

令和 2 年度

修士学位論文

脳活動の時空間パターンに基づく深層学習デコーディングによって明らかにされた視覚情報処理:
fMRI と EEG の同時計測による研究

Visual Information Processing Revealed by Deep Neural Decoding

Based on Spatiotemporal Patterns of Brain Activity:

A Concurrent fMRI and EEG Study

1235072 古谷 玲

指導教員 中原 潔

2021 年 2 月 5 日

高知工科大学大学院 工学研究科 基礎工学専攻

情報システム工学コース

要旨

脳活動の時空間パターンに基づく深層学習デコーディングによって
明らかにされた視覚情報処理: fMRI と EEG の同時計測による研究

古谷 玲

視覚情報の脳内処理は、視覚カテゴリごとに処理を行う脳領域が異なることが知られている。しかし、カテゴリ内の個々の視覚情報を表彰する脳活動の時空間パターンについては不明な部分が多い。カテゴリ内の視覚情報処理について明らかにするために、以下の仮説を立てて検証を行った：(1)カテゴリの異なる視覚刺激は別々の脳領域にて処理されるという点で、fMRI が捉える脳活動が識別に貢献する、(2)カテゴリ内の個別の視覚刺激は同一領域内において時間的に異なる反応を示すという点から脳波が捉える脳活動が貢献する。

本研究では空間分解能に優れた機能的磁気共鳴画像法(fMRI)と時間分解能に優れた脳波の同時脳活動データを深層学習により学習させ、trial based で被験者が見た画像のデコードを行うことでこれらの仮説の検証を行った。解析の結果、カテゴリごとの視覚情報処理は、fMRI 単変量解析の結果が先行研究と一致しており、深層ニューラルネットワーク

(DNN) を用いた識別でも高いスコアが得られたことから仮説 (1) が支持されたといえる。また、カテゴリ内の視覚情報処理については、脳波の事象関連電位 (ERP) の解析により仮説 (2) を部分的に支持する結果が得られた。しかし、DNN によりカテゴリ内識別は chance level (50%) に対して統計的に有意な差はあったものの、実用的な数値には至らなかった。この理由としては、trial based の解析では脳波の S/N 比が十分ではないことが挙げられる。ただし、Grad CAM による可視化の結果、カテゴリ内については脳波が識別に fMRI より強く貢献していることが分かった。

Abstract

Visual Information Processing Revealed by Deep Neural Decoding

Based on Spatiotemporal Patterns of Brain Activity:

A Concurrent fMRI and EEG Study

Rei Furutani

Brain processing of visual information is known to involve different brain regions that process different visual categories. However, it remains unclear how the spatiotemporal pattern of brain activity represents individual visual objects within a category. To clarify the visual information processing within a category, we tested the following hypotheses: (1) brain activity captured by fMRI identifies different brain regions in which different categories of visual stimuli are processed, and thus contributes to between-category-discrimination, and (2) brain activity captured by EEG identifies different temporal patterns in which different visual identities in a category are processed, and thus contributes to within-category-discrimination.

In this study, we tested these hypotheses by conducting simultaneous recording of (a) functional magnetic resonance imaging (fMRI), which has excellent spatial resolution, and (b) EEG, which has excellent temporal resolution, to decode the presented images in a trial-based manner. FMRI univariate analysis for categorical processing identified distinct brain regions (e.g., fusiform face area for face category and lateral occipital cortex for non-face object category) which were reported in previous studies. Deep learning-based decoding of visual objects showed high decoding accuracy in between-category classification in combined deep neural networks (DNN) for fMRI data and EEG data, as well as DNN for only fMRI data, suggesting that the hypothesis (1) is supported. As for within-category visual information processing, the analysis of event-related potentials (ERPs) in EEG partially showed distinct

temporal patterns of neural activities for subcategories (e.g., natural and artificial objects in the object category). The DNN model for combined fMRI/EEG as well as EEG only showed a statistically significant classification accuracies within categories compared to the chance level (50%). However, in comparison with the performance for between-category classification, the performance for within-category classification was quite low. We speculate that the signal-to-noise ratio of the EEG is not sufficient for trial-based analysis in within-category classification. Nevertheless, visualization of brain activity using Grad-CAM showed that EEG contributes to the discrimination of visual information within a category. Taken together, our results partially support hypotheses (1) and (2), and suggest that our brain processes information of visual objects in both spatial and temporal manner.

目次

第 1 章	背景	1
第 2 章	実験方法	3
2.1	参加者	3
2.2	実験デザイン	3
2.3	実験手順	4
2.4	手順と前処理	5
2.4.1	fMRI の手順と前処理	5
2.4.2	脳波の手順と前処理	6
2.4.3	呼吸の記録	7
2.4.4	アイトラッキング記録	7
2.4.5	統計解析	7
2.4.6	刺激画像の解析	8
2.5	解析	11
2.5.1	fMRI 単変量解析	11
2.5.2	脳波の ERP 分析	12
2.5.3	ディープニューラルネットワーク(DNN) 分類器	12
第 3 章		14
3.1	behavior 結果	14
3.2	fMRI 単変量解析	15
3.2.1	顔カテゴリ – 物カテゴリ	15
3.2.2	男性顔サブカテゴリ – 女性顔サブカテゴリ	16
3.2.3	自然物サブカテゴリ – 人工物サブカテゴリ	17

3.3	脳波 ERP	18
3.3.1	顔カテゴリ/物カテゴリ	18
3.3.2	男性顔サブカテゴリ/女性顔サブカテゴリ	19
3.3.3	人工物サブカテゴリ/自然物サブカテゴリ	20
3.4	DNN 学習結果	21
3.4.1	DNN スコア	21
3.4.2	Grad CAM による識別特徴マップの可視化	24
 第4章 考察		32
4.1	結果のまとめと仮説について	32
4.2	過去の研究との比較	32
4.3	アイデンティティごとの識別について	33
4.4	デコーディング成績について	33
4.5	今後の展開	33
 謝辞		35
 参考文献		36

目次

1 脳内視覚表象の時空間モデル	2
2 視覚刺激となるカラー画像	4
3 実験課題シーケンス	5
4 サブカテゴリごとの低次視覚特徴	9
5 ピクセル強度	10
6 平均空間周波数分布	10
7 RGB ヒストグラム	11
8 Deep Neural Network モデル	13
9 課題正答率	14
10 課題における画像提示からの反応速度	15
11 カテゴリ間（顔カテゴリ / 物カテゴリ）fMRI 単変量解析結果	16
12 カテゴリ内（男性顔サブカテゴリ/女性顔サブカテゴリ）fMRI 単変量解析結果	17
13 カテゴリ内（自然物サブカテゴリ/人工物サブカテゴリ）fMRI 単変量解析結果	18
14 カテゴリ間（顔カテゴリ/物カテゴリ）ERP	19
15 カテゴリ内（男性顔サブカテゴリ/女性顔サブカテゴリ）ERP	20
16 カテゴリ内（人工物サブカテゴリ/自然物サブカテゴリ）ERP	21
17 fMRI DNN 学習結果	22
18 脳波 DNN 学習結果	23
19 combined DNN 学習結果	24

20	カテゴリ間識別（顔カテゴリ/物カテゴリ）における fMRI の Grad CAM 結果	. . 26
21	カテゴリ間識別（顔カテゴリ/物カテゴリ）における脳波の Grad CAM 結果	. . 26
22	カテゴリ内識別（男性顔サブカテゴリ/女性顔サブカテゴリ）における fMRI の Grad CAM 結果 27
23	カテゴリ内識別（男性顔サブカテゴリ/女性顔サブカテゴリ）における脳波の Grad CAM 結果 27
24	カテゴリ内識別（自然物サブカテゴリ/人工物サブカテゴリ）における fMRI の Grad CAM 結果 28
25	カテゴリ内識別（自然物サブカテゴリ/人工物サブカテゴリ）における脳波の Grad CAM 結果 28

表目次

1	fMRI 「顔カテゴリ - 物カテゴリ」コントラストによる賦活領域	29
2	fMRI 「物カテゴリ - 顔カテゴリ」コントラストによる賦活領域	30
3	fMRI 「女性顔サブカテゴリ - 男性顔サブカテゴリ」コントラストによる賦活領域	31
4	fMRI 「人工物サブカテゴリ - 自然物サブカテゴリ」コントラストによる賦活領域	31

第 1 章

背景

人間を含めた霊長類の視覚処理に関する研究は多く行われており[1]、それらによりおおまかな脳内の視覚情報の処理のプロセスが明らかになっている。その中でも腹側視覚経路では主に物体の形状、色や質感への選択性があり、特に脳の前方領域ではより複雑な視覚刺激の処理を行っている。例えば、FFA は顔の情報処理を担い[2]、PPA は場所の情報処理を担い[3]、LO などはオブジェクトの情報処理を担っている[4-5]。このように、おおまかなカテゴリごとに処理の領域が異なることが明らかになっているものの、カテゴリ内の個々の視覚情報がどのような脳活動の時空間パターンでコードされているかは不明な点が多く、サルを対象とした MUA(Multi unit activity), LFP(local field potential)といった計測手法を用いた研究[6]はあるが、ヒトを対象とした研究でよく用いられる計測手法である機能的磁気共鳴画像法(fMRI)や脳波(EEG)を用いた研究はほとんどない。

上記のカテゴリ内ごとの視覚情報処理様式についての研究が行われていない要因の 1 つについて、前述の MUA, LFP といった空間的に限定しかつ時間的に速い事象を捉える計測手法が、人を対象とした場合には用いることが難しいことが考えられる。fMRI は空間分解能に優れているものの、時間分解能は高いとは言えない[7]。一方で EEG の計測は高い時間分解能を持つものの、空間分解能は低い[7]。これらのことから、ヒトを対象として視覚情報のカテゴリ間・カテゴリ内情報表現を検証する一つの手法として、fMRI と脳波の同時計測が有用であると考えられる。

本研究では fMRI と EEG の同時計測脳活動データを深層学習により学習させ、被験者が見た画像を脳活動データからデコードし、trial based で識別させた。我々は以下の 2 つの仮説を立てた。(1)カテゴリの異なる視覚刺激は別々の脳領域にて処理されるという点で、高い空間分解能を持つ fMRI が識別に貢献する。(2)カテゴリ内の個別の視覚刺激は同一領域内において、時間的に異なる反応を示すという点から脳波が貢献する。

fMRI と脳波の両方を用いることで、カテゴリ間とカテゴリ内のどちらにも精度を持つ識別が可能になると考えられる（図 1）。

Spatio-temporal model of visual object representation

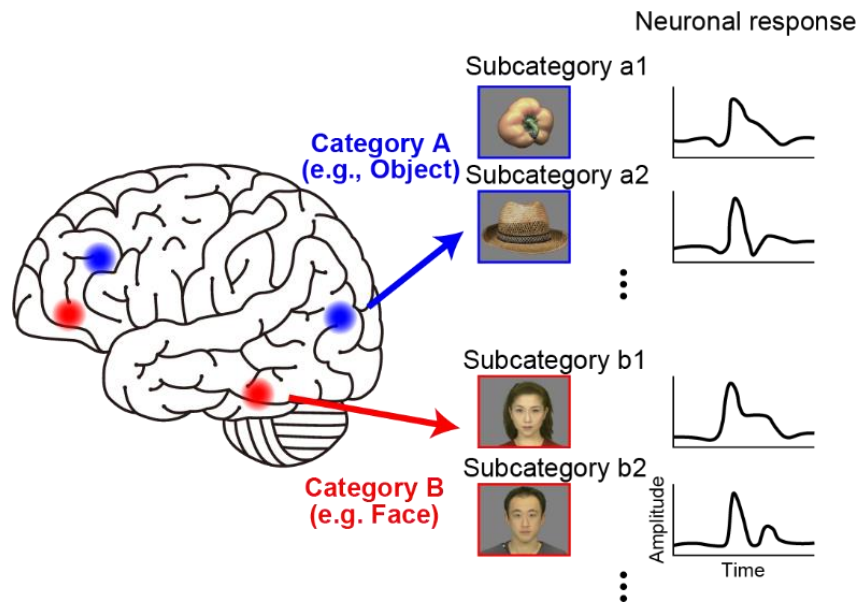


図 1 脳内視覚表象の時空間モデル カテゴリの異なる視覚刺激は別の脳領域で処理され、カテゴリ内の個別刺激は時間的に異なる反応を示す。

第2章

実験方法

2.1 参加者

正常または正常に矯正された視力を持つ健常な右利きの参加者 53 名（男性 34 名、女性 19 名、年齢範囲：18～25 歳）が、書面による同意を得て、本研究に参加した。実験は高知工科大学（香美市）の倫理審査委員会の承認を得て行った。参加者は、各セッションへの参加費として 1 時間につき 1,000 円を受け取った。参加者数は、今回の実験を開始する前に、過去のディープラーニング研究[8]で必要とされた試行回数をもとに決定した。

2.2 実験デザイン

実験刺激としては、顔と物のカテゴリから 40 枚のカラー画像を用いた（図 2）。顔画像 20 枚と物画像 20 枚は、それぞれ ATR 表情データベース DB99（<http://www.atr-p.com/products/face-db.html>）と The Amsterdam Library of Object Images¹ から選択した。顔画像には、男性 2 名（“男性顔サブカテゴリ”）と女性 2 名（“女性顔サブカテゴリ”）の 4 つのアイデンティティが含まれており、5 つの異なる角度の中立的な表情で撮影されている。物画像には、2 つの自然物（パプリカとブロッコリー、“自然物サブカテゴリ”）と 2 つの人工物（麦わら帽子と笛、“人工物サブカテゴリ”）の 4 つのアイデンティティが含まれており、5 つの異なる角度で撮影されている。すべての画像は、図と背景を分離した後、各 RGB チャンネルについて SHINE ToolBox（lumMatch 関数）を改変した MATLAB（Mathworks, Natick, MA）コードによって輝度マッチングを行った。

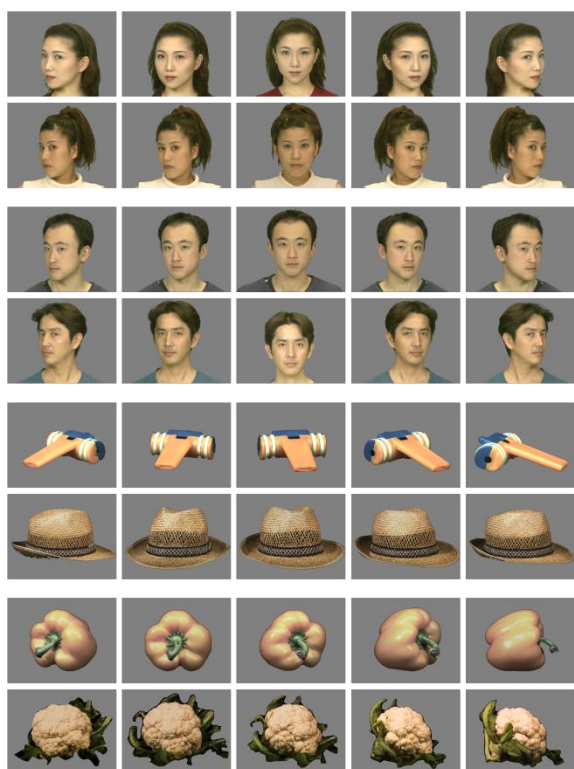


図 2 視覚刺激となるカラー画像 男性 2 名 + 女性 2 名からなる顔画像と、2 つの自然物（パプリカとブロッコリー）と 2 つの人工物（麦わら帽子と笛）からなる物画像。5 つの異なる角度で撮影。

2.3 実験手順

参加者には実験刺激のシーケンスが提示され、ボタンを押すことで対応する視覚サブカテゴリ（男性顔、女性顔、自然物、人工物）を判定してもらった。具体的には、各 Run は 50 試行のシーケンスで構成されており、40 試行では顔または物の画像が提示され、10 試行では、試行間の BOLD response の解離を改善するために空白の画像を提示した。すべての視覚画像は 1 回の試行で 1 回提示され、刺激提示の順序は疑似ランダム化した。視覚フィードバックは行わなかった。Run の最初に、4 つのボタンの位置と 4 つのサブカテゴリの対応を示すタスク指示が提示された（図 3）。その後、fMRI と脳波の同時記録を行った。タスク内容は、まず 30 秒間十字の固視点が提示され、その後 50 回の試行が行われた。各試行では 500 ms の刺激が提示され、その後 9,500 ms の固視点が提示された。そのため、1 回の Run 実行時間は合計 530 秒となる。Run 実行中、参加者は画面の中央を固視するように要求され、目の位置の動きは EyeLink システム（SR リサーチ、カナダ）を使用してモニターした（下記の「アイトラッキング記録」のセクションを参照）。参加者は、fMRI と脳波の同時記録を開始する前に、タスクの練習を行った。各参加者は 3~12 run (9.9 ± 2.8 、

平均±SD、合計 523 run) の fMRI と脳波の同時記録を行った。その結果、顔試行 10,460 回、物試行 10,460 回、空白試行 5,230 回の合計 26,150 試行の神経活動データが得られた。タスクは Psychtoolbox version 3 (<http://psychtoolbox.org/>) を用いてプログラムした。ProPixx DLP LED プロジェクター (VPixx Technologies、カナダ) を使用して、視覚刺激をスキャナの後ろにあるスクリーンに投影した。参加者は、ヘッドコイルに取り付けられた鏡を通して投影された視覚刺激を見た。

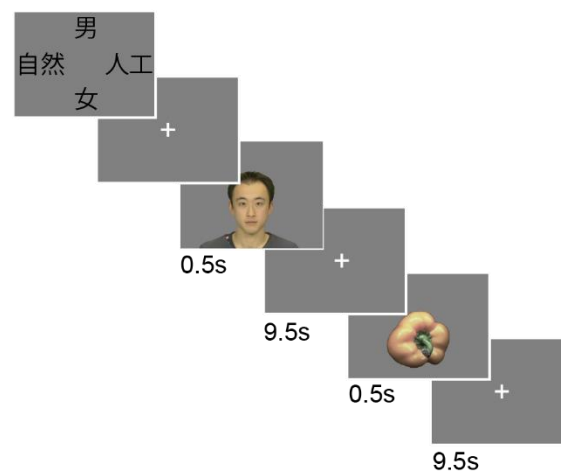


図 3 実験課題シーケンス 初めに 4 つのボタンと 4 つのサブカテゴリの対応を示す画像を提示、その後 30 秒間十字の固視点を提示する。その後 500 ms の刺激提示と 9,500 ms の固視点提示を 1 試行として 50 回の試行が行われた。50 試行のうち 40 試行では顔または物の画像、10 試行では空白の画像を提示した。すべての視覚画像は 1 回の試行で 1 回提示され、刺激提示の順序は疑似ランダム化した。

2.4 手順と前処理

2.4.1 fMRI の手順と前処理

MRI スキャンは、64 ch ヘッドコイルを搭載した 3T MRI スキャナ (Siemens Prisma, Germany) を用いて行った。機能画像は、Echo Planer Imaging sequence [繰り返し時間 (TR): 2.0 s, エコー時間 (TE): 27 ms, フリップ角 (FA): 90 度, 36 スライス, スライス厚さ: 3 mm, 面内分解能: 2×2 mm] を用いて取得し、各 run は、265 volume で構成されている。

最初の 5 volume は、磁場の安定性を考慮するために解析から外した。高解像度の解剖学的画像は、MP-RAGE の T1 強調シーケンス[TR : 2,400 ms、TE = 2.32 ms、FA : 8 度、192 スライス、スライス厚さ : 0.9 mm、面内解像度 0.93×0.93 mm]を使用して取得した。すべての画像の前処理は、SPM12 ソフトウェア (<http://fil.ion.ac.uk/spm/>) を使用して行われた。機能画像は、最初に volume と run をまたいで時間的に realign し、解剖学的画像を機能画像の平均画像に coregister した。次に、解剖学的スキャンに対して推定された正規化パラメータを用いて、MNI テンプレートに空間的に正規化した。画像は 2 mm の等方性ボックスにリサンプリングし、6 mm の FWHM ガウスカーネルで空間的に smoothing を行った。

2.4.2 脳波の手順と前処理

64 個の焼結 Ag/AgCl 電極を用いて、頭部の記録部位 63 箇所の脳波 (EEG) と心電図 (ECG) を記録した (BrainCap-MR、Brain Products、ドイツ)。電極 AFz および FCz は、それぞれ接地および基準電極として用いた。すべての信号は、0.016-250 Hz のバンドパスで記録し、5,000 Hz でデジタルサンプリングした (BrainAmp MR Plus、Brain Products)。電極インピーダンスは、記録前に <10 k Ω に下げた。オフライン前処理は、まず、Brain Vision Analyzer 2 (Brain Products) で行った。まず、平均テンプレート減算法を用いて MR グラディエントアーチファクトを除去し、結果として得られた脳波データを 250Hz にダウンサンプリングした。次に、脳波データを 1-100 Hz の間でバンドパスフィルタリングした。そして、心血管アーチファクトをテンプレートマッチング法により脳波データから除去した。最後に、独立成分分析 (ICA) により瞬き補正を行った。我々は、垂直/水平眼電図 (EOG) のための特定のセンサを使用していないので、EOG チャンネルの垂直方向の活動の代わりに FP1 チャンネルにおける脳波データを使用した。結果として得られたアーチファクト補正後の脳波データは、以下のように EEGLAB でさらに処理された。あるチャンネルで連続的にノイズが発生している場合には、そのチャンネルの脳波データを近隣チャンネルの平均化された脳波データに置き換えることで補間を行った [3,339 チャンネル (53 人 \times 63 スカルプチャンネル) のうち 16 チャンネルを補間した]。脳波データを刺激の開始時刻にタイムロックした 1.5 秒エポック (刺激前 0.5 s \sim 刺激後 1.0 s) に分割し、ベースライン正規化し

た。その後、瞬き、眼球運動、過度のノイズを含む試行を検出するために、データを手動で検査した。その結果、全参加者の 26,150 試行のうち、1,328 試行 (5.1 %) が解析から除外された。分析に使用されたすべての参加者は、アーチファクト除去後に少なくとも 80 % の試行を保持していた。その後、データはすべての頭部チャンネルの平均値にデジタル再リファレンスした。

2.4.3 呼吸の記録

呼吸データは、MRI 対応の完全空気式呼吸トランスデューサと AcqKnowledge ソフトウェア (BIOPAC Systems, Inc.) を用いて 2,000 Hz でサンプリング計測した。MATLAB 用の PhysIO ツールボックスで以下のように前処理した。まず 250 Hz にダウンサンプリングし、0.1~5.0 Hz の間でバンドパスフィルタリングした。次に、データの呼吸位相と時間あたりの呼吸量 (RVT) の両方を計算した。これらの生理学的決定因子とノイズモデリングから、(1)呼吸位相の 4 次フーリエ展開を介して 8 つの RETROICOR (RETROspective Image CORrection) regressor を計算し、(2)RVT と呼吸応答関数との畳み込みにより呼吸応答を計算した。以上の手続きにより、合計 9 個の regressor を得て、fMRI 単変量解析において nuisance regressor として扱った (後述の「fMRI 単変量解析」の項を参照)。

2.4.4 アイトラッキング記録

課題実行中の眼球運動は、MRI 対応の赤外線カメラベースのアイトラッカー EyeLink 1000 (SR リサーチ、オンタリオ州、カナダ) を使用して、1,000 Hz のサンプリングレートで記録された。瞬きは、瞳孔が検出されなかった期間として定義され、分析から除外した。画面の中心 (固視点) からのユークリッド距離を調べるために、頭部参照校正された x および y 座標を使用した。

2.4.5 統計解析

一般的な解析を行った。すべての統計分析は、MATLAB、および Python で実行した。

2.4.6 刺激画像の解析

複数の低次視覚特徴を評価し、以下のように異なるサブカテゴリ間で比較した。Stigliani et al (2015)[9]に基づき、RGB カラー画像を MATLAB の `rgb2gray` 関数でグレースケール画像に変換した後に、(1) コントラスト、(2) 輝度、(3) 同じカテゴリの他の刺激との類似性 (図 4)、(4) 視野率 (図 5)、(5) 空間周波数パワー分布を計算した (図 6) (図 7)。

(1) 画像の Michelson contrast は、最大画素強度と最小画素強度の差と和の比として算出した。 $p < 0.05$ の閾値において、異なるサブカテゴリの画像間でコントラストレベルに有意な差が認められた (サブカテゴリの主効果、 $F(3,36) = 3.13$, $p = 0.0376$)。

(2) 画像の平均画素輝度を算出した。閾値 $p < 0.05$ (サブカテゴリの主効果、 $F(3,36) = 3.65$, $p = 0.0214$) では、異なるサブカテゴリの画像間で輝度レベルに有意な差があることがわかった。

(3) サブカテゴリの画像間の視覚的類似性を推定するために、我々は、各刺激と同じサブカテゴリの他のすべての刺激の正規化されたグレースケール値 (0 と 1 の間であるようにスケールリング) の間の平均ユークリッド距離を測定した。ペアワイズ類似度の値は、0 (各ピクセルでの強度差が反転した画像) から 1 (同一の画像) までの範囲を用いた。カテゴリ間の平均類似度には数値的に有意な差があるが (カテゴリの主効果、 $F(3,176) = 20.65$ 、 $p = 1.64 \times 10^{-11}$)、類似度値の範囲はサブカテゴリ間で類似しているだけでなく、効果の大きさの order (0.0062、正規化されたグレースケール値でのカテゴリ平均の標準偏差(SD)) が、ディスプレイの解像度の order (0.0039、255 刻みの 1 つのグレースケール値の変化(1/255)) と同程度であるため、平均値の差は知覚的には関係ない。

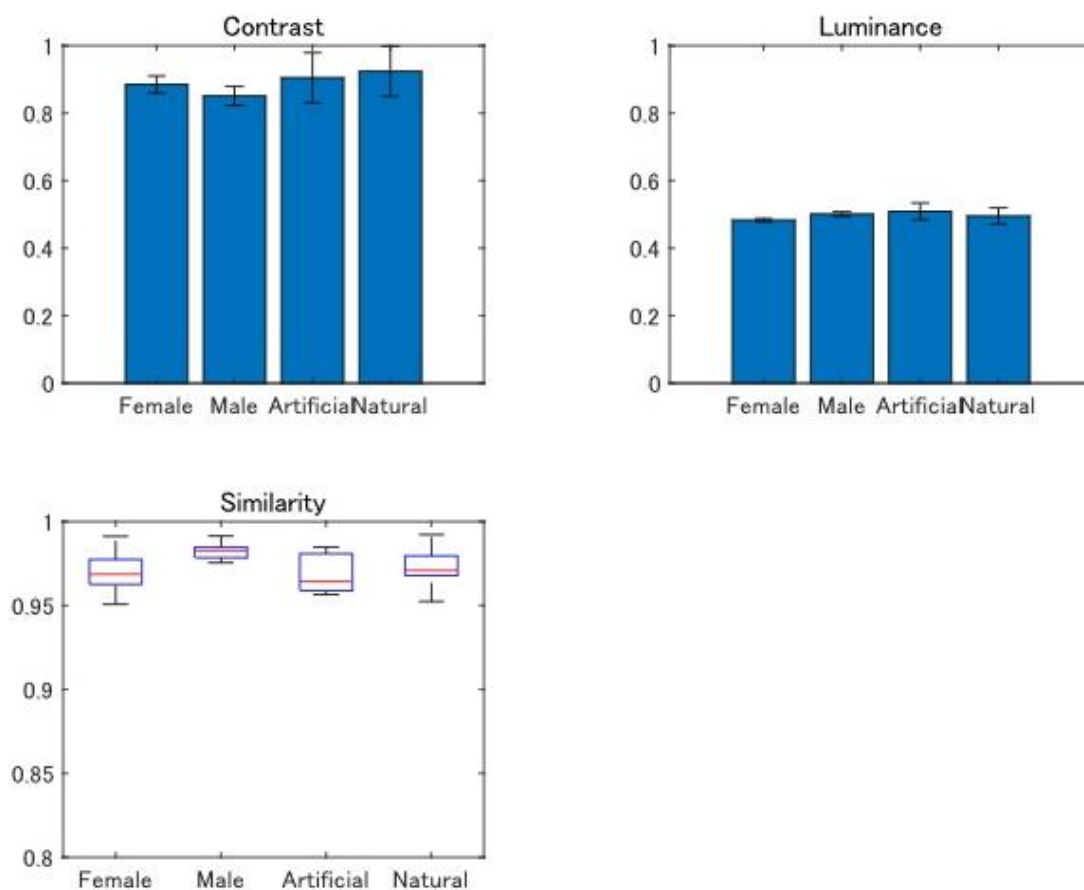


図 4 サブカテゴリごとの低次視覚特徴 (左上)コントラストレベルをサブカテゴリごとに表した。p<0.05 において有意差が認められた。(右上)平均画素輝度を表示。サブカテゴリ間で有意差あり。(左下)視覚類似性を表した。数値的な有意差が認められたものの、order を考慮すると知覚的には差がないと考えられる。

(4) 我々は、サブカテゴリ内の全画像を平均した各ピクセル位置の平均輝度としてピクセル強度を計算した。

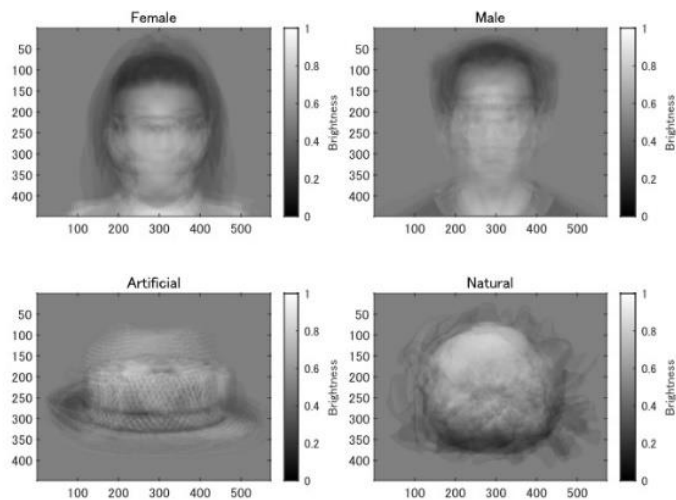


図 5 ピクセル強度 すべての画像のピクセル強度を、サブカテゴリごとに重ねて表した。

(5) 各サブカテゴリの画像全体の平均空間周波数分布を計算した。

これらの指標に加えて、各サブカテゴリの RGB ヒストグラムも算出した。

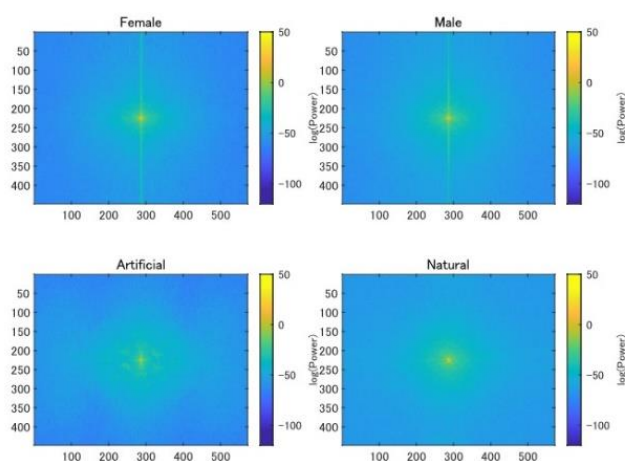


図 6 平均空間周波数分布 すべての画像の周波数分布をサブカテゴリごとに表した。

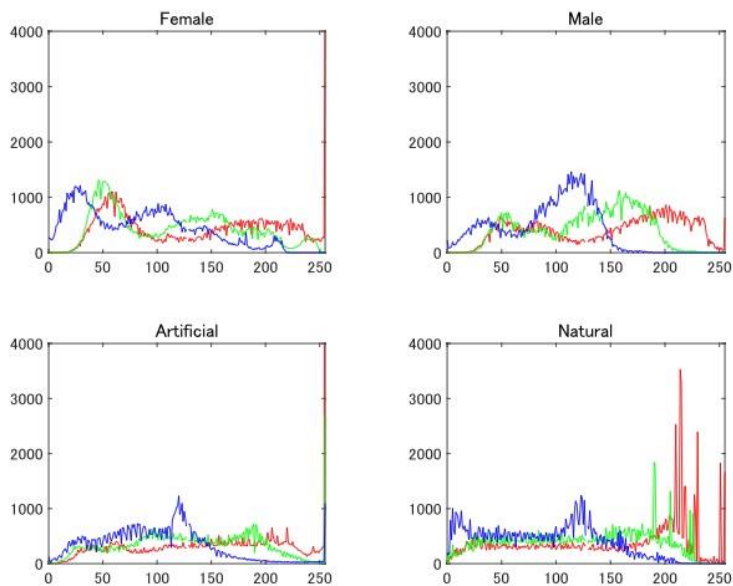


図 7 RGB ヒストグラム 全ての画像の RGB の値をサブカテゴリごとに表した。

2.5 解析

2.5.1 fMRI 単変量解析

参加者ごとの解析では、一般線形モデル(GLM)を用いてタスク event のパラメータ値を推定した。対象とした event は、男性顔、女性顔、自然物、人工物、空白刺激 (9 regressor) の提示である。正しい試行のみを regressor としてコード化した。これらのタスク event は、刺激提示の開始時にタイムロックされ、SPM で実装された標準的な血流動態応答関数(HRF)を用いて展開された。さらに、6 軸の頭部運動パラメータ、白質信号、CSF 信号、呼吸の効果を nuisance 効果として GLM に組み込んだ (17 の nuisance regressor)。次に、脳全体の各ボクセルについてパラメータを推定した。パラメータ推定値のマップは、最初に個々の参加者内で作成した。グループレベルの解析では、各参加者のコントラストマップについて、グループ平均検定を行った。ボクセルクラスタは、まず、 $P < 0.001$, uncorrected のボクセル単位の有意水準を用いて同定された。ボクセルクラスタは、FSL スイート (<http://fmrib.ox.ac.uk/fsl/>) を用いて randomise 関数を用いて ($N = 5,000$)、ファミリーワイズエラー (FWE) 率で補正された $P < 0.05$ のしきい値を持つ全脳全体で有意性をテス

トした。次に、有意なクラスタのピークを同定し、表に列挙した。12 mm 以内に複数のピークが同定された場合は、最も有意なピークを保持した。

2.5.2 脳波の ERP 分析

視覚サブカテゴリごとに、すべての正解試行のデータを平均化することで、単一の事象関連電位 (ERP) を作成した。我々は、あるチャンネルの ERP がカテゴリ内で異なる時間的特徴を示しているかどうか調べた (例えば、男性顔の ERP の時間パターンが女性顔の ERP と異なるかどうかなど)。

2.5.3 ディープニューラルネットワーク(DNN)分類器

視覚的なカテゴリカル表現とサブカテゴリカル表現を含む時空間神経活動を探索するために、ディープニューラルネットワーク分類器(REF) (DNN)を構築した(図 8)。モデルの train/validation は、PyTorch (<https://pytorch.org/>)を用いて実施した。DNN モデルは、(1)fMRI データ用の 3 次元畳み込みニューラルネットワーク(REF) (CNN)と、(2)脳波データ用の 2 次元 CNN(REF)から構成した。fMRI データ用の 3 次元 CNN は、空間特徴を抽出するための 3 つの畳み込み層 (conv3d 関数) と、2 つの全結合層によって構成した。EEG データの 1D CNN は時間的特徴を抽出するための 4 つの畳み込み層(conv1D 関数)と 2 つの完全に全結合層から構成した。これら 2 つのネットワークを結合し、2 値分類 (顔/物、男性顔/女性顔、または自然物/人工物) を行った。両ネットワークは事前に別々に学習し [fMRI/バッチサイズ: 10; エポック: 最大 100(早期停止あり); 学習率. 0.001; オプティマイザ. アダム(REF)] 脳波/バッチサイズ: 10; エポック: 早期停止で最大 100; 学習率. 0.01; オプティマイザ. 確率勾配降下 (SGD)]、複合モデル (fMRI-EEG DNN) の畳み込み層の初期パラメータは事前学習したパラメータ (REF) に設定した。入力データセットは、トラリアルベースの fMRI ボリュームと脳波時系列データであり、以下のように用意した。(1) BOLD response の血流動態応答の遅れを考慮して、視覚刺激提示の開始から 6 秒後(3TR) の fMRI ボリュームを、刺激によって誘発された空間神経活動として収集した。入力 fMRI ボリュームは、79×95×79 のボクセルサイズで、標準的な MNI 空間 (fMRI の手順と前処

理のセクションを参照) に空間的に正規化され、脳外の信号強度はゼロにマスクした。そして、各ボクセル内の信号強度を体積内の信号強度の平均値と SD 値で z スコア化した。(2) また、同じ試行で 63 個の頭部チャンネルから前処理された脳波時系列データを収集した (脳波の手順と前処理の項を参照)。DNN 分類のために、刺激の開始から 1 秒間のデータ (250 サンプル、250Hz のサンプリングレート) を使用した。得られたセグメント化されたデータは、各チャンネルの信号振幅の平均値と SD 値によってさらに z スコア化を行った。最後に、EEGLAB の pop_eegfiltnew 関数を用いて、データをバンドパスフィルタリングし、5 つの周波数帯域のデータ (デルタ、1-4 Hz、シータ、4-8 Hz、アルファ、8-13 Hz、ベータ、13-30 Hz、ガンマ、30-70 Hz) を得た。全データセットには、間違えた試行とノイズの多い試行を除去した後、これを無作為に 10 のサブセットに分割し、10-fold クロスバリデーションテストを可能にした。

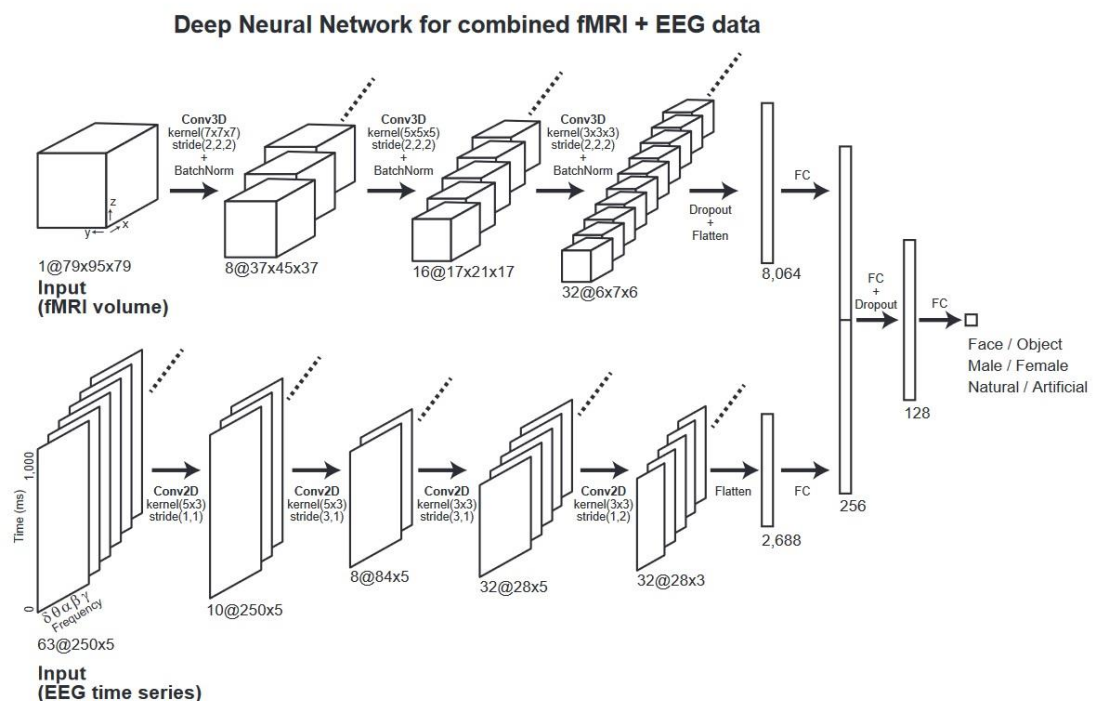


図 8 Deep Neural Network モデル 上部が fMRI 用、下部が脳波用の DNN モデルである。fMRI と脳波の特徴マップを統合し、視覚情報のカテゴリ間分類(顔/物)やカテゴリ内分類(男性顔/女性顔, 人工物/自然物)を行うモデルになっている。

第 3 章

結果

3.1 behavior 結果

課題の正答率を図 9 に示す。カテゴリ間（顔カテゴリ，物カテゴリ）では有意差(ウィルコクソン符号順位検定: $p = 4.19 \times 10^{-5}$)があったが、いずれの課題も平均 95 %以上という高いスコアになっていた。

また、同様に課題の反応速度をカテゴリごとに示す（図 10）。カテゴリ間（顔カテゴリ，物カテゴリ）とカテゴリ内（男性顔サブカテゴリ，女性顔サブカテゴリ）で有意差(ウィルコクソン符号順位検定: 顔カテゴリ vs 物カテゴリ, $p = 2.64 \times 10^{-8}$; 男性がおサブカテゴリ vs 女性顔サブカテゴリ, $p = 4.63 \times 10^{-4}$)がみられた。

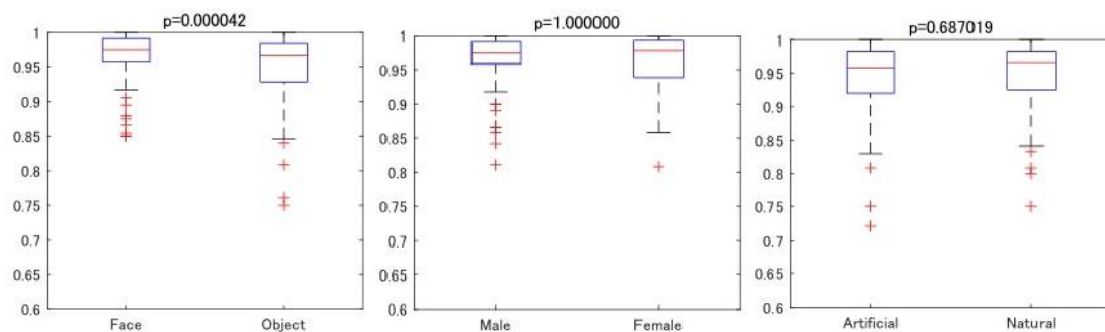


図 9 課題正答率 全参加者の課題正答率を、カテゴリ間またはカテゴリ内ごとに表示。

左：顔/物、中央：男性顔/女性顔、右：人工物/自然物。縦軸は正答率。いずれも平均は 95 %以上の正答率になっていた。また、カテゴリ間（顔カテゴリ，物カテゴリ）については有意差があった。

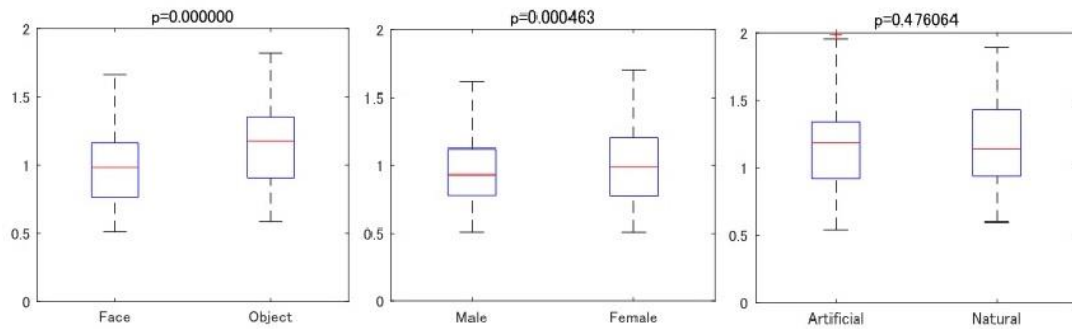


図 10 課題における画像提示からの反応速度 全参加者の画像提示からの応答速度について、カテゴリ間またはカテゴリ内ごとに表示。左：顔/物、中央：男性顔/女性顔、右：人工物/自然物。縦軸は秒。カテゴリ間（顔カテゴリ，物カテゴリ）とカテゴリ内（男性顔サブカテゴリ，女性顔サブカテゴリ）に有意差があった。

3.2 fMRI 単変量解析

3.2.1 顔カテゴリ - 物カテゴリ

fMRI 単変量解析のうち、顔カテゴリと物カテゴリのコントラストの結果を図 11 に示す。顔カテゴリ - 物カテゴリのコントラストでは FFA などの領域で有意な賦活が見られた(表 1)。一方で、物カテゴリ - 顔カテゴリのコントラストでは LO など有意な賦活が見られた(表 2)。

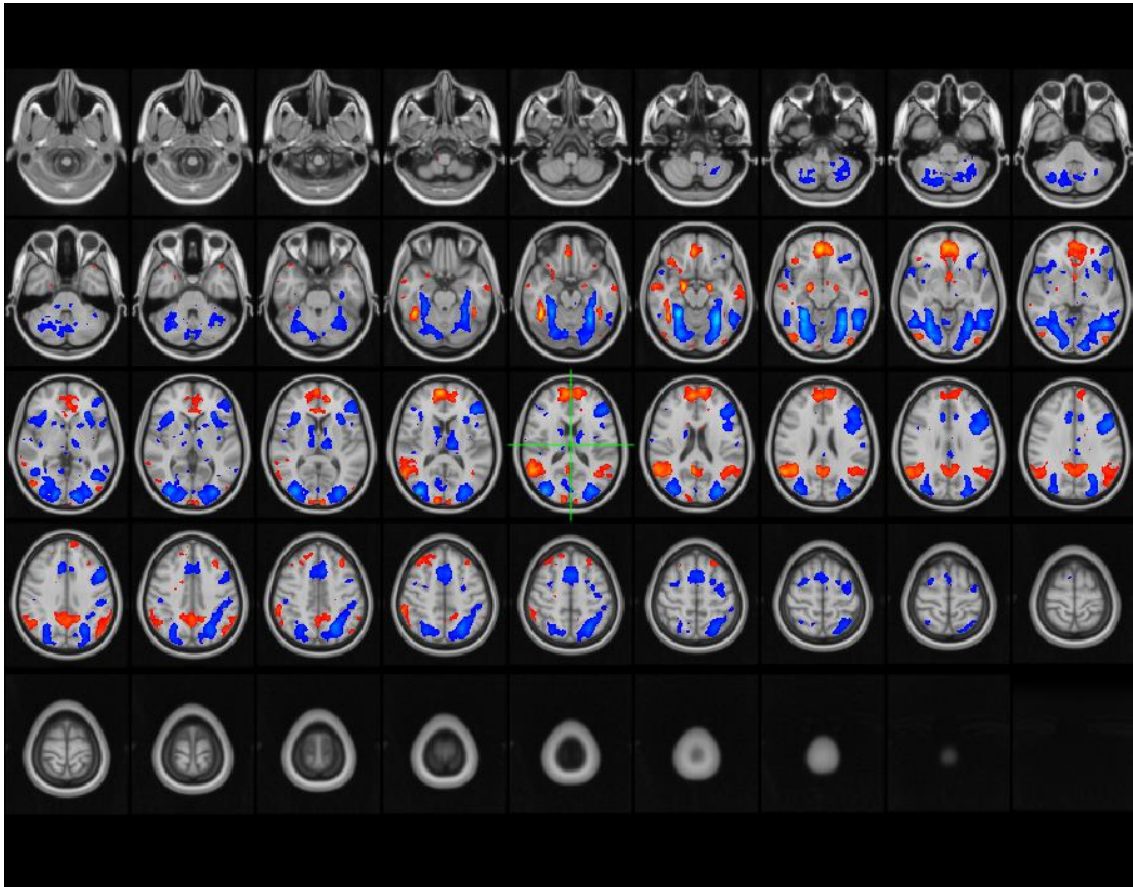


図 11 カテゴリ間（顔カテゴリ / 物カテゴリ）fMRI 単変量解析結果 水平断の解剖画像の上に統計結果(t 値) をカラーマップとして表示した。カテゴリ間(顔カテゴリ - 物カテゴリ)の解析結果のうち、顔カテゴリで有意により強く賦活した領域を赤、物カテゴリで有意により強く賦活した領域を青で表示した。

3.2.2 男性顔サブカテゴリ - 女性顔サブカテゴリ

男性顔サブカテゴリと女性顔サブカテゴリのコントラストの結果を図 12 に示す。男性顔サブカテゴリ - 女性顔サブカテゴリのコントラストにおいては、有意な賦活は見られなかった。女性顔サブカテゴリ - 男性顔サブカテゴリのコントラストの有意な賦活を表 3 に示す。

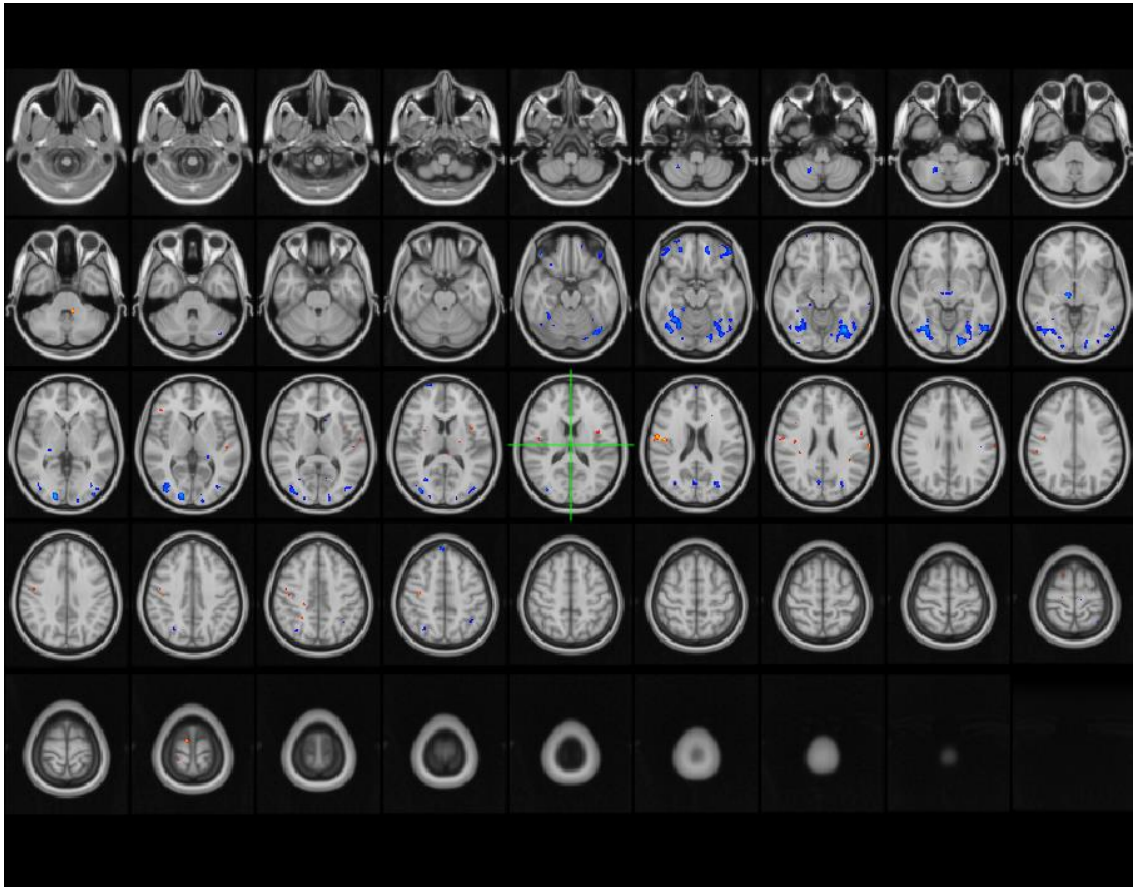


図 12 カテゴリ内（男性顔サブカテゴリ/女性顔サブカテゴリ）fMRI 単変量解析結果 水平断の解剖画像の上に統計結果(t 値) をカラーマップとして表示した。カテゴリ内(男性顔サブカテゴリ - 女性顔サブカテゴリ)の解析結果のうち男性顔サブカテゴリで有意により強く賦活した領域を赤、女性顔サブカテゴリにおいて有意により強くした領域を青で表示した。

3.2.3 自然物サブカテゴリ - 人工物サブカテゴリ

自然物サブカテゴリと人工物サブカテゴリのコントラストの結果を図 13 に示す。自然物サブカテゴリ - 人工物サブカテゴリのコントラストにおいては、有意な賦活は見られなかった。人工物サブカテゴリ - 自然物サブカテゴリのコントラストの有意な賦活を表 4 に示す。

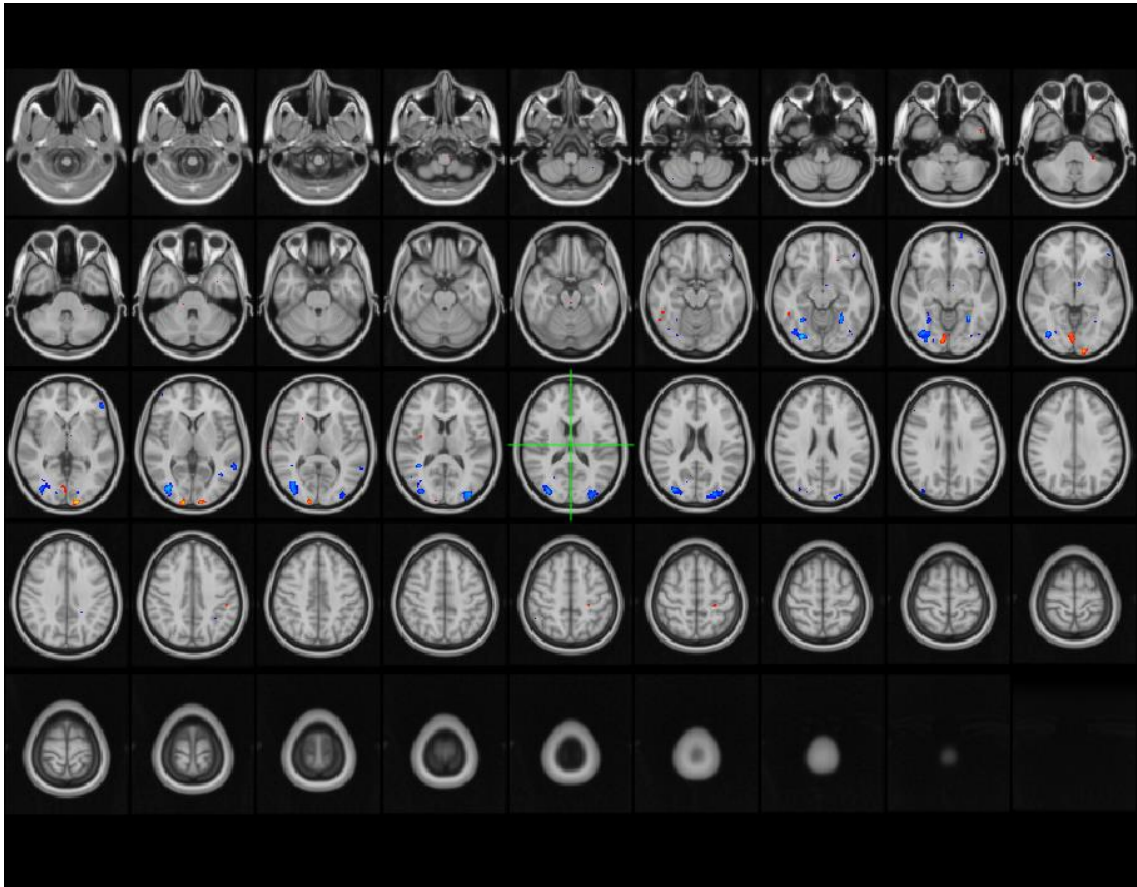


図 13 カテゴリ内（自然物サブカテゴリ/人工物サブカテゴリ）fMRI 単変量解析結果 水平断の解剖画像の上に統計結果(t 値) をカラーマップとして表示した。カテゴリ内(自然物サブカテゴリ, 人工物サブカテゴリ)の解析結果のうち自然物サブカテゴリで有意により強く賦活した領域を赤、人工物サブカテゴリで有意により強く賦活した領域を青で表示した。

3.3 脳波 ERP

3.3.1 顔カテゴリ/物カテゴリ

顔カテゴリ/物カテゴリの ERP 結果を図 14 に示す。複数のチャンネルにおいて波形の差分が視覚的に確認された。顔カテゴリ ERP から物カテゴリ ERP の差分において特徴的な negative 方向のピークがみられた(200 ms)。

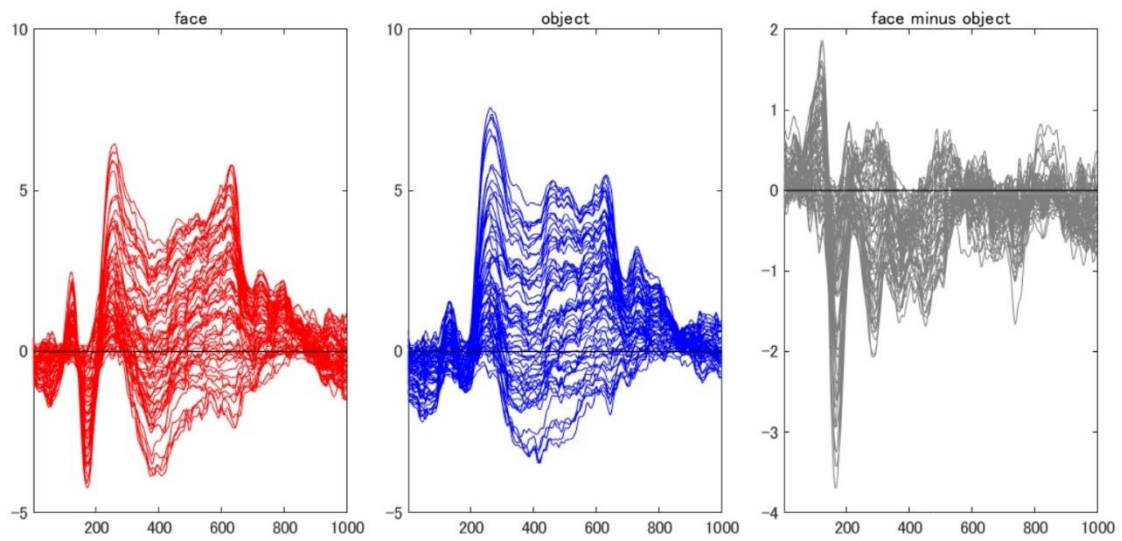


図 14 カテゴリ間（顔カテゴリ/物カテゴリ）ERP 画像提示から 1,000 ms のタイミングにおける ERP を 63 ch 重ねて表示した。顔カテゴリに対する各チャンネルの ERP を（赤色、左）、物カテゴリに対する各チャンネルの ERP を（青色、中央）でプロットしている。図右に両者の差分を表示した。

3.3.2 男性顔サブカテゴリ/女性顔サブカテゴリ

男性顔サブカテゴ/女性顔サブカテゴリの ERP 結果を図 15 に示す。2 つのサブカテゴリ間における特徴的な差分は確認できなかった。

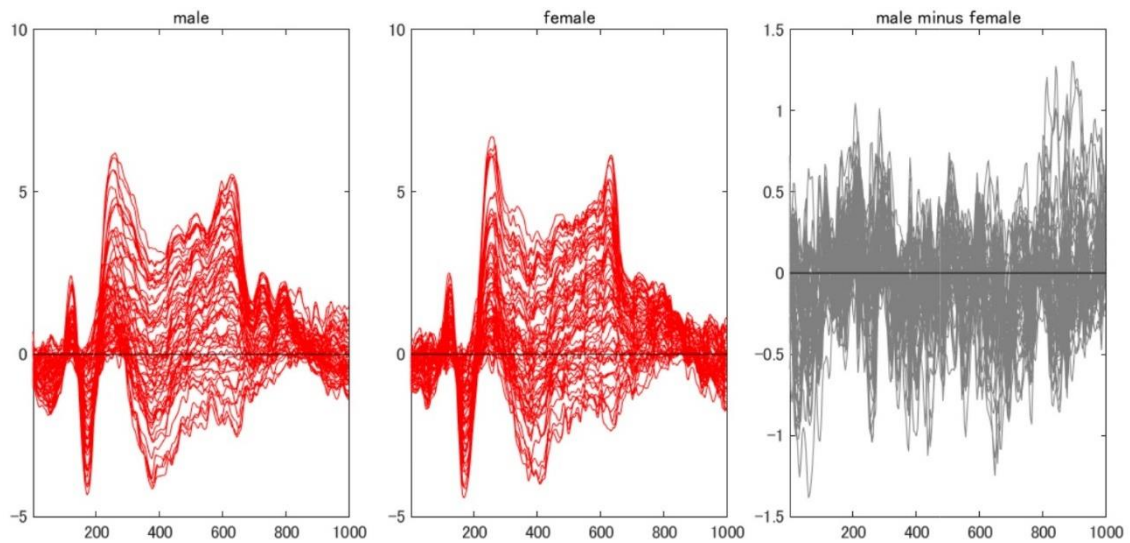


図 15 カテゴリ内（男性顔サブカテゴリ/女性顔サブカテゴリ）ERP 画像提示から 1,000 ms のタイミングにおける ERP を 63 ch 重ねて表示した(男性顔：左、女性顔：中央、差分：右)。

3.3.3 人工物サブカテゴリ/自然物サブカテゴリ

人工物サブカテゴリ/自然物サブカテゴリの ERP 結果を図 16 に示す。複数のチャンネルにおいて波形の差分が視覚的に確認された。自然物カテゴリ ERP から人工物カテゴリ ERP の差分において特徴的な Negative 方向のピークがみられた(200 ms)。

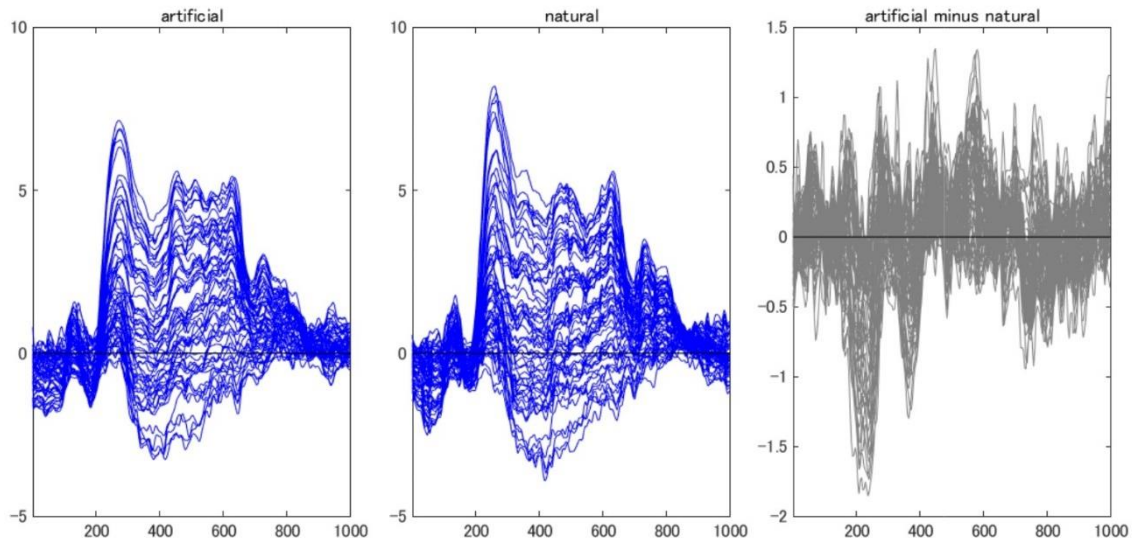


図 16 カテゴリ内（人工物サブカテゴリ/自然物サブカテゴリ）ERP 画像提示から 1,000 ms のタイミングにおける ERP を 63 ch 重ねて表示した(人工物サブカテゴリ：左、自然物サブカテゴリ：中央、差分：右)。

3.4 DNN 学習結果

3.4.1 DNN スコア

fMRI データの DNN 学習結果のスコアを図 17 に示す。カテゴリ間（顔/物）の平均識別精度は Train = 83.4 %、Validation = 75.8 %であった。Train の 1 標本 t 検定の結果は $p = 9.87\text{e-}13$ 、 $t = 55.63$ 、Validation の 1 標本 t 検定の結果は $p = 3.46\text{e-}17$ 、 $t = 174.11$ となり、chance level (50 %)に対して統計的に有意な差が認められた。

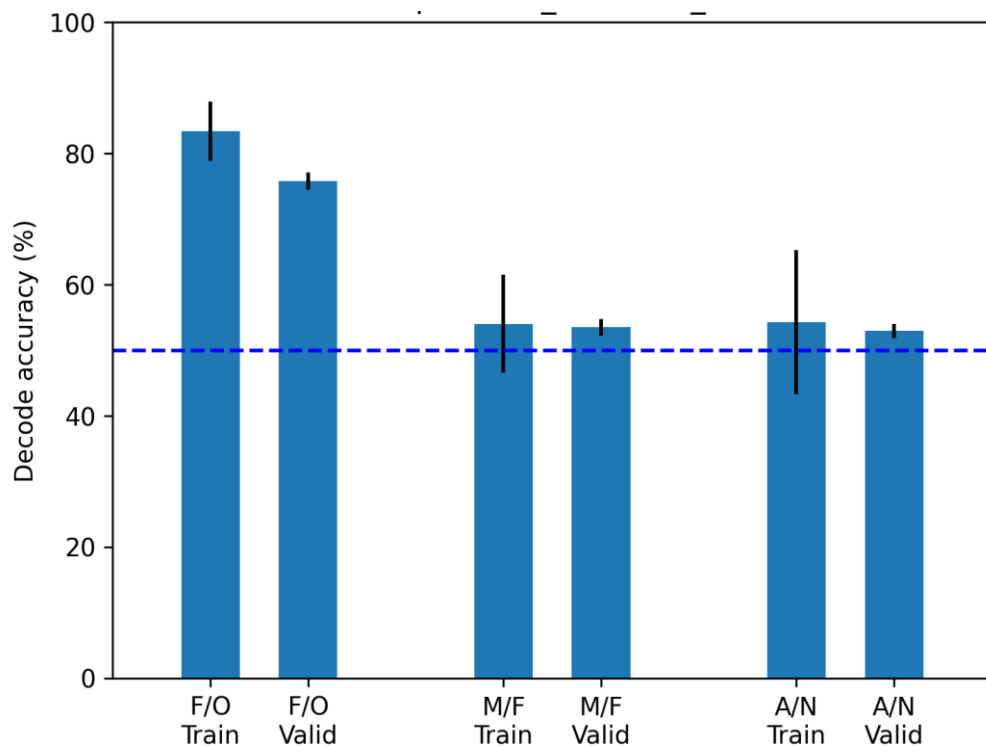


図 17 fMRI DNN 学習結果 今回の DNN モデルを用いた識別結果。カテゴリ間（顔/物カテゴリ）の平均識別精度は Train = 83.4 %、Validation = 75.8 %。カテゴリ内は男性顔/女性顔サブカテゴリが Train = 54.0 %、Validation = 53.5 %。人工物/自然物サブカテゴリが Train = 54.3 %、Validation = 53.0 %。エラーバーは標準偏差。

EEG データの DNN 学習結果のスコアを図 18 に示す。カテゴリ内（男性顔/女性顔）の平均識別精度は Train = 73.3 %、Validation = 53.4 %であった。Train の 1 標本 t 検定の結果は $p = 1.82e-07$ 、 $t = 14.19$ 、Validation の 1 標本 t 検定の結果は $p = 1.06e-16$ 、 $t = 153.66$ となり、chance level (50 %)に対して統計的に有意な差が認められた。

カテゴリ内（人工物/自然物）の平均識別精度は Train = 74.9 %、Validation = 54.0 %であった。Train の 1 標本 t 検定の結果は $p = 7.21e-07$ 、 $t = 12.09$ 、Validation の 1 標本 t 検定の結果は $p = 8.18e-18$ 、 $t = 204.37$ となり、chance level (50 %)に対して統計的に有意な差が認められた。

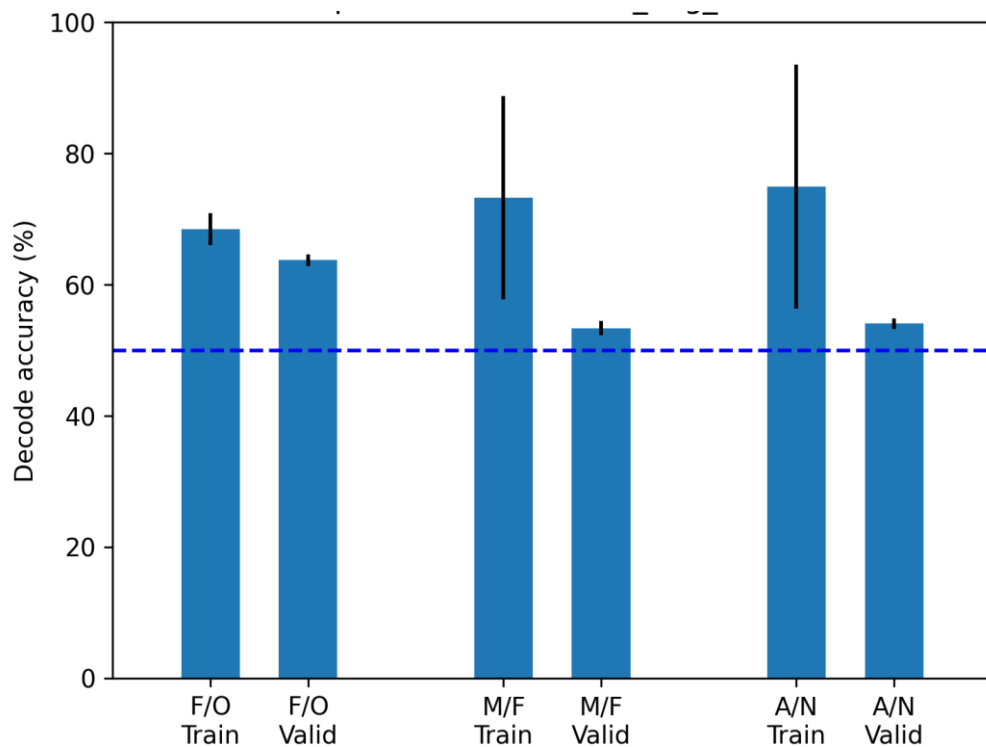


図 18 脳波 DNN 学習結果 今回の DNN モデルを用いた識別結果。カテゴリ間（顔/物カテゴリ）の平均識別精度は Train = 68.5 %、Validation = 63.7 %。カテゴリ内は男性顔/女性顔サブカテゴリが Train = 73.3 %、Validation = 53.4 %。人工物/自然物サブカテゴリが Train = 74.9 %、Validation = 54.0 %。エラーバーは標準偏差。

fMRI データと EEG データを両方使用した DNN 学習結果のスコアを図 19 に示す。カテゴリ間（顔/物）の平均識別精度は Train = 78.1 %、Validation = 77.2 %であった。Train の 1 標本 t 検定の結果は $p = 8.91e-12$ 、 $t = 43.52$ 、Validation の 1 標本 t 検定の結果は $p = 2.71e-18$ 、 $t = 231.10$ となり、chance level (50 %)に対して統計的に有意な差が認められた。

カテゴリ内（男性顔/女性顔）の平均識別精度は Train = 68.0 %、Validation = 53.8 %であった。Train の 1 標本 t 検定の結果は $p = 3.65e-06$ 、 $t = 9.98$ 、Validation の 1 標本 t 検定の結果は $p = 1.11e-13$ 、 $t = 70.97$ となり、chance level (50 %)に対して統計的に有意な差が認められた。

カテゴリ内（人工物/自然物）の平均識別精度は Train = 64.9 %、Validation = 53.9 %であった。Train の 1 標本 t 検定の結果は $p = 1.50e-06$ 、 $t = 11.10$ 、Validation の 1 標本 t 検

定の結果は $p = 1.52e-11$ 、 $t = 41.01$ となり、chance level (50 %) に対して統計的に有意な差が認められた。

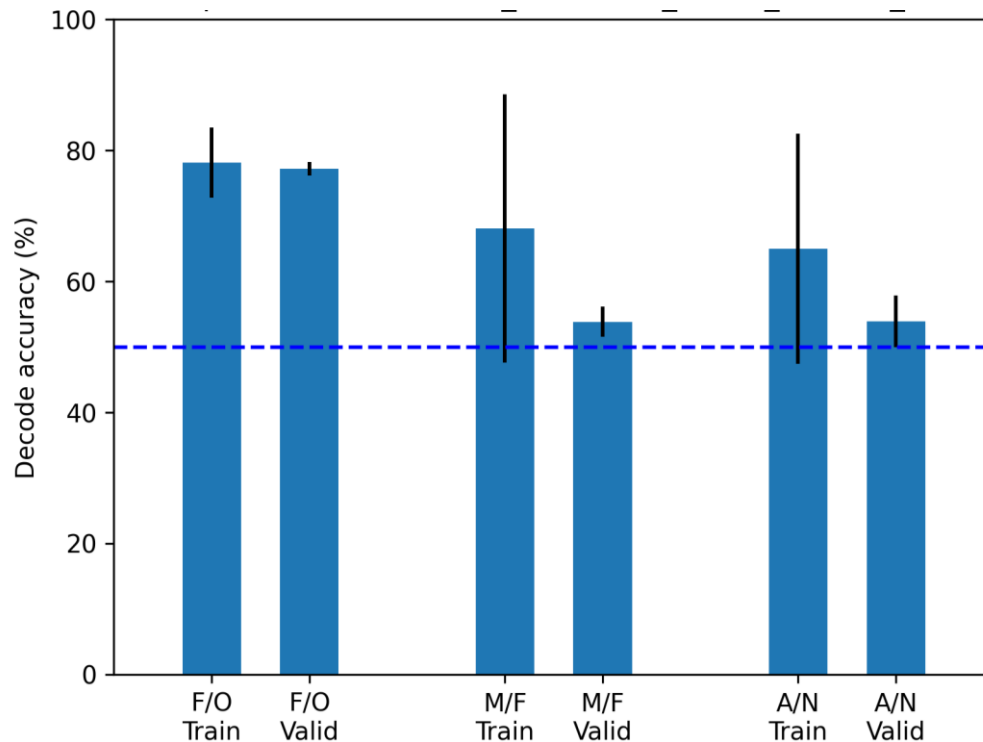


図 19 combined DNN 学習結果 今回の DNN モデルを用いた識別結果。カテゴリ間（顔/物カテゴリ）の平均識別精度は Train = 78.1 %、Validation = 77.2 %。カテゴリ内は男性顔/女性顔サブカテゴリが Train = 68.0 %、Validation = 53.8 %。人工物/自然物サブカテゴリが Train = 64.9 %、Validation = 53.9 %。エラーバーは標準偏差。

3.4.2 Grad CAM による識別特徴マップの可視化

fMRI と脳波どちらがどの程度 DNN 識別結果に影響を与えたかを Grad CAM により図示した。カテゴリ間識別（顔カテゴリ/物カテゴリ）の Grad CAM 結果について fMRI 結果を図 20、脳波結果を図 21 に示す。図 20 と図 21 の weight のスケールの違いから、fMRI の方が脳波より強く学習に影響を与えていることが分かった。特に fMRI の結果からは、顔カテゴリ識別では FFA 領域を、物カテゴリ識別では両側の LO 領域などが学習に影響を与えていることが分かった。一方で脳波についてあまり時間的特徴は見られなかった。

カテゴリ内識別（男性顔サブカテゴリ/女性顔サブカテゴリ）の Grad CAM 結果を図 22 と図 23 に、カテゴリ内識別（人工物サブカテゴリ/自然物サブカテゴリ）の Grad CAM 結果を図 24 と図 25 に示す。どちらも weight のスケールの違いから、脳波の方が fMRI より強く学習に影響を与えていることが分かった。しかしカテゴリ内における時間的特徴の違いについて、カテゴリ内の結果どちらについても見られなかった。

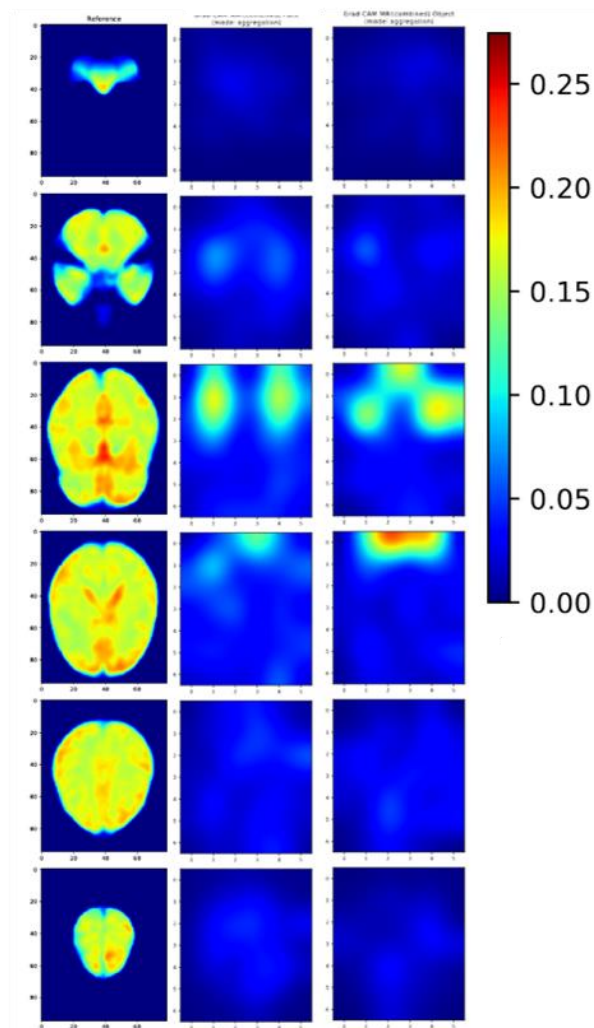


図 20 カテゴリ間識別（顔カテゴリ/物カテゴリ）における fMRI の Grad CAM 結果 左列は脳画像の水平段を並べた reference 画像（画像の下が脳前方、画像の右が脳右方向）。中央列が識別結果を顔とした際、右列が識別結果を物とした際それぞれ識別に貢献した領域を表している。Grad CAM の反応部位は reference の水平段と対応した位置を示している。右のカラーバーは weight を表している。

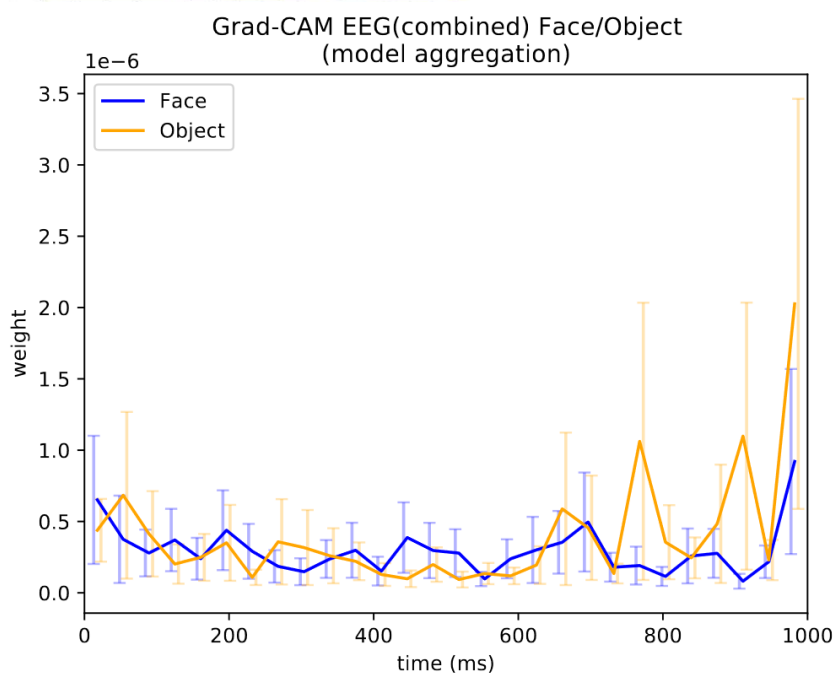


図 21 カテゴリ間識別（顔カテゴリ/物カテゴリ）における脳波の Grad CAM 結果 青が顔、オレンジが物と識別した際に、識別に貢献した脳波データの時間を表している。横軸が ms、縦軸が weight。

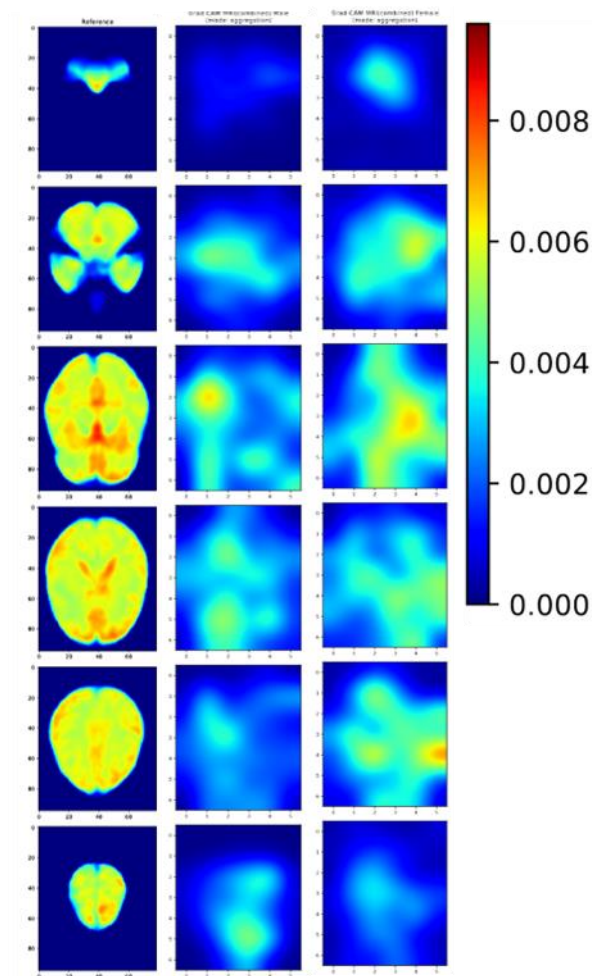


図 22 カテゴリ内識別（男性顔サブカテゴリ/女性顔サブカテゴリ）における fMRI の Grad CAM 結果 図の説明は図 20 と同様。中央列が男性顔、右列が女性顔と識別した際における識別に貢献した領域。

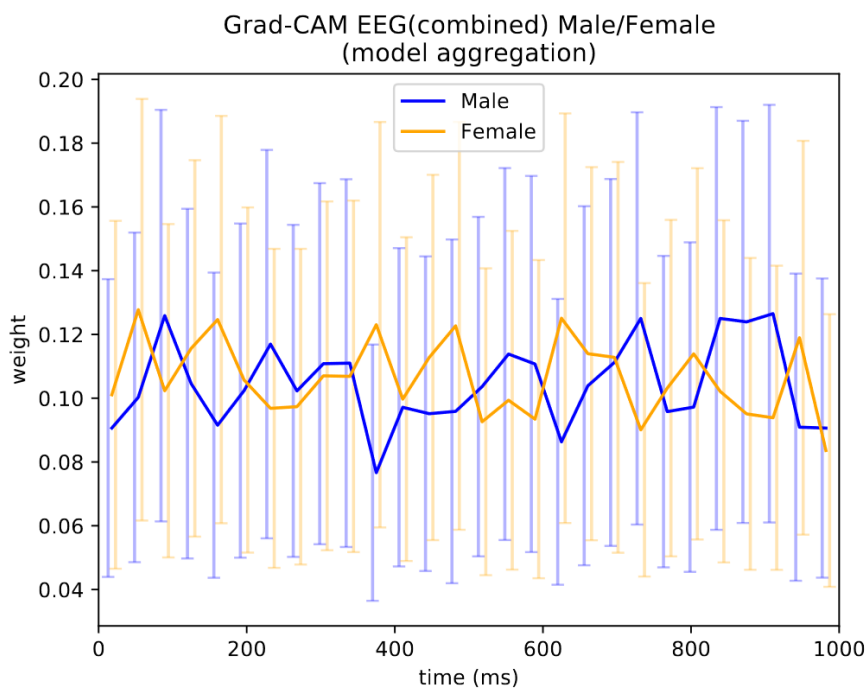


図 23 カテゴリ内識別（男性顔サブカテゴリ/女性顔サブカテゴリ）における脳波の Grad CAM 結果 図の説明は図 21 と同様。青が男性顔、オレンジが女性顔。

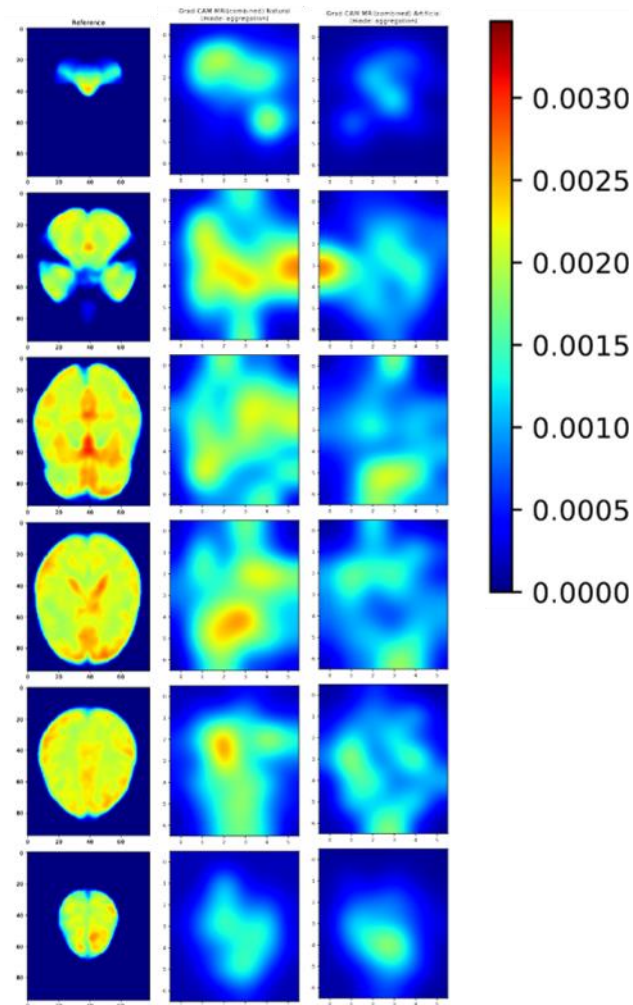


図 24 カテゴリ内識別（自然物サブカテゴリ/人工物サブカテゴリ）における fMRI の Grad CAM 結果 図の説明は図 20 と同様。中央列が自然物、右列が人工物と識別した際における識別に貢献した領域。

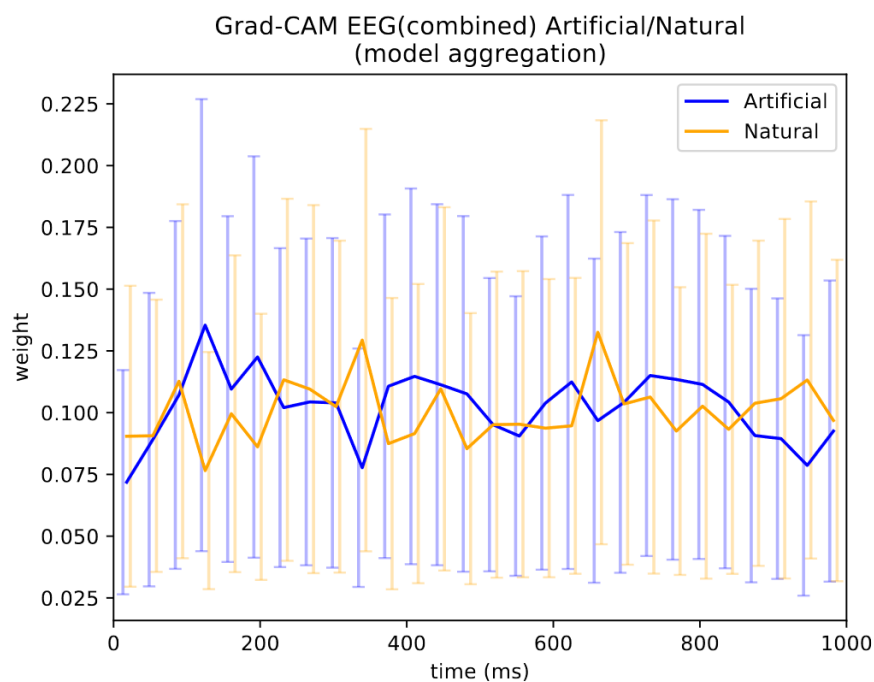


図 25 カテゴリ内識別（自然物サブカテゴリ/人工物サブカテゴリ）における脳波の Grad CAM 結果 図の説明は図 21 と同様。青が男性顔、オレンジが女性顔。

表 1 fMRI 「顔カテゴリ - 物カテゴリ」コントラストによる賦活領域

Cluster Index	Max t	Max x (mm)	Max y (mm)	Max z (mm)	Harvard-Oxford Atlas
12	8.84	2	48	-10	R. Frontal Medial Cortex
12	8.12	6	58	12	R. Frontal Pole
12	7.71	-10	60	18	L. Frontal Pole
12	5.72	-12	48	8	L. Paracingulate Gyrus
12	5.15	-4	38	0	L. Cingulate Gyrus, anterior division
12	4.98	4	32	8	R. Cingulate Gyrus, anterior division
12	4.63	2	24	-2	Subcallosal Cortex
12	4.19	-6	54	34	L. Superior Frontal Gyrus
11	7.31	56	-56	24	R. Angular Gyrus
11	6.53	40	-56	22	R. Angular Gyrus
11	5.88	44	-80	-10	R. Lateral Occipital Cortex, inferior division
11	5.8	62	-40	48	R. Supramarginal Gyrus, posterior division
11	5.32	48	-56	36	R. Angular Gyrus
11	4.37	46	-40	10	R. Supramarginal Gyrus, posterior division
11	4.24	58	-48	40	R. Angular Gyrus
11	3.98	64	-44	2	R. Middle Temporal Gyrus, temporooccipital part
11	3.73	50	-62	50	R. Lateral Occipital Cortex, superior division
10	6.86	6	-56	22	R. Precuneous Cortex
10	6.11	-16	-50	32	L. Cingulate Gyrus, posterior division
10	5.74	4	-56	34	R. Precuneous Cortex
10	4.06	18	-44	34	R. Cingulate Gyrus, posterior division
9	5.55	-50	-52	20	L. Angular Gyrus
9	5.41	-38	-60	18	L. Angular Gyrus
9	5.33	-48	-70	34	L. Lateral Occipital Cortex, superior division
9	4.93	-58	-56	38	L. Angular Gyrus
9	4.47	-60	-60	26	L. Angular Gyrus
9	4.27	-64	-42	34	L. Supramarginal Gyrus, posterior division
9	3.79	-50	-44	32	L. Supramarginal Gyrus, posterior division

表 2 fMRI 「物カテゴリ – 顔カテゴリ」コントラストによる賦活領域

Cluster Index	Max t	Max x (mm)	Max y (mm)	Max z (mm)	Harvard-Oxford Atlas
9	15.5	-28	-54	-18	L. Temporal Occipital Fusiform Cortex
9	12.2	-48	-62	-8	L. Lateral Occipital Cortex, inferior division
9	11.8	-26	-70	-12	L. Occipital Fusiform Gyrus
9	11.5	-28	-42	-20	L. Temporal Fusiform Cortex, posterior division
9	10.2	-30	-86	4	L. Lateral Occipital Cortex, inferior division
9	9.84	-28	-82	18	L. Lateral Occipital Cortex, superior division
9	9.62	-12	-94	-2	L. Occipital Pole
9	9.42	-24	-64	46	L. Lateral Occipital Cortex, superior division
9	8.13	-32	-30	-20	L. Temporal Fusiform Cortex, Posterior division
9	7.97	-26	-68	34	L. Lateral Occipital Cortex, superior division
9	7.9	-24	-86	30	L. Lateral Occipital Cortex, superior division
9	6.93	-34	-50	44	L. Superior Parietal Lobule
9	6.86	-46	-30	38	L. Supramarginal Gyrus, anterior division
9	6.07	-24	-62	-28	L. Cerebellum
9	5.36	-12	-76	40	L. Precuneous Cortex
9	5.18	-20	-68	62	L. Lateral Occipital Cortex, superior division
9	4.67	10	-70	14	R. Intracalcarine Cortex
9	4.42	-40	-60	-30	L. Cerebellum
9	4.18	2	-82	10	R. Supracalcarine Cortex
9	4.07	-54	-52	2	L. Middle Temporal Gyrus, temporooccipital part
9	3.93	-50	-38	56	L. Supramarginal Gyrus, anterior division
9	3.79	-40	-48	-36	L. Cerebellum
8	16.3	28	-46	-16	R. Temporal Occipital Fusiform Cortex
8	14.3	38	-84	14	R. Lateral Occipital Cortex, superior division
8	13.5	28	-62	-12	R. Occipital Fusiform Gyrus
8	10.7	24	-78	-10	R. Occipital Fusiform Gyrus
8	10.1	48	-60	-8	R. Inferior Temporal Gyrus, temporooccipital part
8	9.18	16	-88	0	R. Occipital Pole

表 3 fMRI 「女性顔サブカテゴリ – 男性顔サブカテゴリ」 コントラストによる賦活領域

Cluster Index	Max t	Max x (mm)	Max y (mm)	Max z (mm)	Harvard-Oxford Atlas
2	5.06	-16	-88	-6	L. Lingual Gyrus
2	5.01	-32	-70	-12	L. Occipital Fusiform Gyrus
2	4.55	-54	-68	-8	L. Lateral Occipital Cortex, inferior division
2	4.49	-18	-80	-16	L. Occipital Fusiform Gyrus
2	4.28	-38	-90	0	L. Lateral Occipital Cortex, inferior division
2	4.23	-32	-56	-18	L. Temporal Occipital Fusiform Cortex
2	4.2	-44	-70	-18	L. Lateral Occipital Cortex, inferior division
2	3.9	-16	-96	8	L. Occipital Pole
2	3.75	-48	-56	-16	L. Cerebellum
2	3.49	-32	-82	-8	L. Lateral Occipital Cortex, inferior division
1	4.77	38	-70	-16	R. Occipital Fusiform Gyrus
1	4.63	32	-54	-16	R. Temporal Occipital Fusiform Cortex
1	4.55	36	-82	14	R. Lateral Occipital Cortex, superior division
1	4.35	26	-68	-14	R. Occipital Fusiform Gyrus
1	4.23	44	-70	-4	R. Lateral Occipital Cortex, inferior division
1	3.94	24	-82	-6	R. Occipital Fusiform Gyrus
1	3.92	54	-60	-6	R. Lateral Occipital Cortex, inferior division
1	3.55	30	-38	-18	R. Temporal Fusiform Cortex, posterior division

表 4 fMRI 「人工物サブカテゴリ – 自然物サブカテゴリ」 コントラストによる賦活領域

Cluster Index	Max t	Max x (mm)	Max y (mm)	Max z (mm)	Harvard-Oxford Atlas
2	5.66	30	-78	-12	R. Occipital Fusiform Gyrus
2	5.26	34	-78	4	R. Lateral Occipital Cortex, inferior division
2	4.99	34	-82	18	R. Lateral Occipital Cortex, superior division
2	3.32	36	-62	-14	R. Occipital Fusiform Gyrus
1	4.89	-34	-90	12	L. Lateral Occipital Cortex
1	4.03	-18	-94	22	R. Occipital Pole

第 4 章

考察

4.1 結果のまとめと仮説について

fMRI 単変量解析の結果について、カテゴリ間では顔カテゴリは FFA、物カテゴリは LO において賦活がみられた。一方でカテゴリ内については有意な賦活は見られなかった。これは fMRI ではカテゴリ間の違いの検出が可能だが、カテゴリ内の検出は困難であるという仮説(1)を支持しているといえる。

脳波 ERP の結果について、カテゴリ間コントラスト(顔カテゴリ ERP – 物カテゴリ ERP)と、カテゴリ内コントラスト(自然物サブカテゴリ ERP – 人工物サブカテゴリ ERP)において特徴的な波形の差分が観測された。これはカテゴリ内の検出に脳波が貢献するという仮説(2)を部分的に支持しているといえる。

DNN 学習結果のスコアのうち、カテゴリ間は Validation = 77.2 % と高いスコアになっていた。一方でカテゴリ内の男性顔/女性顔サブカテゴリは Validation = 53.8 %、人工物/自然物サブカテゴリは Validation = 53.9 % であり、chance level (50 %) に対して統計的に有意な差はあったものの、カテゴリ間と比較して低いスコアとなっていた。

Grad CAM による識別特徴マップの可視化について、カテゴリ間については fMRI、カテゴリ内については脳波のデータが貢献していることが分かった。この結果は fMRI がカテゴリ間、脳波がカテゴリ内の識別に貢献をするという仮説を支持する結果になった。

4.2 過去の研究との比較

カテゴリ間について、fMRI 単変量解析の結果が先行研究と一致しており、DNN 学習結果でも高いスコア (Validation = 77.2 %) が得られたといえる。また、Grad CAM を用いた識別特徴の可視化においても fMRI 単変量解析と賦活領域の一致が見られた。我々が調べた範囲では、trial based によるデコーディングは既存の研究ではほとんど行われておらず、run

based のものが大半である[10-12]。run based のデコーディングに比べて trial based のデコーディングは S/N 比が低いため、本研究で一定の結果が得た要因の 1 つは近年開発された深層学習をデコーダとして使用したことが考えられる。また、trial based によるデコーディングという特徴を活かして、例えば自身では意思の表示が困難な患者の認知プロセスを可視化してコミュニケーションを可能とするといった応用も考えられる。

4.3 アイデンティティごとの識別について

デコーディングによる識別を行う研究においては、1000 種を超える視覚刺激を用いることで、同一の画像が提示されることのない実験デザインから構成されるものもある。本研究ではそのような実験デザインを用いず、8 種×5 角度からなる計 40 種の画像を使用した。その理由として、現在設定しているカテゴリ内よりも更に詳細な分類である「アイデンティティの識別」を考慮していたためである。

今回はカテゴリ内の識別結果があまり高くなかったため、より識別が難しいとされるアイデンティティの識別までは行わなかった。

4.4 デコーディング成績について

カテゴリ内の視覚情報デコーディングは、chance level (50 %) に対して統計的に有意な差はあったものの、実用的な数値には至らなかった。この理由としては、前述したように trial based の解析では脳波の S/N 比が十分ではないことが挙げられる。実際、平均波形である ERP で（物サブカテゴリ内は特徴的な差分が見られたが）顔サブカテゴリ内では大きな波形の違い（男性顔に対する反応と女性顔に対する反応の違い）が見られなかった。ただし、Grad CAM による可視化の結果、カテゴリ内については脳波が識別に fMRI より強く貢献していることが分かった。しかし、特定の時間パターンにおいて特徴がみられることはなかった。

4.5 今後の展開

DNN 識別結果については、モデルのハイパーパラメータの検証が不十分であり、最善のモデル探索が行えていなかった可能性がある。この点について、今後 grid search などを用いた探索を行う必要がある。また、現在の解析ではカテゴリ内（特に顔カテゴリ）の特徴の違いを視覚的に十分に捉えられているとはいえなかった。そこで、視覚サブカテゴリごとの

信号源推定を追加で行うことを考えている。今回は脳波の解析について、視覚サブカテゴリごとの ERP による定性的な評価のみ行ったが、追加の検証として周波数解析を行うことが挙げられる。また、run based によるデコーディング結果との学習スコアの比較を行うことも考えられる。

謝辞

本研究に関わられた皆様に感謝いたします。様々なご指導を頂いた高知工科大学情報学群中原潔教授に感謝申し上げます。また、実験指導および論文の添削のご指導を頂いた竹田真己特任教授、ディープニューラルネットワークの構築と検証にご助力頂いた三好康祐氏に心より感謝申し上げます。本研究に参加いただいた 53 名の学生の皆様にも心より感謝を申し上げます。

参考文献

- [1] Grill-Spector, K. & Weiner, K. S. The functional architecture of the ventral temporal cortex and its role in categorization. *Nat Rev Neurosci* 15, 536-548, (2014).
- [2] Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17, 4302-4311, (1997).
- [3] Epstein, R. & Kanwisher, N. A cortical representation of the local visual environment. *Nature* 392, 598-601, (1998).
- [4] Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *PNAS* 92, 8135-8139, (1995).
- [5] Grill-Spector, K., Kourtzi, Z. & Kanwisher, N. The lateral occipital complex and its role in object recognition. *Vision Res* 41, 1409-1422, (2001).
- [6] Miyakawa, N., Majima, K., Sawahata, H., Kawasaki, K., Matsuo, T., Kotake, N., Suzuki, T., Kamitani, Y. & Hasegawa, I. Heterogeneous redistribution of facial subcategory information within and outside the face-selective domain in primate inferior temporal cortex. *Cereb Cortex* 28, 1416-1431, (2018).
- [7] Huster, R. J., Debener, S., Eichele, T., & Herrmann, C. S. Methods for simultaneous EEG-fMRI: an introductory review. *J Neurosci.*, 32, 6053-6060, (2012).
- [8] Shen, G., Horikawa, T., Majima, K. & Kamitani, Y. Deep image reconstruction from human brain activity. *PLoS Comput Biol* 15, e1006633, (2019).
- [9] Stigliani, A., Weiner, K. S. & Grill-Spector, K. Temporal processing capacity in high-level visual cortex is domain specific. *J Neurosci* 35, 12412-12424, (2015).
- [10] Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D. & Brovelli, A. Multivoxel Pattern Analysis for fMRI Data: A Review. *Comput Math Methods Med*, 961257, (2012).

- [11] Cichy, R. M., & Pantazis, D. Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. *NeuroImage* 158, 441-454, (2017).
- [12] Cichy, R., Pantazis, D. & Oliva, A. Resolving human object recognition in space and time. *Nat Neurosci* 17, 455–462 (2014).