

研究ノート

ディープラーニングを用いた顔表情に対する 脳反応と性格特性指標との相関に関する研究

庭野賀津子、田邊素子、庭野道夫

東北福祉大学

要旨

言語獲得前の乳児の顔表情は感情や欲求の表出手段であり、養育者にとって養育行動を取る上での重要な情報源となる。しかし、笑顔や泣き顔などの顔表情を見たときに喚起される情緒的反応には個人差が認められ、様々な要因が脳内の活動に影響を与えていると考えられる。その個人差が生じる要因の一つとして、個人の性格特性がある。そこで本研究では、青年期成人を対象として、乳児及び成人の表情顔に対する脳反応を機能的近赤外線分光法（fNIRS）で測定し、ビッグ・ファイブの5次元性格特性との関連性をディープラーニング（深層学習）の手法を用いて検討した。前頭前野の脳賦活のパターンを入力信号、性格特性の各指標値を出力信号として、畳み込みニューラルネットワーク（CNN）の手法で解析した結果、計測された脳賦活パターンから性格指標を20%以内の誤差で予測できることが分かった。また、予測精度はビッグ・ファイブの各次元によって異なり、各次元のスコアの標準誤差に依存した。この結果から、顔刺激に対する脳反応が個人の性格特性によって異なることが示唆された。今後、顔刺激以外の視覚刺激に対しても同様の実験を行い、脳賦活という生理的反応と性格特性という心理的要因の相関をディープラーニングの手法を用いてより詳細に明らかにしていく予定である。

キーワード：機能的近赤外線分光法（fNIRS）、ディープラーニング、顔表情認知、ビッグ・ファイブ、ニューラルネットワーク

1. 緒言

言語獲得前の乳児の顔表情は基本的な感情や欲求の表出手段であり、養育者にとって重要な情報源となる。乳児の笑顔は養育者に幸福感をもたらす愛着や養育行動を促進させる一方、乳児の泣き顔や泣き声等のネガティブな表情から否定的感情が惹起され育児不安やストレスにつながるケースもある。我々はこれまで育児経験の無い青年期成人が乳児の顔表情を見ている時の脳活動やその他の生理心理的反応について研究してきた。非侵襲的脳機能計測法である機能的近赤外線分光法（fNIRS）によって乳児顔表情認知時の脳反応を測定したところ、顔表情に対する脳反応には個人差が認められ、様々な要因がその個人差を生じさせていることが示唆された¹⁾。また、顔表情刺激に対する脳反応を、機能的核磁気共鳴画像法（fMRI）によって測定するとともに心理検査指標との相関を調べた結果、脳反応は個人の性格特性による影響を受けることが明らかとなった²⁾。さらに乳児の顔表情に対する脳反応をfMRIによって乳児の保育体験の前後に測定したところ、脳の賦活部位や強度が変化することが明らかとなり、顔表情に対する脳反応の仕方は育児経験の有無によって影響を受け、変化することが示唆された³⁾。そのため、乳児の顔表情への脳反応の個人差とその要因をさらに詳細に検討することにより、個人の特性に対応した親への育児支援や、産後うつ予防及び乳幼児への虐待防止のための基礎資料が得られると考える。

近年、脳の情報処理機構を解明する目的で機械学習を活用する研究が盛んになりつつある。機械学習の技法の進展を背景に、fNIRS のデータ解析に機械学習を適用する試みが増えてきている。例えば、fNIRS データを畳み込みニューラルネットワーク (CNN) の手法で解析することで被験者の性別を分類することが試みられている⁴⁾。また、痛みのバイオマーカーの同定のために、fNIRS と機械学習を使用する試みもなされている⁵⁾。痛みは脳が感じるものであり、その痛みの度合いを定量化できれば、痛みの客観的な評価法の開発に貢献すると期待される。

そこで本研究では、顔表情に対する脳活動を fNIRS で測定した結果とビッグ・ファイブ性格特性との関連をディープラーニング (深層学習) の手法を用いて解析し、明らかにすることを目的とする。なお、fNIRS は fMRI と比較して、対象者の身体的拘束性が低く、座位や立位などの自然な姿勢のまま測定でき、測定場所も特別な検査室で実施する必要はないことから、対象者への精神的圧迫感や検査中の騒音、あるいは閉所恐怖症等によるデータへの影響を低く抑えることができる。また、本研究における我々の主たる関心領域は前頭前皮質である。そこで我々は本研究では、脳機能測定の方法として fNIRS を選択した。

2. 研究方法

2.1 研究対象者

対象者は、医療系学部に在籍する健常な大学生50名 (男女各25名、平均年齢 21.3 ± 0.7 歳) である。全員右利きで、裸眼視力もしくはメガネ等による矯正視力が両眼で0.7以上であった。対象者には、本研究内容について文書と口頭で十分な説明を行い、研究参加について書面にて同意を得た。本研究の実施に先立ち、東北福祉大学研究倫理委員会の審査・承認を受けた (承認番号: RS181101)。

2.2 性格検査

各対象者の心理的評価のために、NEO-PI-R の日本語版を用いた。NEO-PI-R は質問項目が全240問からなる、ビッグ・ファイブ人格5因子モデルに基づく質問紙法的人格テストで、個人の性格特性を、開放性 (O: Openness)、誠実性 (C: Conscientiousness)、外向性 (E: Extraversion)、調和性 (A: Agreeableness)、神経症傾向 (N: Neuroticism) の5つの次元でとらえるものである。この5因子モデルにおいては人格を表す用語には5因子が見られるとしており、その5因子は人格構造において普遍的なものとして、性格特性を測定するために世界的に広く使用されている⁶⁾。質問紙の各項目は、「そう思わない」と「そう思う」を両端とする5段階のリッカート尺度に基づいて評価され、5つの次元ごとにスコアが算出される。

NEO-PI-R の開発者である Costa ら⁷⁾ 及び日本語版の開発者の下仲ら⁸⁾ によると、5次元の基本的定義は以下の通りである。

- (1) 神経症傾向 (N) : 精神科の臨床においてもっとも馴染みのある特性であり、この N 得点が高いと、恐怖、悲しみ、情緒的不安定の傾向が強い。心理的ストレスを受けやすく、不安症や統合失調症のボーダーラインとの強い関連性が示唆される。また、この N 得点が低いと、精神的に安定していると言える。
- (2) 外向性 (E) : E 得点が高いと、社会的で上昇志向であり、活動的である。E 得点が低いと内向的と言えるが、外向的の対極ということではなく外向性が欠如しており控えめで依存心が高いとされている。
- (3) 開放性 (O) : O 得点が高いと、開放性が高く、芸術に対してより想像力があり、知的好奇心が強く、行動に柔軟性がある。O 得点が低いと、保守的で外見も控えめであり、情動的反応も乏しいとされている。

- (4) 調和性 (A) : A 得点が高いと、調和性があり、同情的で、社会的にも好ましいとされる一方、極端に高いと心理的健康度からみて好ましくない。A 得点が低いと冷淡で、自分の信念を貫き、敵対的である。
- (5) 誠実性 (C) : C 得点が高いと、目的を持ち、意志が強く、学業や職業の達成に向けてがんばるといふポジティブな面と、気難しさやワーカホリックにつながるというネガティブな面を持ち合わせている。C 得点が低いと、目標に向かってがんばるひたむきさに欠け、快楽主義であるとされる。

2.3 顔表情認知課題と実験手順

顔表情認知課題（タスク）における刺激呈示は、Platform of Stimuli and Tasks（日立製作所中央研究所開発ソフトウェア）を用いた。刺激に用いた画像は、乳児（Infant）と成人（Adult）それぞれ、①笑い顔または喜び顔（以下、happy）、②泣き顔または悲しみ顔（以下、sad）、③快・不快の不明確な曖昧顔または中性顔（以下、neutral）の3種類の顔表情を正面から撮影した静止画である。乳児画像は、我々の過去の研究で使用した顔画像セット³⁾の4名（男女各2名）の乳児画像を用いた。一方、成人画像は、Fujimura et al.の顔表情データベース⁹⁾より4名（男女各2名）の顔画像を選び、使用した。これら、乳児と成人それぞれ、4名×3表情の写真を呈示刺激とした。

脳機能計測はブロックデザインで行った。課題（タスク）はFig. 1に示す通り、レスト課題（黒地に白字の十字の固視点を提示）を20秒間、ターゲット課題（顔表情刺激）を12秒間、交互に3回ずつ呈示した。脳機能計測は、一人の対象者につき、乳児顔、成人顔、それぞれ顔表情3種類、計6回の計測を行った。対象者ごとに刺激画像呈示の順番のカウンターバランスをとった。脳機能計測の所要時間は一人当たり約30分であった。（なお、本稿の紙媒体はモノクロ印刷であるため色の違いが判別しにくいだが、電子版はカラーとなっているため、本稿の図は電子版を参照されたい。Fig. 1以降の図も同様である。）

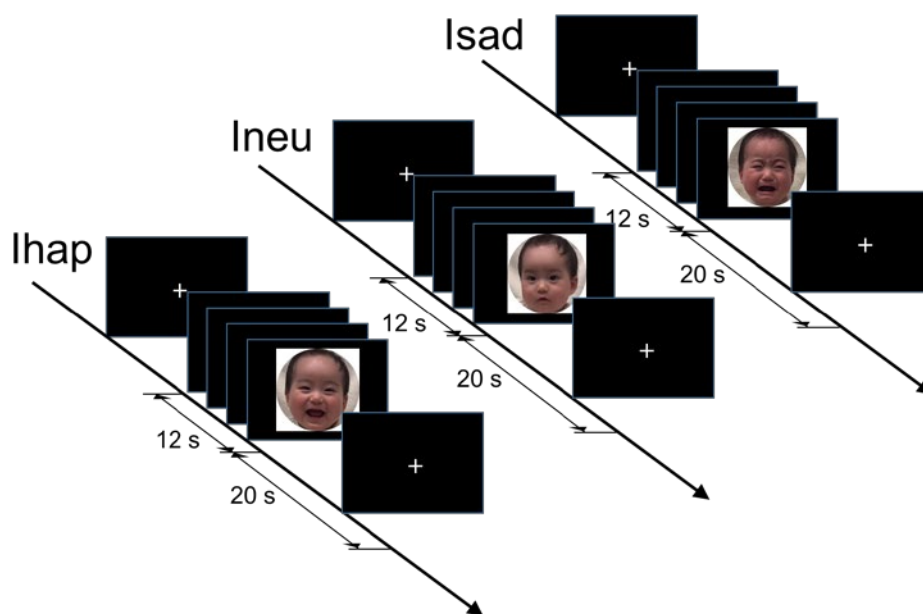


Fig. 1 顔表情呈示課題と計測デザイン

注) 図は幼児の3種類の表情呈示課題 (lhap, lneu, lsad) の例を示している。レスト課題（黒字に十字の固視点を提示）を20秒間、ターゲット課題（顔表情刺激）を12秒間とし、交互に3回呈示するブロック・デザインとした。

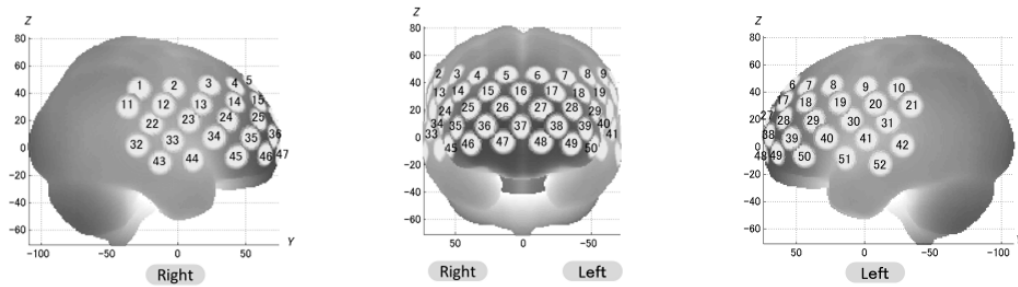


Fig. 2 NIRS 計測位置 (52チャンネルの配置)
 注) 左の図が右半球、中央の図が正面、右の図が左半球

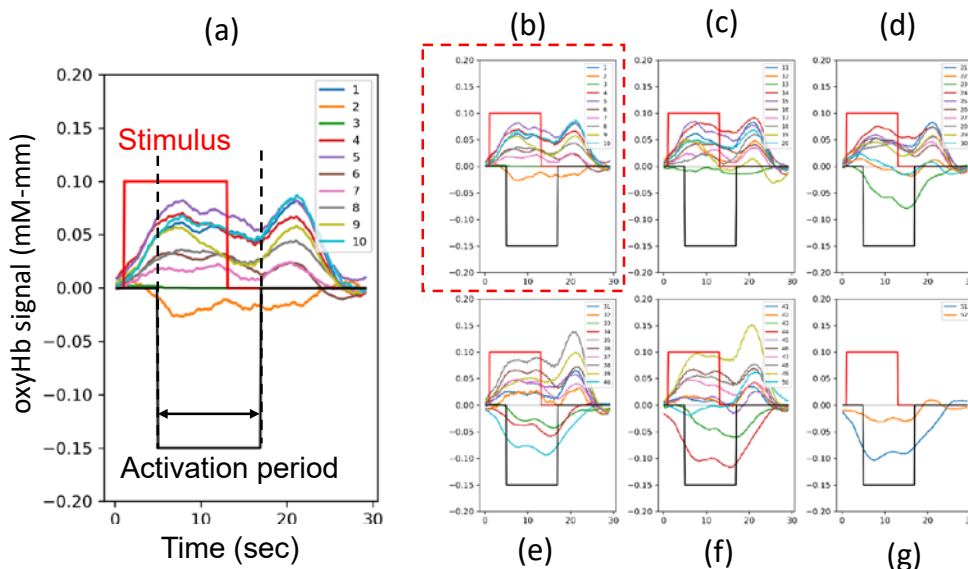


Fig. 3 fNIRS 装置によって計測した酸素化ヘモグロビン濃度の時間変化の例

注) 各グラフ中の赤の実線の枠は顔表情の提示時間窓を表す。一方、黒の破線及び実践の枠は解析対象とした時間窓を表す。(a) は (b) の拡大図。(b) ~ (g) はそれぞれ、fNIRS の測定チャンネルごとの測定値を示している。グラフ内右上の数値はチャンネル番号。

2. 4 fNIRS 計測

脳機能計測は、静寂な実験室において行われた。対象者は背もたれの高い椅子に寄りかかる安定した座位姿勢で、1.5m 前方にある刺激呈示用モニター(画面サイズ21インチ)を注視するよう指示された。計測は、多チャンネル fNIRS 装置 (ETG-4000、日立メディコ) を使用し、52チャンネルホルダーを前頭部から両側頭部にかけて装着した。52のチャンネルの配置を Fig. 2 に示す。最下列のプロープを国際10-20法の Fp1-Fp2 ラインとなるよう配置した。本研究における大脳皮質の解剖学的な部位とチャンネル位置の対応は先行研究の¹⁰⁾ のバーチャルレジストレーションの結果に準じた。

我々の用いた fNIRS 装置では、2種類の短波長 (695nm, 830nm) の近赤外光により、脳表層の酸素化ヘモグロビン濃度変化 (oxyHb)、脱酸素化ヘモグロビン濃度変化 (deoxyHb)、総ヘモグロビン濃度変化 (totalHb) を計測できる。脳の局所的な神経活動が増加すると血管が拡張して動脈の血流量が増加し、oxyHb 濃度が増大する。そのため、oxyHb 濃度変化が最も敏感に神経活動を反映すると考えられることから、本研究では oxyHb 濃度変化を脳機能計測の指標とし、後述の方法で酸素化ヘモグロビン濃度変化量の加算平均処理を行った。

2.5 脳機能計測のデータ解析

各課題に対する個人の脳活性化を評価するために、各課題条件の oxyHb 濃度の変化を fNIRS 計測装置により計測した。刺激呈示の12秒間のうち、ニューロン活動に対する血行動態反応の遅れを考慮して、課題開始後5秒後から課題終了5秒後までの12秒の区間を「時間窓」として解析対象とした。Fig. 3に示すように、この時間窓には、課題ブロック中の oxyHb 濃度変化量増加のピークが含まれていた。解析対象の時間窓における12秒間の oxyHb 濃度変化量の積分値 (Δ oxyHb) を算出した。

その後の統計分析では、課題条件間のチャンネルごとの Δ oxyHb の違いを評価した。チャンネルごとの賦活の例を Fig. 4に示す。そして、性格と脳の賦活および賦活部位との関係を分析するために、性格検査スコアとチャンネルごとの Δ oxyHb の間の相関係数を算出した。さらに、脳部位ごとの機能と性格との相関を明らかにするために、52チャンネルを先行研究に準じて脳解剖学上の部位に分類し^{10,11}、それぞれの脳部位ごとの Δ oxyHb の平均と性格検査結果との相関も求めた。

2.6 ディープラーニングによるデータ解析

2.6.1 ネットワーク構造モデル

本研究で用いたネットワーク構造モデルは「5層ニューラルネットワーク (5-layer neural network, 「5-LNN」と略す)」、「VGG19」、「ResNet」の3種類である。5-LNNは5層の中間層を持つ一般的なネットワークであり、畳み込みニューラルネットワーク (CNN) の手法を用いていない。この後述べる CNN の手法を用いた残り2種類のネットワーク構造との比較のために採用した。VGG19は、画像認識のために開発されたネットワークで¹²、ネットワークを発表した研究グループ名 (Visual Geometry Group, Department of Engineering Science, University of Oxford) が名前の由来である。畳み込み層と全結合層を連結しシンプルに層を増やしたネットワーク構造で、使い勝手がよく、画像認識に幅広く用いられるようになっている。層数については、11、13、16、19層などいくつかのバリエーションがあるが、本研究では19層のネットワーク (VGG19) を採用した。後述するように、VGG19はオックスフォード大学の K. Simonyan らによって提唱された16層の畳み込み層と5層の Max Pooling 層、そして、3層の全結合層からなるネットワーク構造である¹²。

Residual Network (ResNet) は、残差を用いたネットワークモデルである。この手法に関する論文は2016年に発表され¹³、コンピューター科学分野において最近の引用回数が上位にランクされている。この論文は、深層学習 (ディープラーニング) に使われる畳み込みニューラルネットワークにおいて、「ResNet (Residual Network)」と呼ぶ新しいモデルを考案し、画像分類の性能の飛躍的な向上を実現した。画像分類問題において、ネットワークの層の深さは重要であり、深いほど精度向上すると考えられているが、より深いネットワークを学習させようとする、精度が劣化する。ResNet では、このような深いネットワークにおいて発生する精度劣化の問題を解決した。後述するように、いくつかの積み重ねられた層 (ブロック) が、直接最適な写像 (変換) になるよう学習するのではなく、残差の写像が最適になるよう学習する構造になっている。

2.6.2 学習方法

合計データ数は50 (男性:25、女性:25)、説明変数は6種類の顔表情に対する52チャンネルの脳活性化値であり、6種類の顔表情刺激からの取得データを1つの説明変数とした。5-LNN においては312 (52x6) 次元ベクトル、VGG19と ResNet においては、5x11の画像にするために3つの0セル (ダミーセル) を加え、5x11x6の画像を説明変数とした。また、目的変数は NEO 5因子 (N, E, O, A, C) のスコア (心理指標値) とした。

このデータを訓練用データ：40（男性：20、女性：20）と評価用データ：10（男性：5、女性：5）にランダムに分割し、訓練用データで学習を実施し、評価用データでモデルの予測性能を評価した。なお、学習時のロス関数は、本タスクが回帰であるため、以下で与えられる平均二乗誤差（Mean Squared Error: MSE）を用いた。

$$\text{MSE} = \frac{1}{40} \sum_i \sum_{j \in \text{N,E,O,A,C}} \frac{(y_{ij} - t_{ij})^2}{2}$$

ここで、 y_{ij} はモデルに説明変数を入力した際に得られる予測値、 t_{ij} はその説明変数に対する正解値である。なお、一般的に、学習時、入力される説明変数の順番の影響を抑えるためと、学習速度の向上を目的に64や128、256といったデータ数程度でバッチ化を行うが、ここでは、学習データ数の40をバッチサイズとした。そのため、上式において MSE の算出は40の除算を行っており、また、epoch 数と学習回数が一致する。

最適化（オプティマイザー）には Adaptive moment estimation を用い、学習回数は10,000 epochs とし、学習が安定化し、収束を良くするために、学習率は学習開始時に0.001からスタートし、学習終了の10,000 epoch において0となるよう、

$$\text{lr} = 0.001 \times \frac{10,000 - \text{学習回数}}{10,000}$$

の式で割り引いている。

2.6.3 モデル評価方法

本稿では、後述する3つのモデルの性能を比較するために、合計50のデータに対して、訓練用データ（40）と訓練データ（10）の組み合わせを変更して、10,000 epochs の学習を20回実施し、各学習に対して、評価データに対する平均絶対値誤差（Mean Absolute Error: MAE）の記述統計情報（平均値、標準偏差、最小値、25% タイル、50% タイル（中央値）、75% タイル、最大値）を算出した。なお、20種類の訓練用データと評価用データの組み合わせは、乱数シードを固定することで、3つのモデルで統一し、N, E, O, A, C の各項目において、MAE の値が25% タイル $-1.5 \times$ （75% タイル -25% タイル）よりも少ないもの、および、75% タイル $+1.5 \times$ （75% タイル -25% タイル）よりも大きいものは外れ値として除外した。

3. 結果と考察

3.1 性格検査

NEO-PI-R の結果に基づく、対象者の5次元の性格特性のスコアの基本統計量を Table 1 に示す。各スコアは40から160の範囲に分布していた。また各スコアの平均値及び標準偏差は、日本人大学生の標準的標本⁸⁾ から得られたものと近似していた。

3.2 脳賦活のタスク依存性

NIRS 計測した52チャンネルの各チャンネルの活性化値（脳血流量の平均値）のタスク依存性を対象者ごとに調べた。乳児の3種類の顔表情に対する対象者（女性）の活性化値の分布（活性化マップ）の一例を Fig. 4 に示す。図中の番号はチャンネル番号、図の左側が右脳、右側が左脳である。図で色の濃いチャンネル（脳部位）は活性化値が高いところである。図から分かるように、明らかに顔表情によって脳活性の反応が異なることが分かる。各対象者について、Fig. 4 に示すような活性化マップ（賦活イメージ）が、先に述べ

Table 1 NEO-PI-R の結果の基本統計量

Factors	Mean	S.D.	Range
Neuroticism (N)	111.8	23.0	54-165
Extraversion (E)	107.0	18.8	74-150
Openness (O)	113.4	16.0	79-159
Agreeableness (A)	113.2	16.2	76-155
Conscientiousness (C)	98.7	22.5	41-143

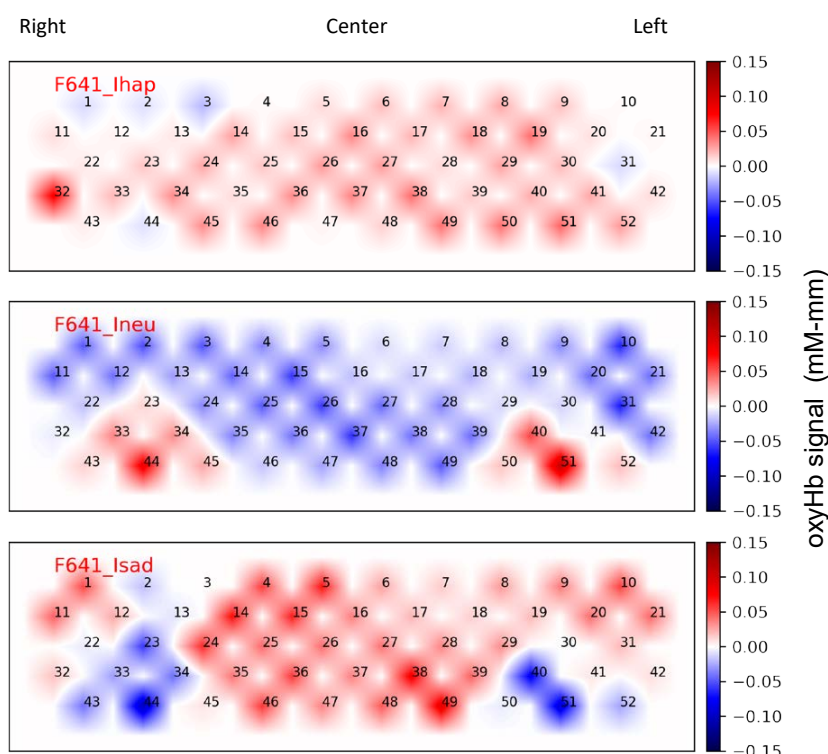


Fig. 4 刺激画像を見ているときの脳の賦活の例
 注) 上から乳児の happy, neutral, sad の表情を見ている時の脳の賦活。図中の数値はチャンネル番号。
 各チャンネルの色及び濃淡は右のインディケーターに対応した oxyHb 濃度。

たディープラーニングにおける説明変数となる。

Fig. 5 には対象者50名すべての平均をとった活性化マップを示す。左側が女性の結果、右側が男性の結果、そして中央が両者の平均である。この図からも、顔表情によって特異的に賦活する領域があることが分かる。特に注目すべきは、happy と sad の表情に対する脳賦活の度合いが、曖昧表情に比べて大きくなっていることである。また、happy 表情に対しては側頭部が、sad 表情に対しては中央部がより賦活している。さらに、幼児の sad 表情に対して、女性の方が男性より賦活の度合いが大きくなっていることも注目に値する。対象者は育児未経験の男女であるが、育児経験が無くとも母性に関わる脳部位の活動がこのような活性化の違いをもたらしているとも考えられる。今後、脳賦活のタスク依存性の要因について、脳機能との関連でさらに詳しく追及していく予定である。

3.3 ディープラーニング解析結果

先に説明した、5-LNN、VGG19、ResNet の3種類のネットワーク構造のモデルについて、学習と評価データの組み合わせを変えて20回学習を行って得られた予測値の MAE と、それらから推察される予測結果の傾向について以下に述べる。

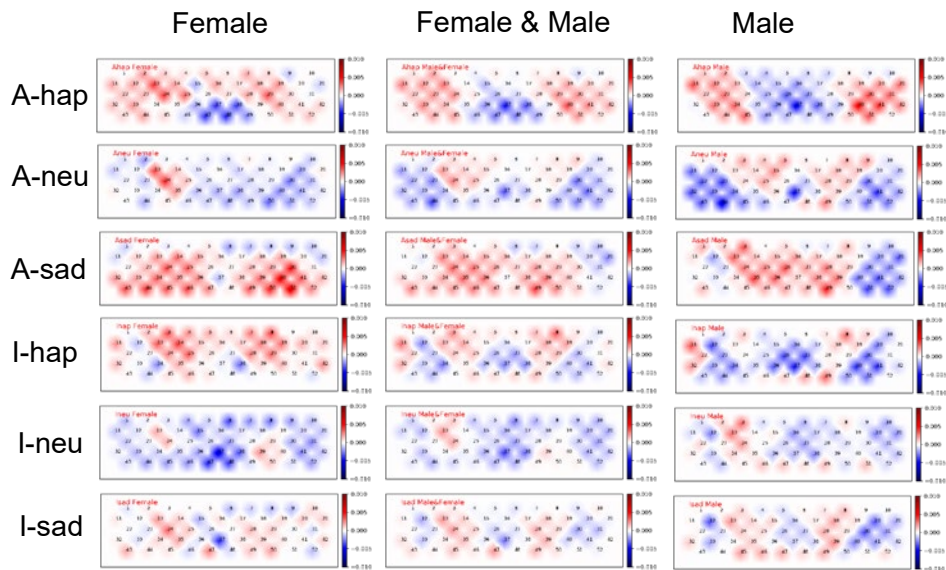


Fig. 5 対象者すべての脳賦活の平均の活性化マップ

注) 各グラフ縦軸は oxyHb 濃度の計測値 (mM-mm)。グラフ内の数値は測定チャンネルの番号。
 A-happy の A は成人、I-happy の I は乳児、hap (happy), neu (neutral), sad はそれぞれ表情の種類。
 Female は女性対象者のデータ、Male は男性対象者のデータ。

3.3.1 5-LNN (5層ニューラルネットワーク)

このネットワーク構造の模式図は Fig. 6 の通りである。まず、入力には6種類の顔表情刺激に対する脳賦活マップの結果を並べ、312 (52 x 6) 次元のベクトルの入力層 (Input layer) とし、その後、隠れ層 (Hidden layers) として、1024、512、256、128、64の5つの全結合層を重ね、最後に出力層 (Output layer) において、心理指標の5要素を予測値として出力した。なお、活性化関数には、深層学習で多用される Relu 関数を用いた。

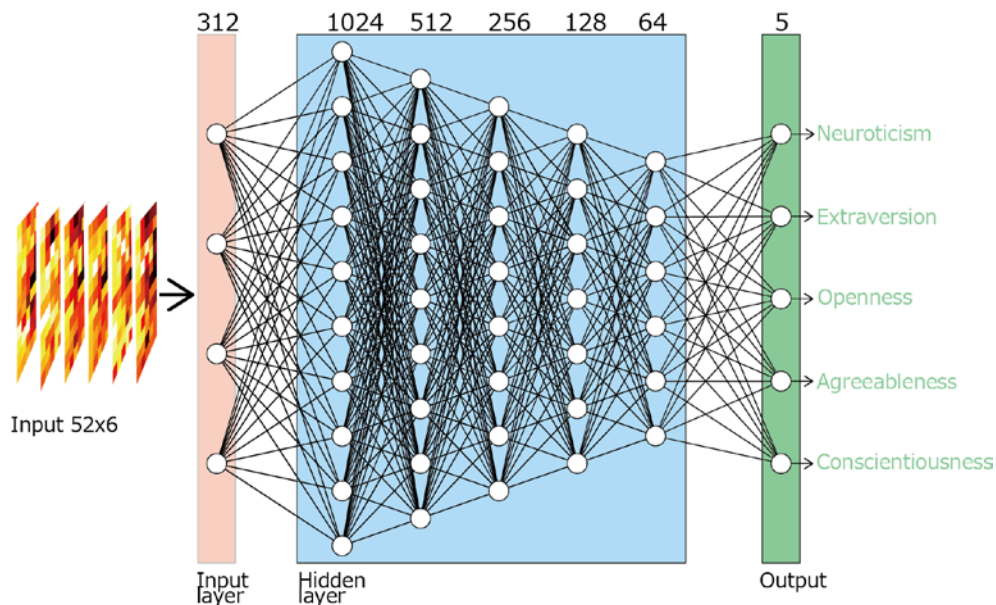


Fig. 6 5層ニューラルネットワーク構造モデルの概念図

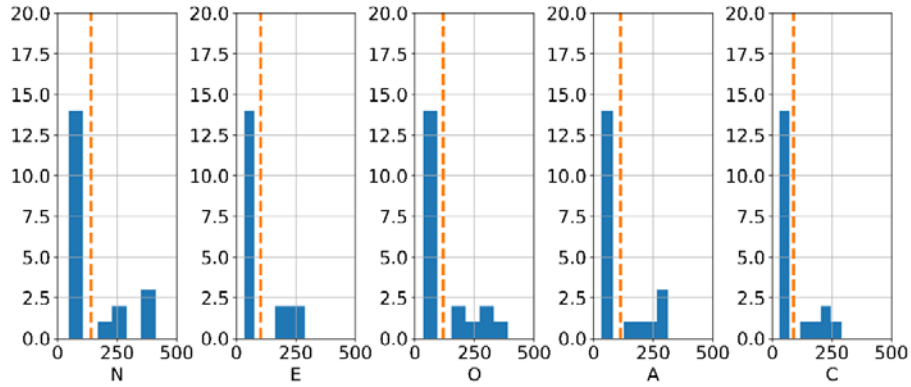


Fig. 7 5層ニューラルネットワーク構造モデルの平均絶対値誤差の度数分布
 注) NEO-PI-Rの5次元の性格特性ごとの平均絶対値誤差 (MAE) の度数。各グラフの横軸は MAE、縦軸は度数を示す。

Table 2 平均絶対値誤差 (5-LNN) の記述統計量

	N	E	O	A	C
mean	127.25	94.28	108.51	102.93	89.78
std	117.58	81	108.14	93.68	82.35
min	46.9	33.9	36.1	34.1	29.8
25%	58.55	45.3	46.85	45.7	39.05
50%	66.2	52.8	56.2	56.7	47.4
75%	151.25	122.3	128.9	119.8	105.7
max	411	289.5	390.3	312.9	289.5

Fig. 7は、5層ニューラルネットワーク構造のモデルにおいて、20種類の訓練データと評価データの組み合わせで、20回の学習を行い、評価データに対して得られたMAEの度数分布（ヒストグラム）である。すなわち、横軸がMAE、縦軸が度数である。また、縦の直線はMAEの平均値である。MAEの平均値が100前後であり、このネットワーク構造では殆ど予測できていない。Table 2に示す平均絶対値誤差の記述統計量の表も、標準偏差が100前後であり、単純な多層ネットワークでは予測できないことを示している。

3.3.2 VGG19

Fig. 8にオックスフォード大学のK. Shimonyanら¹²⁾によって提唱された16層の畳み込み層と5層のMax Pooling層、そして、3層の全結合層からなるネットワーク構造VGG19を示す（19の由来は、畳み込み層と全結合層の層数の合計に由来する。）。ここで、畳み込み層とは、イメージに対して、フィルタ、または、カーネルと呼ばれる重みパラメータを有した領域をスライドさせながら作用させていき、イメージの特徴を取得するための層を指す。この畳み込み層はイメージを入力にして、イメージを出力とする層であるため、入力は正方形、または、長方形の配列である必要があり、イメージをイメージのまま扱う点の特徴である。しかし、本研究で対象としたfNIRS活性化マップの賦活イメージは、Fig. 4に示すように、[10列, 11列, 10列, 11列, 10列]の配列構造であるため、1行目と3行目、5行目それぞれに0を要素として加え、5行×11列の配列にした。さらに、6種類の表情に対してそれぞれ1枚のイメージを作成し、入力は5行×11列×6枚のイメージとした。本稿では、全てのフィルタサイズを3x3とし、ストライドは1（NIRSのチャネル1個に相当）とした。畳み込み層においては、1フィルタに対して1イメージを出力するので、64フィルタを用いると、64のイメージを出力することになる。Pooling層も畳み込み層と同様にフィルタを有するが、Pooling層のフィルタは重みパラメータを有さず、領域内の最大値（Max

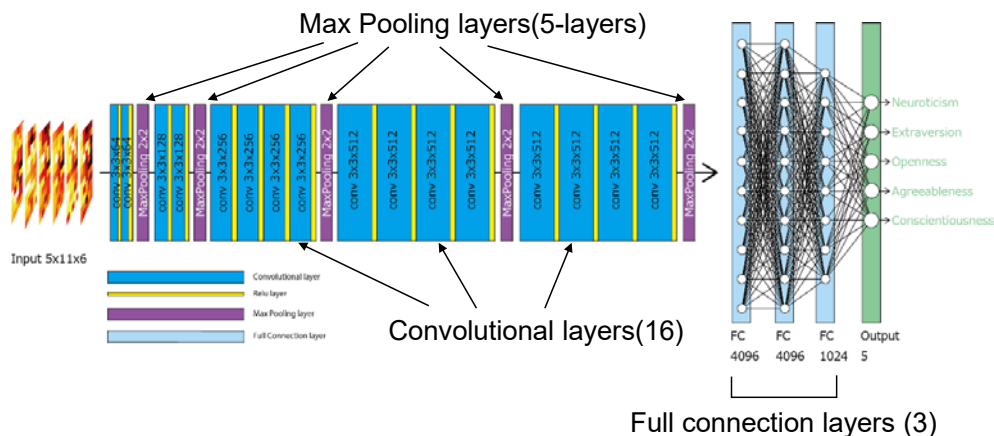


Fig. 8 VGG19構造モデルの概念図

Pooling) や、領域内の平均 (Average Pooling) を 1 要素にマッピングしてイメージを出力する。本稿では Pooling 層においては、 2×2 のフィルタをストライド 1 で用いた。

VGG19では、最後の Max Pooling 層の後には、入力イメージの特徴を反映していると期待される512のイメージが出力され、その512のイメージを通常のニューラルネットワークの入力イメージと同様に扱って、出力を得る。したがって、最後の3層の全結合層部分は一般的なニューラルネットワーク構造と同じであるため、VGG19のようなネットワークは畳み込みニューラルネットワーク (Convolutional neural network, CNN) と呼ばれることがある。このモデル構造の長所は、フィルタを用いることで、イメージ内の特徴のズレに対して堅牢である点である。

5層ニューラルネットワーク構造のモデルの場合と同様に、VGG19モデルの場合も20種類の訓練データと評価データの組み合わせで、20回の学習を行い、評価データに対して MAE を求めた。その結果を Fig. 9 のヒストグラムと、Table 3 の平均絶対値誤差の記述統計量の表に示す。ヒストグラムの図中の縦線は MAE の平均値である。MAE の平均値は、5-LNN に比べて明らかに小さく、おおよそ5分の1程度まで小さくなっている。VGG19が可成りの精度で心理検査指標を予測していることがわかる。さらに、Table 1 に示した性格指標値の標準誤差と比べるとそれぞれの性格因子の MAE は同程度になっている。これらの結果は、ディープラーニングによる予測がかなり精度よく行われたことを示唆している。N と C の性格検査指標値の標準偏差が他の因子の標準偏差より大きくなっているが、この傾向が MAE の値において再現されていることは注目すべき点である。今回の研究において、6表情に対する52チャンネルの脳賦活量を入力データとして用いたことがこのような精度の良い予測につながったと言える。逆に、異なる顔表情に対する脳賦活が性格を反映していると言える。すなわち、性格により前頭前野の脳賦活の仕方が異なることを示唆している。

3.3.3 ResNet

ResNet はマイクロソフト社の K. He らによって提唱されたモデルで、VGG19などの畳み込みニューラルネットワークモデルと同様に、畳み込み層、Pooling 層、全結合層からなるが、残差ブロック (Residual Block) と呼ばれるブロックを有するのが大きな特徴である。Fig. 10に本研究で用いた ResNet のネットワーク構造を示す。Residual Block の概要は図中の赤線枠内に示した通りで、畳み込み層によってイメージの特徴を抽出した後に、元のイメージを加える仕組みになっている。この様に、抽出された特徴にもとのイメージを加えることで、層を多数積んだ際に問題となる逆伝搬時のロスの消失を防ぐことが可能となる。本研究の ResNet モデルは、フィルタ数を変えながら、フィルタ数64、128、256の Residual Block を、

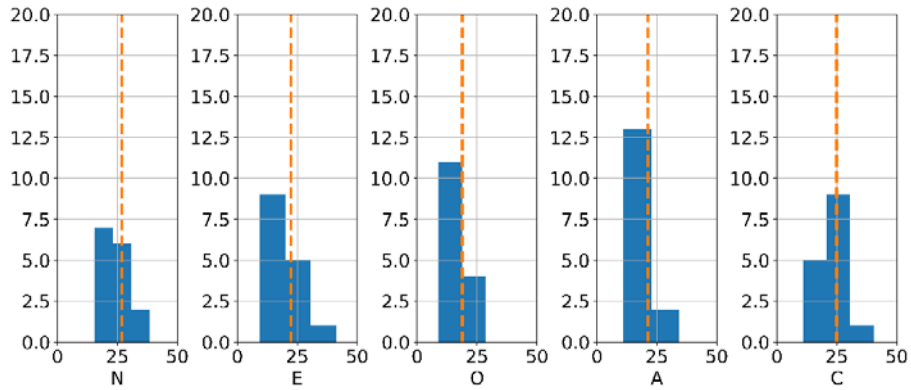


Fig. 9 VGG19構造モデルの平均絶対値誤差の度数分布

注) NEO-PI-Rの5次元の性格特性ごとの平均絶対値誤差 (MAE) の度数。各グラフの横軸は MAE、縦軸は度数を示す。

Table 3 平均絶対値誤差 (VGG19) の記述統計量

	N	E	O	A	C
mean	25.64	18.5	18.19	19.6	24.32
std	5.87	3.17	4.69	5.04	4.87
min	15.4	12.9	10.5	13.7	17
25%	22.28	16.47	15.65	15.07	19.77
50%	25.1	19.6	16.55	19	25.05
75%	27.65	20.85	21.1	23.42	27.23
max	37.2	23.2	29.6	31.6	34.5

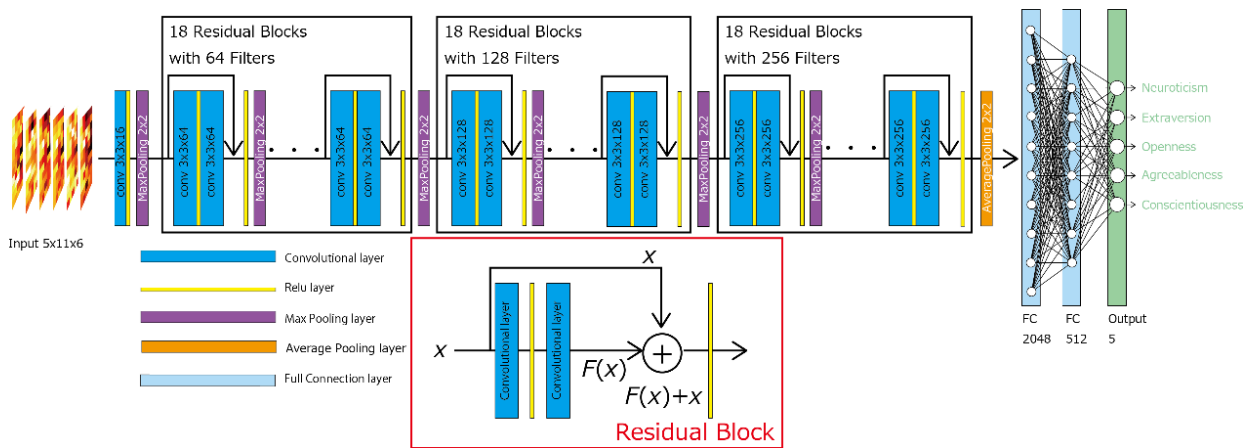


Fig. 10 ResNet 構造モデルの概念図

それぞれ18 Blocks ずつ積んでいる。1 Block は2層の畳み込み層を有するので、108層の畳み込み層と2層の全結合層で合計110層と、非常に深いネットワーク構造を有している。

上述の2種類のネットワーク構造モデルの場合と同様に、20種類の訓練データと評価データの組み合わせで、20回の学習を行い、評価データに対して求めたMAEをFig. 11とTable 4に示す。結果は、VGG19とほとんど同じであり、特に著しく予測の精度が向上していない。この結果は、モデルそのものが不十分というより、Table 1に示したように、そもそもの性格検査指標のスコアにばらつきがあること原因していると考えられる。しかし、MAEの値を注意深く見ると、Table 1で示した各因子の標準偏差

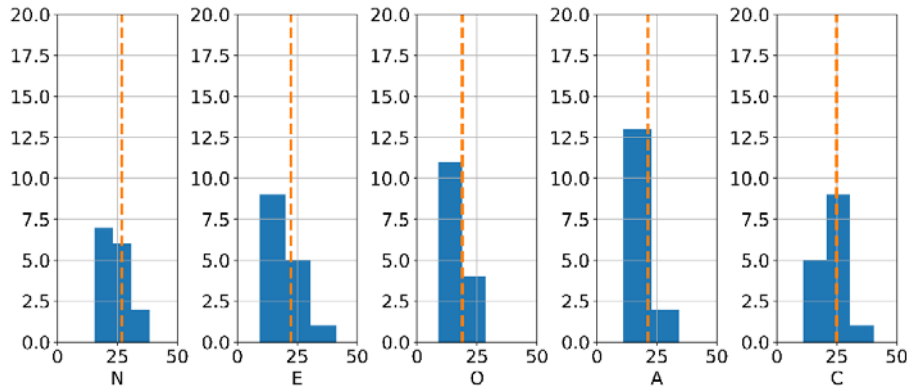


Fig. 11 ResNet 構造モデルの平均絶対値誤差の度数分布

注) NEO-PI-R の 5 次元の性格特性ごとの平均絶対値誤差 (MAE) の度数。各グラフの横軸は MAE、縦軸は度数を示す。

Table 4 平均絶対値誤差 (ResNet) の記述統計量

	N	E	O	A	C
mean	24.53	19.01	15.57	17.25	21.87
std	6.11	5.89	5.06	6.08	6.33
min	15.4	9.3	8.9	11	11
25%	20	15.55	12.45	14.25	17.65
50%	23.5	17.8	14.1	16	20.9
75%	28.75	21.15	17.7	17.8	24.2
max	37.4	34.3	27.6	33.9	36.6

の値により近づいていることがわかる。ResNet は画像分類の性能の飛躍的な向上を実現したモデルであり、本研究で得られた結果は、このネットワーク構造モデルの性能の高さを示している。

4. 今後の展開

本研究では、顔刺激に対する脳反応を fNIRS で測定した結果と性格特性指標値との相関をディープラーニング (深層学習) の手法を用いて明らかにした。ディープラーニングで用いた入力データは脳賦活パターンであった。今後、脳賦活のパターンから脳賦活のネットワーク構造を導き出し、そのネットワークのパターンを入力データ (説明変数) として学習させ、性格特性指標を予測する試みを行いたいと考えている。以下に、現在進めている研究を簡単に紹介する。

脳血流量は顔表情を提示すると、Fig. 3 に示すような時間変化 (タイムコース) を示す。それぞれのチャンネルの時間変化は、似ているものと似ていないものがあり、似ているものは同期していると見なし、互いに連動していると考え。そこで、52個のチャンネル同士の相関を調べる。それらの中から相関係数の大きいものを取り出すと、各チャンネルがどのように連動しているか (繋がっているか) がわかる。こうして相関の強いものだけを取り出して、繋がっているところは 1、それ以外は 0 とした行列で表す。この行列は隣接行列 (adjacency matrix) と呼ばれる。Fig. 12 はある対象者に対して 6 種類の顔表情を呈示した時の脳賦活から求めた 52 行 52 列の隣接行列である。白い点が 1、黒い点が 0 である。隣接行列に基づいてチャンネル同士の連結をネットワーク構造として表したものが Fig. 13 である。顔表情によって隣接行列のパターンやネットワーク構造が大きく異なっていることがわかる。そこで、隣接行列のパターンを一つの画像と考え、この画像を入力データとしてディープラーニングを行えば、性格との相関を明らかにできるはずである。さらに、隣接行列から容易にネットワークの特徴量であるクラスタリング係数や平均パ

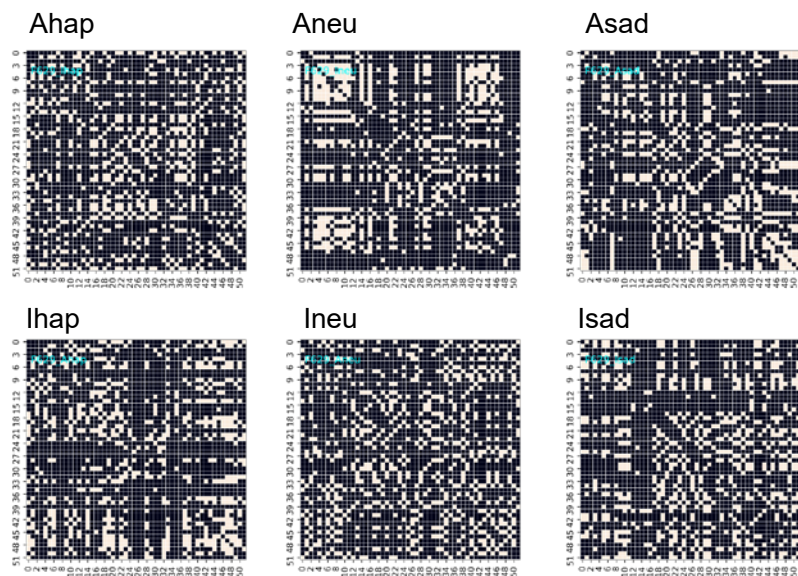


Fig. 12 6種類の顔表情呈示課題に対する脳賦活の状態を表す隣接行列 (adjacency matrix)
 注) 52行 x52列の正方行列で左上から右下に向けた対角線に対して左右対称になっている。

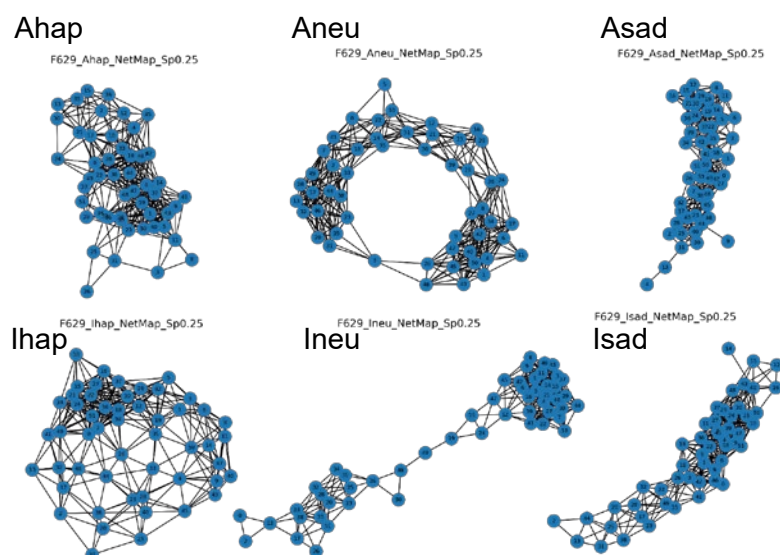


Fig. 13 Fig. 12の隣接行列を可視化したネットワーク構造
 注) 測定した52チャンネルについて、チャンネル同士の連結をネットワーク構造として表したもの。
 図中の数値はチャンネル番号。Ahap, lhap はそれぞれ、成人のhappy顔、乳児のhappy顔を示す。以下、同様。

ス長などを求めることができる。今後さらに詳細な分析を行い、これらの特徴量と性格特性の相関を明らかにしたい。

5. 結言

本研究では、顔刺激に対する脳活動をfNIRSで測定した結果と性格特性との相関をディープラーニング(深層学習)の手法を用いて明らかにした。前頭前野の賦活のパターンを入力信号、性格特性指標値を出力信号として、畳み込みニューラルネットワーク(CNN)の手法で解析した結果、計測された脳賦活パターンから性格特性指標値が20%以内の誤差で予測できることが分かった。今後、顔刺激以外の視覚刺激に対しても同様の実験を行い、脳賦活という生理的反応と性格特性という心理的要因との相関をより詳

細に明らかにしていく予定である。本研究によって得られる知見は、青年期成人の顔刺激の種類に対する反応傾向と性格特性との関連を検討するための基礎資料となり、青年期成人への親性育成教育や、乳幼児の親の産後うつ予防あるいは虐待防止等の育児支援に寄与するものと期待される。

謝辞

本研究は JSPS 科研費（18K03072 研究代表者 庭野賀津子）の助成を受けたものです。

引用文献

- 1) Niwano, K. and Tanabe, M. "Regional brain responses in young adults during producing infant-directed speech: A NIRS study." *Proceeding of The 23rd Biennial Meeting of the International Society for the Study of Behavioural Development*, Shanghai, 2014
- 2) Niwano, K., Ito, A., Tanabe, M., Sato, Y. and Fujii, T. "An fMRI study of the relationship between neuroticism and neural responses to infant emotional faces. *Proceeding of 15th European Congress of Psychology*, Amsterdam, 2017
- 3) Ito, A., Niwano, K., Tanabe, M., Sato, Y. and Fujii, T. "Activity changes in the left superior temporal sulcus reflect the effects of childcare training on young female students' perceptions of infants' negative facial expressions." *Neuroscience Research* 131, 36-44, 2018
- 4) Hiwa, S., Hanawa, K., Tamura, R., Hachisuka, K. and Hiroyasu, T., "Analyzing Brain Functions by Subject Classification of Functional Near-Infrared Spectroscopy Data Using Convolutional Neural Networks Analysis," *Computational Intelligence and Neuroscience*, 9 Volume 2016, Article ID 1841945, 9 pages.
- 5) Rojas, R., F., Huang, X. and Ou, K.-L., "A Machine Learning Approach for the Identification of a Biomarker of Human Pain using fNIRS," *Scientific Reports* 9, Article number: 5645, 2019
- 6) 下仲順子, 中里克治, 権藤恭之, 高山緑「日本版 NEO-PI-R の作成とその因子的妥当性の検討」『性格心理学研究』 6 (2), 138-147, 1998
- 7) Costa, P. and McCrae, R. "Normal personality assessment in clinical practice: The NEO personality inventory." *Psychological Assessment* 14, 5-13, 1992
- 8) 下仲順子, 中里克治, 権藤恭之, 高山 緑『日本版 NEO-PI-R NEO-FFI 使用マニュアル改訂増補版』東京心理株式会社, 2011
- 9) Fujimura, T. and Umemura, H. "Development and validation of a facial expression database based on the dimensional and categorical model of emotions." *Cognition and Emotion* 32, 1663-1670, 2018. DOI: 10.1080/02699931.2017.1419936.
- 10) Tsuzuki, D., Jurcak, V., Singh, A.K., Okamoto, M., Watanabe, E. and Dan, I. "Virtual spatial registration of stand-alone fNIRS data to MNI space." *NeuroImage* 34, 1506-1518, 2007
- 11) Rorden, C. and Brett, M. "Stereotaxic display of brain lesions." *Behavioural Neurology* 12, 191-200, 2000
- 12) Simonyan, K. and Zisserman, A. "Very deep convolutional networks for large-scale image recognition." *Conference paper ICLR*, 2015
- 13) He, K., Zhang, X., Ren, S. and Sun, J. "Deep residual learning for image recognition." *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 770-778, 2016. doi: 10.1109/CVPR.2016.90.