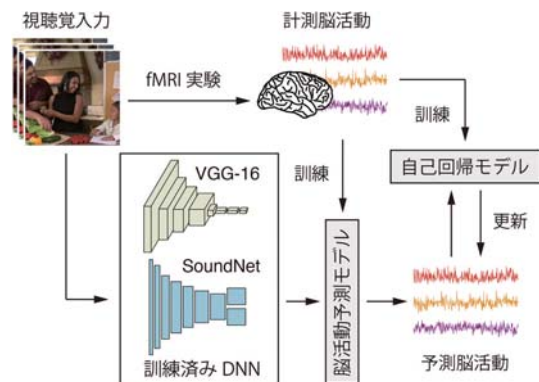


図 1 脳活動予測モデルと自己回帰モデル

VGG-16 [Simonyan et al., 2015, arXiv] および聴覚情報を扱う SoundNet [Aytar et al., 2016, NIPS] を用いた。計測脳活動は 2 時間分の映像を視聴中の被験者 68 名から fMRI を用いて収集した。なお、全被験者から実験前に書面で同意を得た。また、実験プロトコルについては、情報通信研究機構の倫理審査委員会および安全審査委員会から承認を得た。

続いて、これらモデルを用いて DNN の内部情報から変換された予測脳活動を基に、感性情報の推定を行う解読モデルを構築する。解読モデルの概要図を図 2 に示す。

解読モデルは、視聴覚情報を入力したときに脳活動予測モデル・自己回帰モデルが生成する予測脳活動と、視聴覚情報に対応する感性ラベルの対応関係を統計的に学習することで訓練する。いったん訓練が完了すれば、任意の脳活動から感性ラベルを推定する。

本研究では、感性ラベルとして、視聴覚情報に対して人手で付与した「美しい」「大人っぽい」などの 30 種類の印象に対する平均評定結果と、個人が別々に付与した視聴覚情報に対する好き嫌いの評定結果を用いた。

本技術の特長は、脳活動を感性推定の媒介に使いながらも、モデル構築後は追加の脳計測を必要とせず、任意の入力に対して推定を行う、一種の AI システムとして機能する点である。

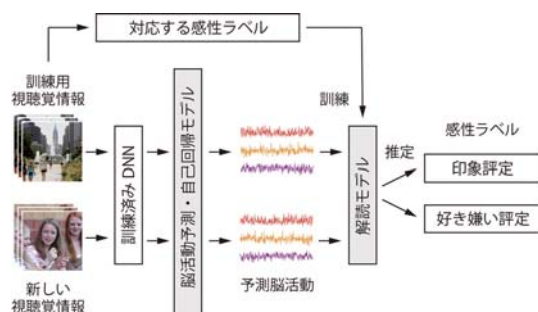


図2 解読モデルの概要図

■感性推定における性能検証

本技術の性能検証として、新しい視聴覚情報に対する感性ラベルの推定課題において、本技

術と既存技術の精度と比較した。既存技術の1つ目は、DNN から脳活動を介さずに感性ラベルを推定する既存 AI 技術である。この比較により、脳活動を媒介することが、推定にどれほどの効果をもたらすかが定量化できる。また、既存技術の2つ目は視聴覚情報を視聴中の計測脳活動から感性ラベルを推定する脳解読技術である。脳解読技術の精度が高いときは、脳情報が有効な課題で、精度が低いときは脳情報が有効でない課題であるといえ、その推定課題における脳情報の有効性を評価できる。表1に比較結果をまとめた。

表1 感性ラベル推定における精度比較

ラベル	印象	好き嫌い
本技術	0.517	0.247
既存 AI 技術	0.501	0.224
脳解読技術	0.349	0.238

表1の数値は、推定ラベルと正解ラベルの間のピアソン相関係数を表す。相関係数が高い方が推定精度が良いことを示す。印象ラベルと好き嫌いラベルの両方において、本技術は既存 AI 技術より有意に高い精度を示した ($P < 0.0001$)。

興味深いのは、既存 AI 技術から本技術への精度向上度合いと、脳解読技術による精度の関係である。印象ラベル推定は、脳解読技術の精度が相対的に低く、脳情報が有効でない推定課題だといえる。そのとき、既存 AI 技術に対する本技術の精度向上は約 3.2% であった。一方で好き嫌い推定は、脳解読技術の精度が相対的に高く、脳情報が有効でない推定課題だといえる。そのとき、既存 AI 技術に対する本技術の精度向上は約 10.3% となり、先の課題より向上度合いが高くなった。つまりこの結果は、脳情報が有効な推定課題ほど、脳情報を媒介する本技術が効果的にはたらくことを示唆している。

また、印象ラベル推定では複数名が行った評定の平均値を推定しているが、好き嫌いラベル推定では個人が行った評定を個人ごとに推定している。つまり、好き嫌いラベル推定は感性の

個人差を含めて推定する難しい問題となっており、その問題において既存 AI 技術より本技術の推定精度が高くなったことは、本技術を感性の個人差の推定へと応用できる可能性を示唆している。

■他種の認知情報の推定

本技術は感性情報の推定だけでなく、他の種類の認知情報を推定する課題においても、既存 AI 技術に比べて高い精度を示した。例えば、広告映像に対する、大規模集団の行動や嗜好を表す指標（集団指標）を推定する課題である。

検証のための集団指標の1つとして用いたのは、Web 広告映像に対する視聴者の反応である。特に、広告映像をスキップせずに最後まで視聴完了した割合（視聴完了率）を大量の Web アクセスから映像ごとに集計し、推定対象とした。さらにもう1つの集団指標は、テレビ広告映像に対して、大規模なモニター調査によって得られた好感度調査の結果である。以上2つの集団指標の推定結果を既存手法と比較した結果を表2に示す。

表2 集団指標推定における精度比較

指標	視聴完了率	好感度調査
本技術	0.444	0.387
既存 AI 技術	0.303	0.365
脳解読技術	0.468	—

これらの推定においても、本技術は既存 AI 技術より有意に高い精度を示した ($P < 0.0001$)。また、ここでも脳解読技術が高い推定精度を示す（脳情報が有効な）視聴完了率の推定において、既存 AI 技術に対する本技術の精度向上度合いは大きく（約 46.5%）、本技術が脳情報を効果的に利用していることを示唆している。

このような集団指標の推定に精度向上をもたらす本技術は、ビジネスや公共サービスへの実装による社会応用の可能性も大いに秘めている。例えば、本研究で検証対象とした Web 広告やテレビ広告などの広告産業は 7 兆円規模の市場であり（2019 年 株式会社電通調べ）、広告を出す前

にその評価を行うサービスを本技術を基に提供できれば、その社会的価値は多大であるといえる。

また本技術の検証においては、感性のみでなく他種の認知の個人差を、本技術を用いて推定できる可能性が示唆された。この検証では、映像の各シーンに対して、多数の人間によるシーン記述を推定対象のラベルとして収集した。一方で、本技術の実装において、個人ごとの計測脳活動から作成したモデルを別々に用いて、同映像に対するラベル推定を行った。そして、人手で行ったシーン記述のばらつきと、本技術を用いたラベル推定の個人モデル間のばらつきをそれぞれ計算し、その関係について調べた。図3は、横軸を推定結果のばらつき、縦軸を正解ラベルのばらつきとして、シーンごとの値を点でプロットした図である。

これら2つのばらつきの間には有意な相関関係が見られた（ピアソン相関係数 0.245, スピアマン相関係数 0.319, $P < 0.0001$ ）。つまり、シーン記述の個人差が大きなシーンほど、本技術の推定結果の個人差も大きくなることを示す。このように、本技術は認知の個人間のばらつきを推定するためにも利用でき、個人差を反映し

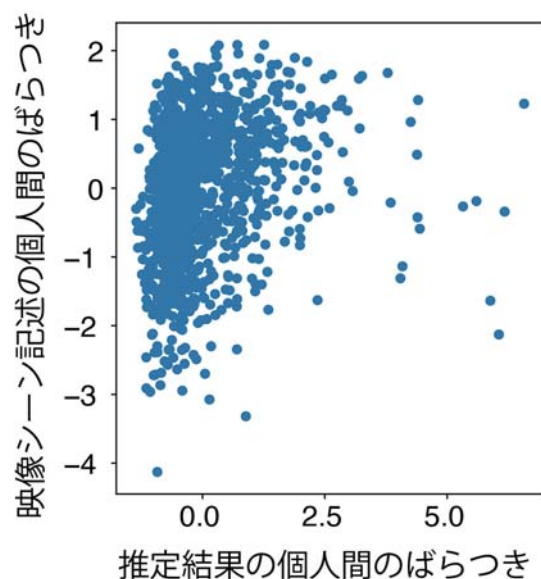


図3 映像シーン記述の個人間のばらつきにおける正解ラベルと推定結果の相関

た認識を行う AI として、大きな利用価値をもつといえる。

■まとめ

本研究では、既存の AI 技術に脳情報を融合して認識性能を強化するための画期的な技術の開発を行った。この技術を、映像に対する印象評定と、個人による映像の好き嫌い評価という2種類の感性ラベルの推定で検証し、性能向上を確認した。また、本技術は広告映像を見たときの集団の行動指標の推定においても、既存 AI 技術に対する優位性を示したと同時に、感性や認知の個人差も推定できる可能性を示した。以上のことから、本技術は人間と機械が協働する未来社会を実現するために重要な、感性を理解する AI を実現するための基盤技術として、未来社会に重要な貢献をもたらす可能性を秘めているといえる。

[今後の研究の方向, 課題]

一つの方向性は、脳活動予測モデルの精度向上である。本技術では、脳情報を取り込んで AI の情報表現を脳の情報表現に近づけることで、人間の感性などの推定精度を向上させることを可能にしている。したがって、任意の入力から脳の情報表現を予測する、脳活動予測モデルの精度を上げ、脳の情報表現の精緻化を行えば、より人間らしい認識を行う AI を実現することが可能だと考える。この脳活動予測モデルの実装には、神経科学分野における符号化モデリング [Naselaris et al., 2011, NeuroImage] と呼ばれる方法論を踏襲しており、現在様々な研究で符号化モデルの開発と改良が試みられているため、そのような知見を取り込むことで、脳活動予測モデルの性能を向上させられると考えている。

別の方向性は、入力モダリティの多様化である。現在の本技術の実装では映像を主な入力として想定し、静止画や音声なども含む視覚・聴

覚情報の一部にも対応させている。これを嗅覚や味覚、触覚などへ拡張し、多種のモダリティで脳情報を利用した AI を実現したいと考えている。また感覚入力だけでなく、文章入力を受けて、その文章に対する感性や認知を推定する脳融合型 AI への応用も視野に入れている。この文章入力を扱う脳融合型 AI に関しては、すでに一部成果が得られており、学会での発表を行い、大会優秀賞を受賞した(成果リスト1)。

また、本技術は個人ごとにモデルを作成し、感性などの個人差を推定することを特長とするが、モデル作成のために要する fMRI 実験のコストも、社会実装を考える上では削減が望ましい。そのため、より低コストな脳計測(脳波計など)への適用や、被験者間でのモデル汎化、行動データからのモデル推定など、様々な方向性から技術の社会応用を視野に入れた研究を継続して行きたいと考えている。

[成果の発表, 論文等]

1. 西田知史, 中野裕介, Blanc Antoine, 前田直哉, 角将高, 西本伸志. 脳情報を組み込んだ単語分散表現による文章からの印象・好感度推定. 第34回人工知能学会全国大会, 3Q1-GS-9, オンライン開催, 2020年6月. (大会優秀賞受賞, 査読付国内学会, 口頭発表)
2. Nishida S, Nakano Y, Blanc A, Maeda N, Kado M, Nishimoto S. Brain-mediated Transfer Learning of Convolutional Neural Networks. Thirty-Fourth AAAI Conference on Artificial Intelligence, 1540, New York, Feb 2020. (査読付国際学会, 採択率 20.6%)
3. 西田知史. DX時代の最先端技術 消費者感性を予測する脳解読基盤「NeuroAI[®]」: 脳情報とAIの融合. NTT DATA Innovation Conference 2020, 東京, 2020年1月. (国内会議, 招待講演)
4. Nishida S, Nakano Y, Blanc A, Maeda N, Kado M, Nishimoto S. Deep transfer learning mediated by human brain information. 脳と心のメカニズム第20回冬のワークショップ, ルスツ, 2020年1月. (Excellent poster award 受賞, 国内学会, ポスター発表)
5. Nishida S. Towards the decoding of human natural perception for real-world applications. Applying Neuroscience to Business, Yokohama, Sept 2019.

(国際学会, 招待講演)
6. 西田知史, 西本伸志, 脳内情報表現の融合による
深層学習ネットワークの認識能力向上, 第33回人

工知能学会全国大会, 4C2-J-1-02, 新潟, 2019年
6月. (査読付国内学会, 口頭発表)