

平成 30 年度

修士学位論文

3 次元 fMRI 画像からの脳情報
デコーディングにおける CNN の適用

Application of Convolutional Neural Network to
Brain Decoding from 3-Dimensional
functional Magnetic Resonance Imaging

1215084 佐々木 泰一

指導教員 吉田 真一

2019 年 2 月 28 日

高知工科大学大学院 工学研究科 基盤工学専攻
情報学コース

要 旨

3次元 fMRI 画像からの脳情報 デコーディングにおける CNN の適用

佐々木 泰一

fMRI を用いた脳情報デコーディングはこれまで視覚を中心に成功報告がいくつかされている。脳情報デコーディングを行う際の機械学習アルゴリズムに CNN を用いたものはいくつか報告されているが、被験者が物体を認識しているときの低次脳機能のデコーディングである。これに対し、本研究では、より高次の脳機能も関わりと考えられる感情のデコーディングを対象とする。また、脳活動データは本来 3次元であるが、多くの研究では 2次元のスライス画像の集まりに 2D-CNN を適用している。3次元の CNN も存在するが、脳情報デコーディングにおける 3D-CNN の適用例はまだほとんど報告されていない。これには、3D-CNN が 2D-CNN に比較して多くのパラメータがあり、また計算資源を要することも問題になっていると考える。そこで本研究では、これまで SVM で行われていた感情の脳情報デコーディングを対象として、3D-CNN の適用に加え、新たに提案する特徴選択を行い 2D-CNN を適用し、その性能を明らかにする。

評価実験では、被験者 8名の脳活動データに対して、SVM、従来の 2D-CNN、提案手法を用いた 2D-CNN、3D - CNN を用いて学習し、識別率を比較する。結果として、多くの被験者において、提案手法を用いた 2D-CNN による学習をした場合に最も高い識別率が算出された。また、提案手法を用いた 2D-CNN は 3D-CNN よりも少ない計算資源で学習することができることがわかった。

キーワード 畳み込みニューラルネットワーク, 脳情報デコーディング, 多変量パターン
解析

Abstract

Application of Convolutional Neural Network to Brain Decoding from 3-Dimensional functional Magnetic Resonance Imaging

Taichi SASAKI

Several success studies have been reported on brain decoding using fMRI especially in vision tasks. Several studies using CNN for machine learning algorithms for brain decoding are reported However these are decoding of low-order brain functions compared to emotions such as object recognition. While brain activity data is originally three-dimensional, 2D-CNN is applied to a collection of two-dimensional slice images in many studies. 3D-CNN also exists, but application examples of 3D-CNN in brain information decoding are still hardly reported. For the reason, we consider that 3D-CNN has many parameters compared to 2D-CNN, and requires expensive computing resources are required. Therefore, in this research, we compare the following two methods: applying 2D-CNN to brain decoding of emotion using SVM and applying 2D-CNN to proposed features.

In the evaluation experiment, learning is performed using brain activity data of eight subjects. We compare the accuracy of SVM, conventional 2D-CNN, 2D-CNN using the proposed method, 3D-CNN. As a result, As a result, accuracy of 2D-CNN is highest. 2D-CNN using the proposed features is also able to learn using fewer computational resources than 3D-CNN.

key words Convolutional Neural Network , Brain Decoding , Multi-Variate Pattern

Analysis , functional Magnetic Resonance Imaging

目次

第 1 章	序論	1
第 2 章	関連技術	3
2.1	Functional Magnetic Resonance Imaging(fMRI)	3
2.2	脳情報デコーディング	3
2.3	Multi Variate Pattern Analysis(MVPA)	4
2.3.1	Searchlight 法	4
第 3 章	機械学習手法	6
3.1	Support Vector Machine(SVM)	6
3.1.1	カーネル法	7
3.2	Convolutional Neural Network(CNN)	8
3.2.1	CNN の基本構造	8
3.2.2	畳み込み層	10
3.2.3	プーリング層	11
3.3	3D-CNN	12
第 4 章	実験	16
4.1	実験環境	16
4.2	データセット	16
4.3	前処理	16
4.4	学習モデル	17
4.4.1	Searchlight + SVM モデル	17
4.4.2	2D-CNN モデル	18
4.4.3	3D-CNN モデル	19

目次

4.5	交差検証法	19
第 5 章	提案手法・実験	26
5.1	被験者 A	26
5.2	被験者 B	26
5.3	被験者 C	27
5.4	被験者 D	27
5.5	被験者 E	28
5.6	被験者 F	29
5.7	被験者 G	29
5.8	被験者 H	30
第 6 章	考察	40
6.1	提案手法と従来の 2D-CNN および SVM との比較	40
6.2	提案手法と 3D-CNN の比較	40
第 7 章	結論	43
	謝辞	44
	参考文献	45

目次

2.1	Searchlight 法の手順	5
3.1	最適識別超平面とサポートベクトル	13
3.2	CNN の基本構造の概略図	13
3.3	畳み込み層でのフィルタ 1 つに関する計算内容	14
3.4	Max Pooling	14
3.5	3D-CNN の基本構造の概略図	15
4.1	提案手法 1	21
4.2	提案手法 2	22
4.3	交差検証法	25
5.1	被験者 A の学習結果	32
5.2	被験者 B の学習結果	33
5.3	被験者 C の学習結果	34
5.4	被験者 D の学習結果	35
5.5	被験者 E の学習結果	36
5.6	被験者 F の学習結果	37
5.7	被験者 G の学習結果	38
5.8	被験者 H の学習結果	39

表目次

4.1	実験環境	21
4.2	2D-CNN の学習モデル	22
4.3	3D-CNN の学習モデル 1	23
4.4	3D-CNN の学習モデル 2	23
4.5	3D-CNN の学習モデル 3	24
4.6	3D-CNN の学習モデル 4	24
6.1	実験結果	42
6.2	実験結果	42

第 1 章

序論

近年，社会の高齢化に伴い，医療や介護の分野で BCI が期待されている．その中で，脳活動データから人間の状態を解読する脳情報デコーディングの研究が行われている．これまで脳情報デコーディングは，物体認識や自然画像認識などで成功報告がされている [1][2][3]．感情の脳情報デコーディングについても研究がされているが，これには SPM(Statistical Parametric Mapping) によって脳の賦活強度の高い部分をいくつか選択し特徴とする方法や，Searchlight 法 (球などの幾何学的に定義された関心領域を脳内で移動させ，それを特徴とする方法) などを用いた MVPA(Multi-Variate Pattern Analysis) が主に用いられてきた．

CNN を用いた脳情報デコーディングはいくつか報告されているが，これらは物体認識などの感情と比較して低次脳機能のデコーディングである．また，脳活動データは本来 3 次元であるが，多くの研究では 2 次元のスライス画像の集まりに 2D-CNN を適用している．3 次元の CNN も存在するが，脳情報デコーディングにおける 3D-CNN の適用例はまだほとんど報告されていない．これには，3D-CNN が 2D-CNN に比較して多くのパラメータがあり，また計算資源を要することも問題になっていると考える．そこで本研究では，これまで SVM で行われていた感情の脳情報デコーディングを対象として，3D-CNN の適用に加え，新たに提案する特徴選択を行い 2D-CNN を適用し，その性能を明らかにする．

本研究では，快・不快感情を誘起した状態の脳画像データを対象に，いくつかの機械学習アルゴリズムを適用した結果，提案手法が 90.1%と最も高い識別率となり，人間の感情の脳情報デコーディングにおいて提案手法が有効であることを示す．

本論文の構成は，第 2 章では，脳情報デコーディングの概要と従来の解析手法について記

す．第 3 章では，本研究で用いる機械学習手法について記す．第 4 章では，本研究におけるデータセットの説明や，解析手法について記す．第 5 章では，実験を行った結果を記す．第 6 章では，結果を従来手法などと比較した考察を述べる．第 7 章では，本研究の結論を述べる．

第 2 章

関連技術

2.1 Functional Magnetic Resonance Imaging(fMRI)

Functional Magnetic Resonance Imaging(fMRI) は、人の脳活動によって引き起こされる血流の変化を、MR 装置 (核磁気共鳴装置) を用いて、可視化する方法である。これにより、脳の機能の画像化を行う。MR 装置は、強力な磁石の中で原子核の共鳴周波数電磁波を用い、体内の水素原子の原子核の分布を測定することにより、体内の組織の形状や動きを測定する。

2.2 脳情報デコーディング

脳情報デコーディングとは、見ているものや体の動きの状態などを、脳によってコード化された情報と見なし、それを脳活動信号から解読することを指す。この技術を用い、コンピュータなどを動かす BCI(Brain Computer Interface) の研究が、医療や介護の分野で期待されている。脳情報デコーディングは大きく分けて以下の 3 つのステップから成る。

1. 脳活動信号を計測し、ラベルをつける
2. 脳活動信号からラベルを予測する統計モデル (デコーダ) を構築する。
3. 新たな脳活動信号を与えた時、デコーダが「ラベル」を正確に予測できるかを評価する。

まずステップ 1 で、撮像した脳活動データに、そのときに見ているものや、体を動かしている状態などを「ラベル」として付加する。本研究では、笑っている状態であるか、怒っている状態であるかを「ラベル」として付加する。ステップ 2 では、機械学習を用いて、付加

2.3 Multi Variate Pattern Analysis(MVPA)

したラベルごとに学習させ、デコーダを構築する。ステップ3では、ステップ2で学習していない脳活動信号を与えた場合に、構築したデコーダで「ラベル」を正確に予測できるかを評価する。

本研究では、まず脳活動データに、快感情を誘起した状態であるか、不快感情を誘起した状態であるかを「ラベル」として付加し、学習を行いデコーダを構築する。そして、構築したデコーダを用いて予測を行う。

2.3 Multi Variate Pattern Analysis(MVPA)

fMRI 研究では、一般的に MVPA が使用される。これは、ある大きな脳領域における賦活の全体的な増減からではなく、ボクセル全体の様々な賦活のパターンから入力された脳画像が属するカテゴリを予測するものである。MVPA の手順として特徴選択があるが、これにはいくつかの方法が存在する。本項ではその中で、Searchlight 法について説明する。

2.3.1 Searchlight 法

Searchlight 法は例えば球などの幾何学的に定義された関心領域を脳内で移動させる方法である。Searchlight 法の具体的な手順について図 2.1 に示す。まず、ボクセルのサブセットの特定と、 T 間隔ごとの各ボクセルにおける BOLD 振幅の特定が行われる。識別アルゴリズムが脳内でボクセルの特異な組み合わせを特定する際、特徴空間を脳内全てのボクセルと同程度に大きくできるが、多くの特徴を含めることは過学習の問題につながってしまう可能性がある。一般的に、脳全体のボクセルではなく、限られた範囲内のボクセル郡を使用することで、結果がより解釈しやすくなる。 T 間隔は、識別したい刺激事象や行動によって引き起こされる BOLD 信号変化を反映するように選ばれたタイムポイントに設定できる。

2.3 Multi Variate Pattern Analysis(MVPA)

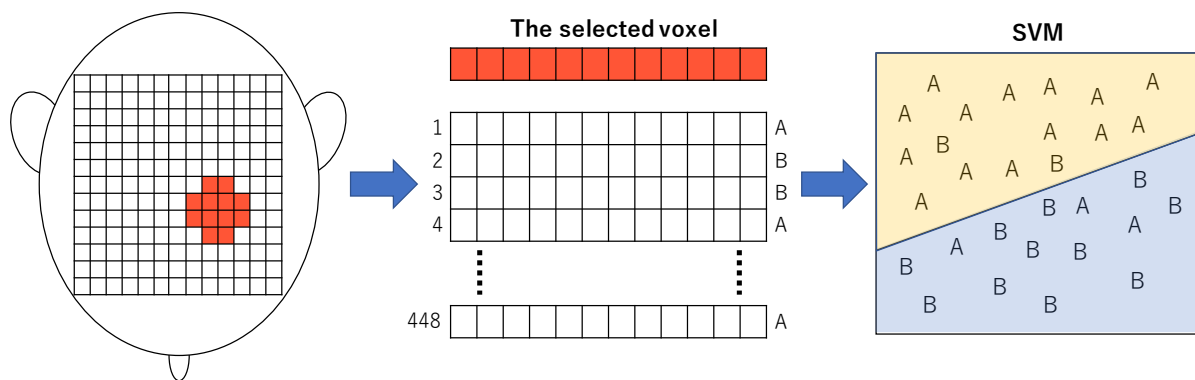


図 2.1 Searchlight 法の手順

第 3 章

機械学習手法

本項では、本論文で用いる Support Vector Machine(SVM) , Convolutional Neural Network(CNN) , 3D-CNN の 3 つの機械学習手法について説明する。

3.1 Support Vector Machine(SVM)

Support Vector Machine(SVM) は、1995 年に Vapnik らにより提案された線形識別関数の構成法であり、最大マージンの考え方を採用している。マージンとは、識別関数と各データ点との距離であり、これを最大化することで、汎化性能を向上させるというものである。線形分離が不可能な問題に対しては、カーネル法を採用することで非線形識別に対応させることや、ソフトマージンによる負のマージンを許すことで、一部のデータに対する誤識別を許容する方法による対応をする。図 3.1 は SVM の特徴空間上における決定関数の定め方を簡潔に表したものである。

SVM の分離超平面を求める最適化問題は式 3.1、制約条件は式 3.2、また、決定関数は式 3.3 で定義される。式 3.1 上の $K(x, x')$ はカーネル関数であり、ベクトル x と x' をより高次元の特徴空間に写像することで、非線形な識別が可能となる。詳細は後述する。

$$\max Q(\alpha) = \sum_{i=1}^M \alpha_i - \frac{1}{2} \sum_{i,j=1}^M \alpha_i \alpha_j y_i y_j K(x, x') \quad (3.1)$$

$$\text{s.t.} \sum_{i=1}^M y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C (i = 1, \dots, M) \quad (3.2)$$

3.1 Support Vector Machine(SVM)

$$D(x) = \sum_{i \in S} \alpha_i y_i K(x, x_i) + b \quad (3.3)$$

最大の利点として、マージンを最大化することで、高い汎化性能を実現できることや、線形分離し辛い高次元データにおいてもサポートベクターを用いた次元削減とカーネルの利用によりデータの分割ができること、解が一意的に収束するため、Neural Networkなどで発生しやすい局所解の問題を回避できることなどが挙げられる。

SVM は脳情報デコーディングにおいてもいくつかの成功報告がされており、人間の視野の方位選択性のデコーディング [4] や、人間の快・不快感情のデコーディング [5] などに用いられている。

3.1.1 カーネル法

SVM は、識別境界が線形の識別超平面となる。識別境界が学習データの線形関数では表せない場合には、非線形特徴写像を用いて非線形特徴空間に写像し、その空間内で線形識別関数を用いる方法がある。これをカーネル法と呼ぶ。カーネル法には、多項式カーネルな

3.2 Convolutional Neural Network(CNN)

どもあるが，本実験では RBF カーネル (Radial Bases Function kernel) を用いる．RBF カーネルは式 (4.1) によって定義される [8] ．

γ はカーネル関数の広がりを制御するパラメータである． γ が小さい場合は識別境界が単純になり， γ が大きい場合は識別境界が複雑になっていく．本実験では，後述する PSO を用いてこの γ の値を定める．

$$K(x, x') = \exp(-\gamma [\|x - x'\|^2]) \quad (3.4)$$

3.2 Convolutional Neural Network(CNN)

Convolutional Neural Network(畳み込みニューラルネットワーク) は，ニューラルネットワークの派生で，画像識別などの，様々な認識タスクにおいて高精度な識別が可能なモデルである．

CNN の源流の一つに福島が提案した Neocognitron がある [9] ．これは，生物の脳の視覚神経系の構造からヒントを得て，単細胞層と複雑細胞層を組み合わせた，2 層の神経回路を基本モジュールとしてモデル化したものである．CNN は，この Neocognitron のモデルを誤差逆伝搬法による勾配計算によって教師あり学習に改良したもので，Yann らによって提案された [10] ．この畳み込み構造は，画像だけでなく，音声などの時系列信号や自然言語テキストの処理に対しても有効であることが知られている．

3.2.1 CNN の基本構造

CNN の基本構造の概略図を 3.2 に示す．CNN の特徴は，畳み込み層およびプーリング層と呼ばれる層を交互に接続した構造を持つことにある．この構造を除けば，CNN は一般的なフィードフォワード型の Neural Network と同じである．

ある層における j 番目のノードは，その直前の層の m このノードからそれぞれ入力 $x_i (i = 1, \dots, m)$ を受け取り，これらの重み付き和にバイアスを加算した，

3.2 Convolutional Neural Network(CNN)

$$u_j = b_j + \sum_{i=1}^m w_{ij}x_i \quad (3.5)$$

を計算する．この u_j を活性化関数に入力したときの応答，

$$y_j = a(u_j) \quad (3.6)$$

がこのノードの出力となり，次の層へと伝搬する．活性化関数 a には，収束性の良さや学習速度の速さから ReLU 活性化関数がよく用いられる [11]．

通常，CNN の出力層付近には隣接層間のノード全てを結合した層（全結合層）を 1 層以上配置する．CNN の出力を与える最終層は通常の NN と同様に設計される．クラス分類を目的とする場合，目的のクラス数と同数のノード n 個を配置し，活性化関数にソフトマックス関数 (softmax function) を用いる．回帰を目的とする場合は，説明したい変数と同数のノードを並べ，変数の値域に合わせ活性化関数を選ぶ．

3.2 Convolutional Neural Network(CNN)

3.2.2 畳み込み層

畳み込み層の基本構造を図 3.3 に示す．図 3.3 は N' 個あるフィルタのうちの一つについての計算である．畳み込み層では入力画像に対してフィルタを畳み込む計算を行う．これは，一般的な画像処理でのフィルタの畳み込み，すなわち小サイズの画像を入力画像に 2 次元的に畳み込んで，画像をぼかしたり，エッジを強調するものと基本的に同じである [12]．

具体的には，入力のサイズ $S \times S$ 画素の各チャンネルごとに $L \times L$ のサイズの 2 次元フィルタを畳み込み，その結果を全 N チャンネルにわたって加算する．この計算の結果は 1 チャンネルの画像の形式 u_{ij} をとる． u_{ij} は，フィルタを w_{ijk} と書くと，

$$u_{ij} = \sum_{k=1}^N \left[\sum_{(p,q) \in P_{ij}} x_{pqk} w_{p-i, q-j, k} \right] \quad (3.7)$$

のように計算される．ただし P_{ij} は，画像中の画素 (i, j) を頂点とするサイズ $L \times L$ 画素の正方領域，

$$P_{ij} = \{(i + i', j + j') \mid i' = 0, \dots, L - 1, j' = 0, \dots, L - 1\} \quad (3.8)$$

である．この場合，出力層のノード数は入力層の $(1/s)^2$ 倍になる．この画素間隔 s をストライドと呼び，入力画像サイズが大きいとき，ネットワークのサイズが大きくなりすぎる

3.2 Convolutional Neural Network(CNN)

のを避けるため，2以上の s を選ぶことがある．ただしその場合，特徴を取りこぼすことに繋がり，一般的に性能は低下する傾向にある．

このように計算された u_{ij} は，その後活性化関数 $a(\cdot)$ を経て，畳み込み層の出力，

$$y_{ij} = a(u_{ij}) \quad (3.9)$$

となる．これにより，一つのフィルタ w_{ijk} につき，入力 x_{ijk} と縦横サイズが同じ $S \times S$ の1チャンネル出力 y_{ij} を得る．同様のフィルタを N' 個用意し，それぞれ独立に上の計算を行えば， N' チャンネル分の $S \times S$ の出力，すなわち $S \times S \times N'$ のサイズの y_{ijk} を得る．これは次の層への入力となる．

3.2.3 プーリング層

プーリングを行うプーリング層は畳み込み層と対で使われ，基本的には畳み込み層の出力がこの層への入力となる．プーリング層の目的は，画像のどの位置でフィルタの応答が強かったかという情報を一部捨て，画像内に現れる特徴の微小な位置変化に対する応答の普遍性を実現することにある [12]．

プーリング層のノード (i, j) は，畳み込み層と同様に，その入力側の層に局所受容野 P_{ij} を持つ．このノード (i, j) の出力は， P_{ij} の内部のノード $(p, q) \in P_{ij}$ の出力 y_{pq} を後述の方法で集約したものである．なお P_{ij} のサイズは，畳み込み層のフィルタサイズとは無関係に設定される．また入力が複数チャンネルある場合，チャンネルごとに以上の処理を独立して行うのが一般的である．つまり，畳み込み層の出力チャンネル数とプーリング層の出力チャンネル数は一致する．

通常，このプーリングの処理はストライド s を2以上に設定する． $s = 2$ とすると出力は入力の縦横半分のサイズとなる．受容野 P_{ij} 内のノードからの入力をまとめて一つの値に集約する方法にはいくつかあるが，本研究では， P_{ij} に属するノードからの入力の最大値を出力とするMax Poolingを用いる．Max Poolingは式3.10で表される．図3.4はストライド s を2にした場合の挙動を示したものである．

3.3 3D-CNN

$$y_{ijk} = \max_{(p,q) \in P_{ij}} x_{pqk} \quad (3.10)$$

3.3 3D-CNN

3D-CNN の基本構造の概略図を 3.5 に示す．3D-CNN は基本的に 2D-CNN と同じように，畳み込み層とプーリング層で特徴を抽出し，全結合層で識別を行う．違いとして特徴抽出に z 軸が加えられ，これにより 3 次元空間での特徴抽出を可能にしている．3D-CNN の既存の適用例として，動画データのように 2 次元データに時間軸を特徴として加える研究がされている [13]．脳情報デコーディングにおいては 3D-CNN を適用した例はあまりないが，3D-CNN は 3 次元空間で畳み込みを行うため，脳画像のスライス間の情報を維持したまま特徴の抽出が行えると考えられる．

3.3 3D-CNN

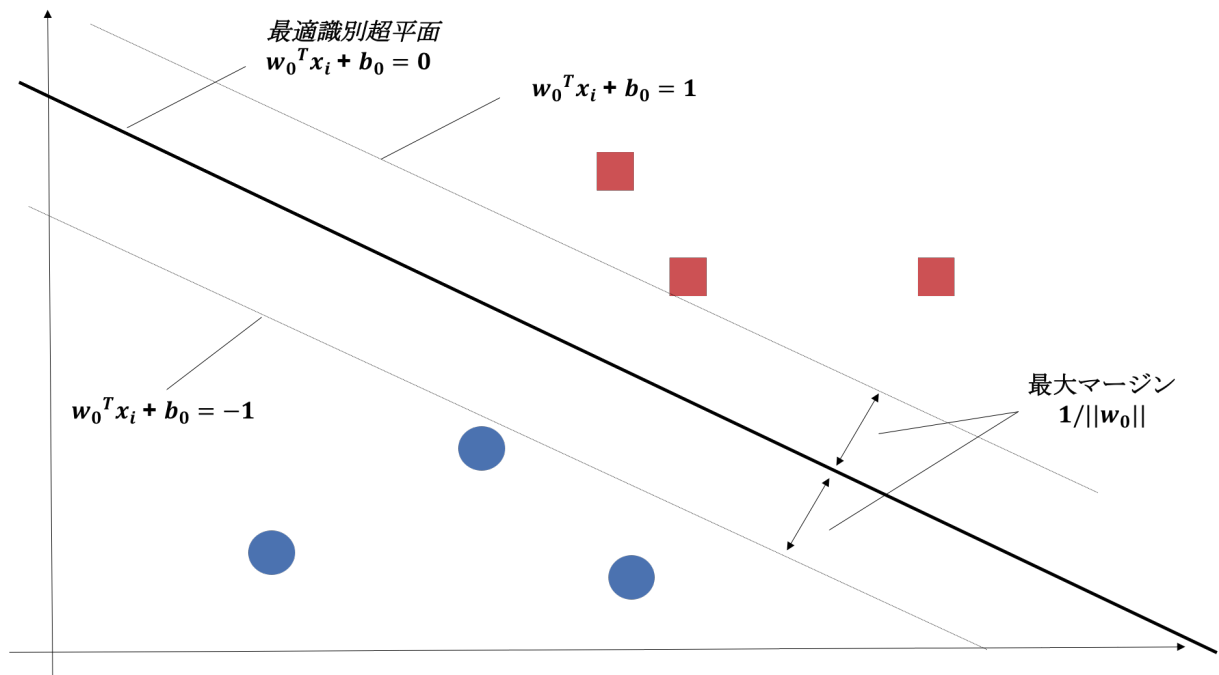


図 3.1 最適識別超平面とサポートベクトル

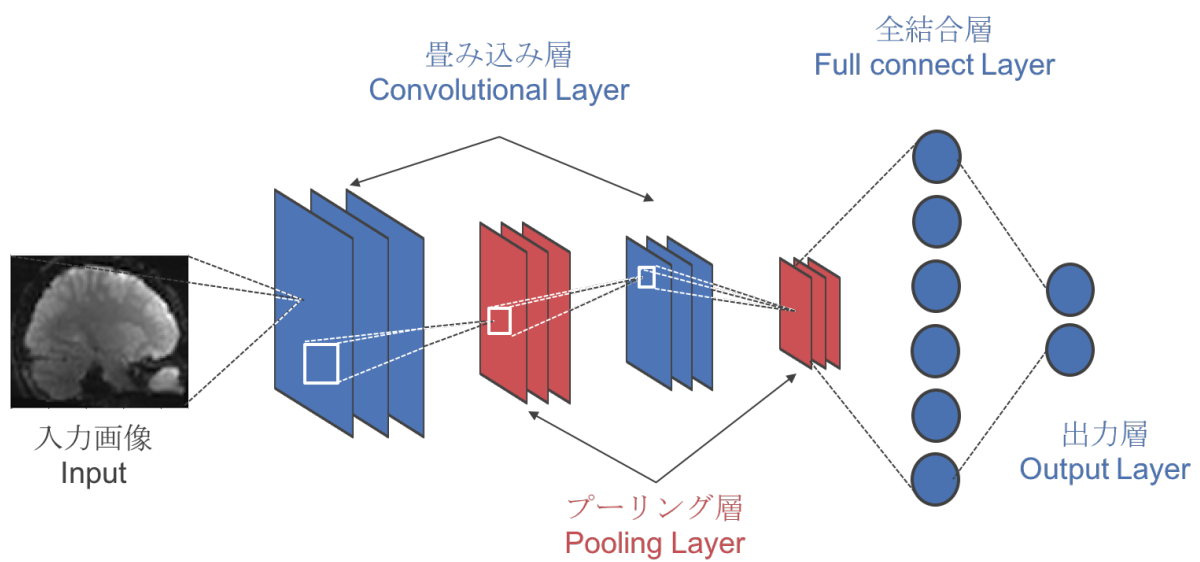


図 3.2 CNN の基本構造の概略図

3.3 3D-CNN

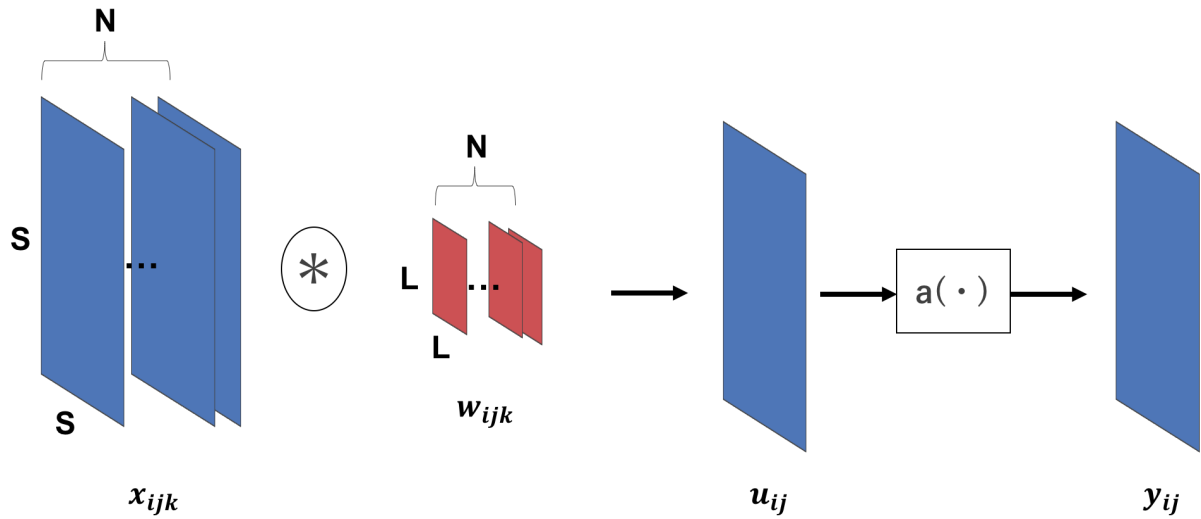


図 3.3 畳み込み層でのフィルタ 1 つに関する計算内容

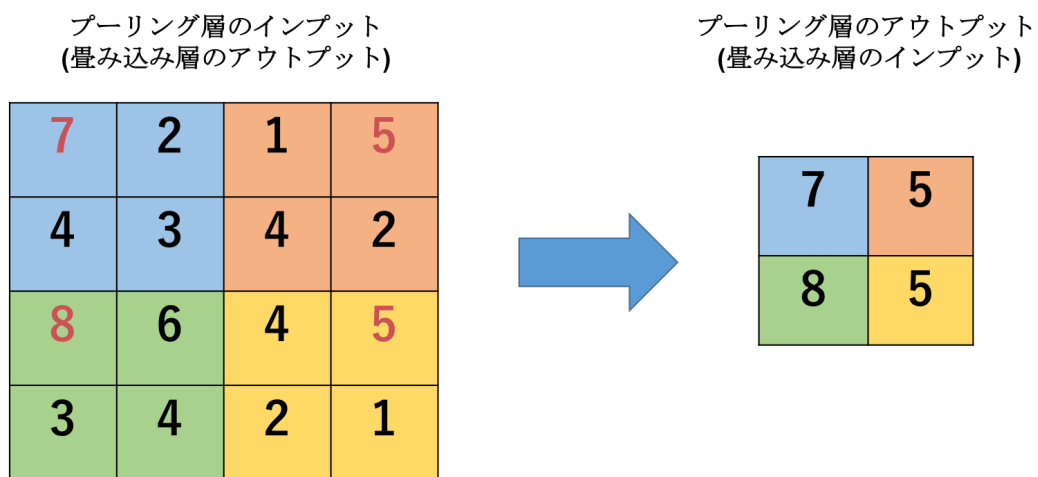


図 3.4 Max Pooling

3.3 3D-CNN

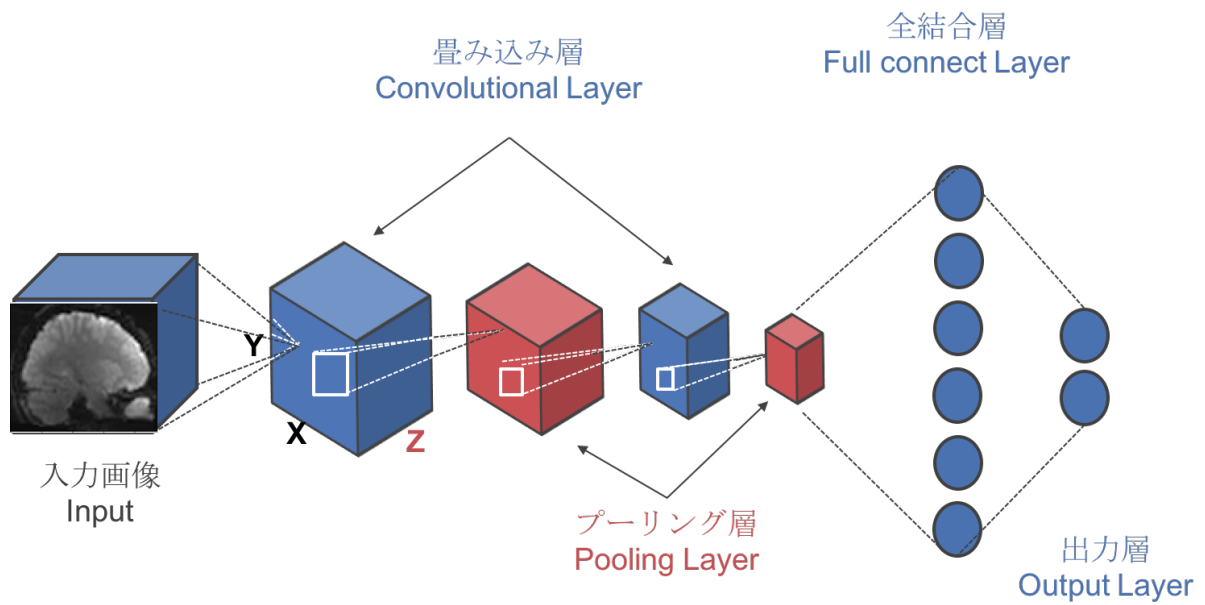


図 3.5 3D-CNN の基本構造の概略図

第 4 章

提案手法・実験手順

4.1 実験環境

本研究で用いる学習モデルは，以下の構成の計算機上にシステムを構築している (表 4.1) .

4.2 データセット

データセットにはあらかじめ感情指標が評定されたカラー画像セットである OASIS を被験者 8 名に提示し，その時の脳活動を fMRI 装置で計測した脳活動データである．OASIS は Benedek Kurdi らにより 2016 年に発表された感情誘起画像である．データセットは，動物，物体，人間，風景の 4 つのカテゴリーから構成されており，カテゴリー別に 200 枚程度のデータを含んでいる．

脳活動データ 362 枚のうち，レスト部分を除いた 144 枚をそれぞれ使用する．144 枚の内訳は，快感情を誘起する画像を見せた際の脳活動データ 72 枚と，不快感情を誘起する画像を見せた際の脳活動データ 72 枚である．

4.3 前処理

脳活動データを解析するために SPM12 を用いて，前処理を施す．

fMRI の実験はたとえ短い時間であっても，体動や心拍などによって被験者の頭も動いてしまうことが多い，また，ボタンを押すような運動を伴う場合もある．これによって各被験者ごとに個人差やノイズが生じてしまう．機械学習においてノイズは誤識別の原因となりう

4.4 学習モデル

るため、できる限り軽減する必要がある。そのため、本研究では脳解析で一般的に用いられている、Realignment 処理、Normalisation 処理の 2 つの処理を行う。Realignment では、2 枚目以降の脳活動データを、1 枚目の脳活動データに合わせることで、頭部の動きに伴う脳活動データの変動を補正する。Normalisation では、Realignment で処理した脳活動データを MNI(Montreal Neurological Institute) 標準脳に合わせる。また一般的に、以上の処理の後に Smoothing 処理を行うことが多いが、この処理は、ボクセルごとの値の違いを平滑化してしまうため、機械学習における特徴を消してしまう可能性がある。そのため、本研究においてはこの処理を行わない。

4.4 学習モデル

本実験においては、下記 3 種類の機械学習法を用いて学習モデルを構築する。

4.4.1 Searchlight + SVM モデル

Searchlight の実装には PyMVPA を用い、Searchlight の半径は 2 とする。Searchlight で得た特徴を SVM にて学習する。SVM の実装には Scikit-learn を用いる。SVM のカーネルは RBF カーネル (Radial Bases Function kernel) を用いる。RBF カーネルは式 (4.1) によって定義される。RBF カーネルは C, γ の 2 種類のハイパーパラメータがあり、本実験ではそれぞれ、PSO(Particle Swarm Optimization) を用いて定める。探索範囲は、 $1 \sim 1000, 10^{-12} \sim 1$ とし、最も識別率が高かった値を用いる。

$$K(x, x') = \exp(-\gamma [\|x - x'\|^2]) \quad (4.1)$$

4.4 学習モデル

4.4.2 提案する特徴と 2D-CNN モデル

2D-CNN は、通常 2 次元画像に対して適用される。しかし、本研究で用いる脳活動データは 3 次元画像であるため、そのまま 2D-CNN に適用することはできない。そこで本研究では、3 次元脳機能画像を 2D-CNN に適用するための特徴抽出・識別を行うための特徴選択を提案する。MRI 画像には 2 次元断面ごとに、以下の 3 つに定義される。

horizontal(水平断面) 頭立地で地面に平行な断面

sagittal(矢状断面) 身体を左右に分割する断面

coronal(冠状断面) 矢状面に垂直に交わり、身体を前後に分割する断面

これらの断面のうち 1 つの断面を用いる方法がまず考えられるが、これでは、スライス間にまたがる局所特徴が失われることが考えられる。本研究では、ある 1 つの断面と、新たにスライス面に直交する 2 方向の面についてリスライスを行い、これも入力画像として用いることを提案する。すなわち元の画像が水平断面のスライスである場合に、新たに矢状断面、冠状断面のスライスを再構成し、これも入力画像として追加する。

具体的な方法を図 4.1, 図 4.2 に示す。それぞれの断面で、スライスごとに画像を並べる。この際、水平断面が他の 2 面に比べサイズが小さく、それを同サイズにする必要がある。ここでは、配列に 0 を代入することでサイズを同じにしている。並べた画像の配列をそれぞれ 1 つのレコードとして扱うパターンを提案手法 1、3 つの 2 次元画像の配列を一つの配列とし、それを 1 レコードとして扱うパターンを提案手法 2 とする。

本研究では、水平断面、矢状断面、冠状断面のみを学習に用いる場合と、提案手法 1 と提案手法 2 の 5 パターンで学習を行う。それぞれのネットワーク構成は表 4.2 のような構成である。

4.5 交差検証法

4.4.3 3D-CNN モデル

3D-CNN モデルは、3次元構造の配列をそのまま入力することができるため、脳機能画像の形状を保持したまま学習することができる。そのため2D-CNNのように3つの断面図を1つのレコードとしてまとめるような処理をする必要がない。本研究では、表4.3～4.6のような4種類のネットワーク構成で検証する。

4.5 交差検証法

構築した学習モデルの評価として、K-分割交差検証法 (K-Fold Cross-Validation:CV) を用いる。K-分割交差検証法は、データセット数が少ない場合に用いられる制度評価手法である。データセットをK等分に分割し、そのうちの1グループをテストデータ、残りのグループを学習データとして用いるということを、全てのグループがテストデータとして用いられるまで繰り返す手法である。これによって、学習データの偏りによる汎化性能の低下を抑えることができる。本研究では10分割で行う。具体的な手順を以下に記す。

4.5 交差検証法

1. データセットを K グループに等分割
2. 1 グループをテストデータ, 残り $K - 1$ グループを学習データとしたときの識別率を算出
3. 手順 1, 2 を K 回繰り返し, 得られた K 回分の識別率の平均をその学習モデルの評価とする

4.5 交差検証法

表 4.1 実験環境

CPU	Core i7-7700K(4.20GHz)
GPU	GTX1080Ti
メモリ	32GB

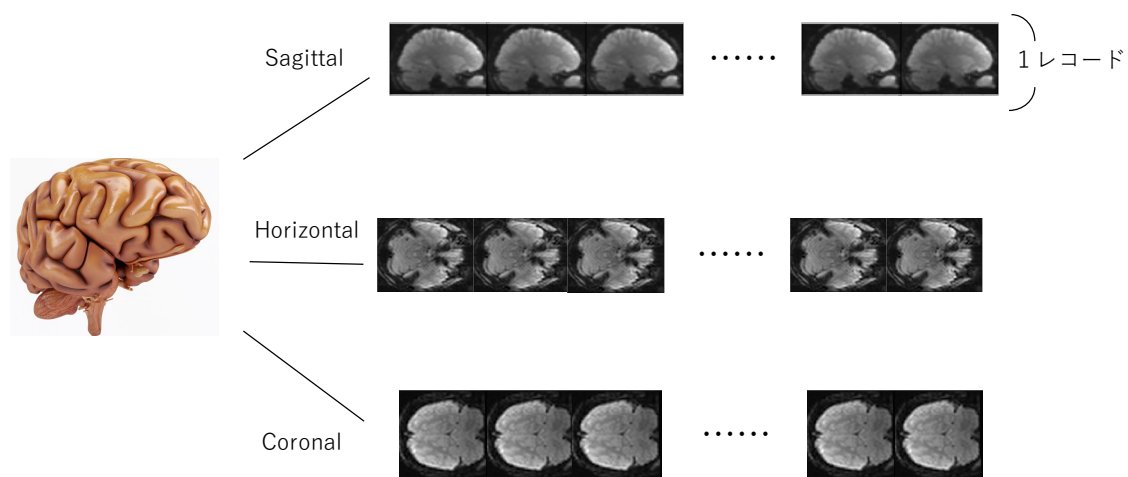


図 4.1 提案手法 1

4.5 交差検証法

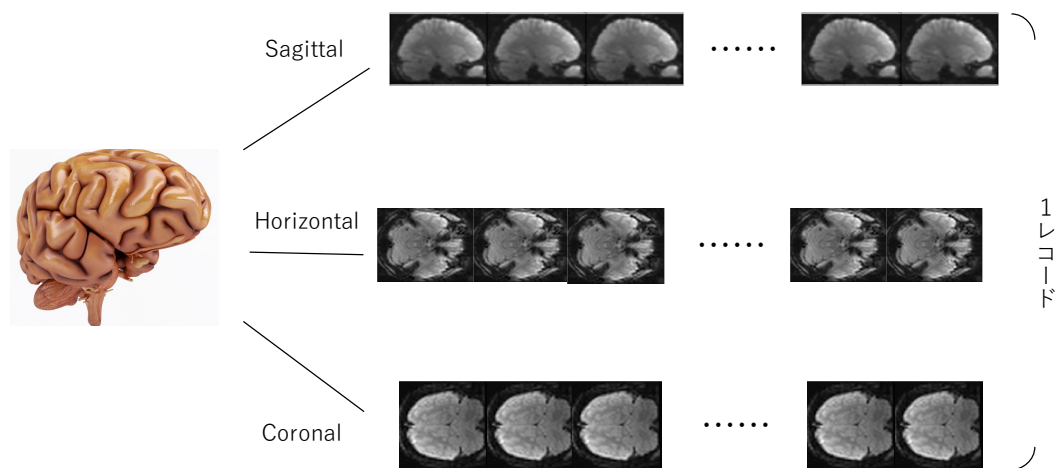


図 4.2 提案手法 2

表 4.2 2D-CNN の学習モデル

layer	filter	stride	unit	f(x)
1 層目：畳み込み層	$3 \times 3 \times 20$	1	-	ReLU
2 層目：プーリング層	2×2	2	-	-
3 層目：畳み込み層	$3 \times 3 \times 20$	1	-	ReLU
4 層目：プーリング層	2×2	2	-	-
5 層目：畳み込み層	$3 \times 3 \times 20$	1	-	ReLU
6 層目：プーリング層	2×2	2	-	-
7 層目：畳み込み層	$3 \times 3 \times 20$	1	-	ReLU
8 層目：プーリング層	2×2	2	-	-
9 層目：全結合層	-	-	1000	ReLU
10 層目：出力層	-	-	2	softmax

4.5 交差検証法

表 4.3 3D-CNN の学習モデル 1

layer	filter	stride	unit	f(x)
1 層目：畳み込み層	$3 \times 3 \times 3 \times 20$	1	-	ReLU
2 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
3 層目：畳み込み層	$3 \times 3 \times 3 \times 20$	1	-	ReLU
4 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
5 層目：畳み込み層	$3 \times 3 \times 3 \times 20$	1	-	ReLU
6 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
7 層目：畳み込み層	$3 \times 3 \times 3 \times 20$	1	-	ReLU
8 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
9 層目：全結合層	-	-	1000	ReLU
10 層目：出力層	-	-	2	softmax

表 4.4 3D-CNN の学習モデル 2

layer	filter	stride	unit	f(x)
1 層目：畳み込み層	$3 \times 3 \times 3 \times 30$	1	-	ReLU
2 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
3 層目：畳み込み層	$3 \times 3 \times 3 \times 30$	1	-	ReLU
4 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
5 層目：畳み込み層	$3 \times 3 \times 3 \times 30$	1	-	ReLU
6 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
7 層目：畳み込み層	$3 \times 3 \times 3 \times 30$	1	-	ReLU
8 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
9 層目：全結合層	-	-	1000	ReLU
10 層目：出力層	-	-	2	softmax

4.5 交差検証法

表 4.5 3D-CNN の学習モデル 3

layer	filter	stride	unit	f(x)
1 層目：畳み込み層	$2 \times 2 \times 2 \times 20$	1	-	ReLU
2 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
3 層目：畳み込み層	$2 \times 2 \times 2 \times 20$	1	-	ReLU
4 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
5 層目：畳み込み層	$2 \times 2 \times 2 \times 20$	1	-	ReLU
6 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
7 層目：畳み込み層	$2 \times 2 \times 2 \times 20$	1	-	ReLU
8 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
9 層目：全結合層	-	-	1000	ReLU
10 層目：出力層	-	-	2	softmax

表 4.6 3D-CNN の学習モデル 4

layer	filter	stride	unit	f(x)
1 層目：畳み込み層	$2 \times 2 \times 2 \times 30$	1	-	ReLU
2 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
3 層目：畳み込み層	$2 \times 2 \times 2 \times 30$	1	-	ReLU
4 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
5 層目：畳み込み層	$2 \times 2 \times 2 \times 30$	1	-	ReLU
6 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
7 層目：畳み込み層	$2 \times 2 \times 2 \times 30$	1	-	ReLU
8 層目：プーリング層	$2 \times 2 \times 2$	2	-	-
9 層目：全結合層	-	-	1000	ReLU
10 層目：出力層	-	-	2	softmax

4.5 交差検証法

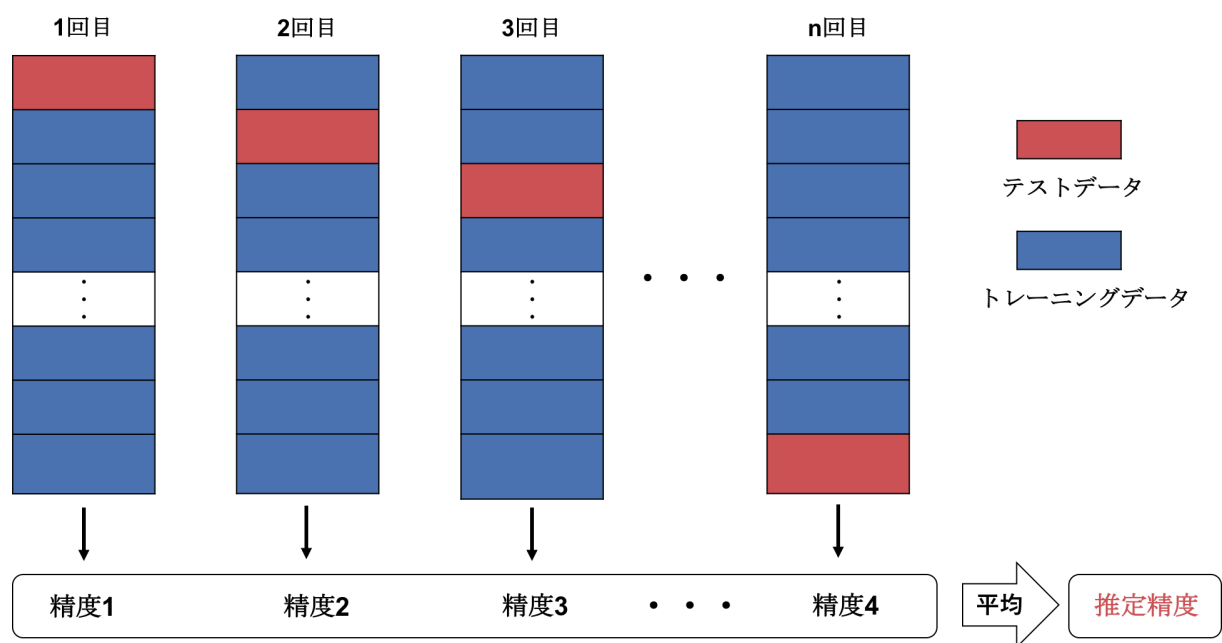


図 4.3 交差検証法

第 5 章

実験結果

本項では，被験者ごとの学習結果を述べる．

5.1 被験者 A

被験者 A の脳活動データを学習し，算出された結果を図 5.1 に示す．提案手法 2 を用いた場合に，識別率 80.2%と最も高くなった．SVM は 66.2%となり被験者 A においては従来手法よりも提案手法 2 を用いた場合のほうが高い識別率となった．3D-CNN は 73.6%となりこちらも提案手法を用いた場合のほうが高い識別率となった．また，coronal，horizontal，sagittal の 3 つを比較した場合には識別率に大きな差は無かった．

5.2 被験者 B

被験者 B の脳活動データを学習し，算出された結果を図 5.2 に示す．SVM を用いた場合に，識別率 85.3%と最も高くなった．提案手法 2 を用いた場合は 80.0%となり，被験者 B においては提案手法よりも従来手法を用いた場合のほうが高い識別率となった．3D-CNN は 82.5%となり提案手法より高い識別率となったが，従来手法である SVM よりは低い識別率となった．また，coronal，horizontal，sagittal の 3 つを比較した場合には識別率に大きな差は無かった．

5.3 被験者 C

5.3 被験者 C

被験者 C の脳活動データを学習し，算出された結果を図 5.3 に示す．horizontal を用いた場合に，識別率 69.4% と最も高くなった．SVM と提案手法を比較すると，提案手法のほうがやや高い識別率とはなったが，大きな差はなかった．3D-CNN と比較した場合も提案手法の方が高い識別率とはなったが，大きな差はなかった．また，coronal，horizontal，sagittal の 3 つを比較した場合には識別率に大きな差は無かった．被験者 C は全体的に他の被験者に比べて識別率が低くなった．

5.4 被験者 D

被験者 A の脳活動データを学習し，算出された結果を図 5.4 に示す．提案手法 2 を用いた場合に，識別率 85.7% と最も高くなった．SVM は 60.4% となり被験者 A においては従来手

5.5 被験者 E

法よりも提案手法 2 を用いた場合のほうが高い識別率となった。3D-CNN は 77.1%となりこちらにも提案手法を用いた場合のほうが高い識別率となった。また、coronal, horizontal, sagittal の 3 つを比較した場合には sagittal がやや高い識別率となった。

5.5 被験者 E

被験者 A の脳活動データを学習し、算出された結果を図 5.5 に示す。提案手法 2 を用いた場合に、識別率 77.2%と最も高くなった。SVM は 70.2%となり被験者 A においては従来手法よりも提案手法 2 を用いた場合のほうが高い識別率となった。3D-CNN は 73.5%となりこちらにも提案手法を用いた場合のほうが高い識別率となった。また、coronal, horizontal, sagittal の 3 つを比較した場合には識別率に大きな差は無かった。

5.6 被験者 F

5.6 被験者 F

被験者 A の脳活動データを学習し，算出された結果を図 5.5 に示す．提案手法 2 を用いた場合に，識別率 86.2%と最も高くなった．SVM は 75.0%となり被験者 A においては従来手法よりも提案手法 2 を用いた場合のほうが高い識別率となった．3D-CNN は 80.6%となりこちらも提案手法を用いた場合のほうが高い識別率となった．また，coronal，horizontal，sagital の 3 つを比較した場合には識別率に大きな差は無かった．

5.7 被験者 G

被験者 A の脳活動データを学習し，算出された結果を図 5.7 に示す．提案手法 2 を用いた場合に，識別率 90.1%と最も高くなった．この結果は全被験者中で最も高い識別率である．SVM は 79.1%となり被験者 A においては従来手法よりも提案手法 2 を用いた場合のほうが

5.8 被験者 H

高い識別率となった．3D-CNN は 82.0%となりこちらも提案手法を用いた場合のほうが高い識別率となった．また，coronal，horizontal，sagittal の 3 つを比較した場合には sagittal だけが少し低い識別率となった．

5.8 被験者 H

被験者 H の脳活動データを学習し，算出された結果を図 5.8 に示す．提案手法 2 を用いた場合に，識別率 85.3%と最も高くなった．また，coronal，horizontal，sagittal の 3 つを比較した場合には識別率に大きな差は無かった．

被験者 A の脳活動データを学習し，算出された結果を図 5.8 に示す．提案手法 2 を用いた場合に，識別率 85.3%と最も高くなった．この結果は全被験者中で最も高い識別率である．SVM は 80.2%となり被験者 A においては従来手法よりも提案手法 2 を用いた場合のほうが高い識別率となった．3D-CNN は 78.8%となりこちらも提案手法を用いた場合のほうが高

5.8 被験者 H

い識別率となった。また，coronal，horizontal，sagital の 3 つを比較した場合には sagittal だけが少し低い識別率となった。

5.8 被験者 H

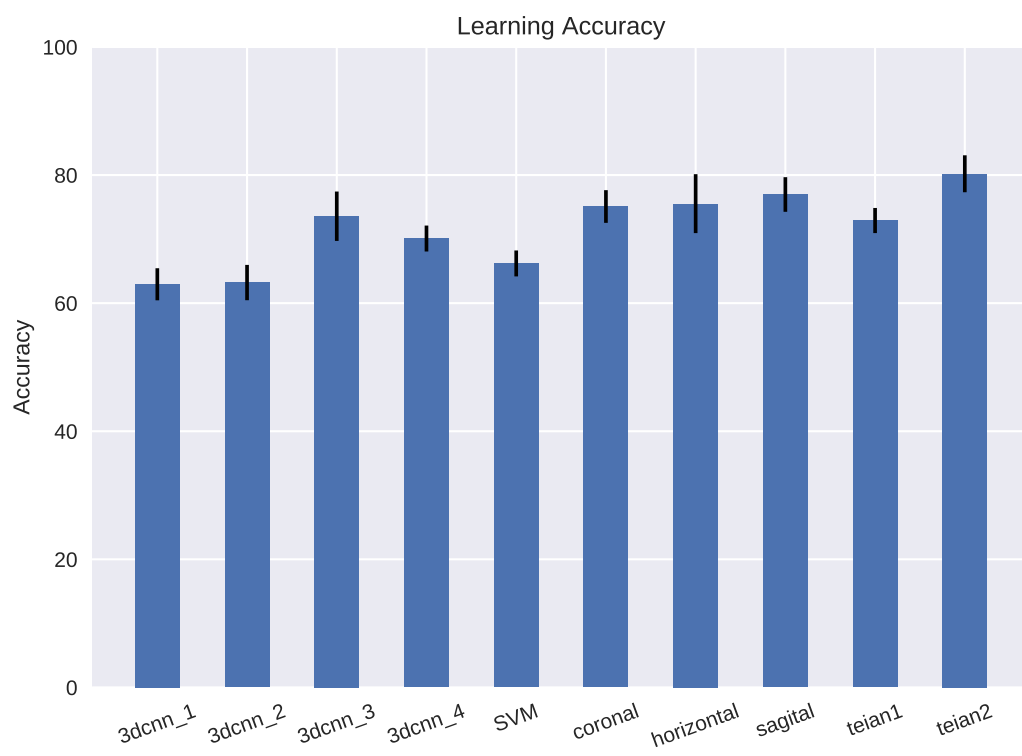


図 5.1 被験者 A の学習結果

5.8 被験者 H

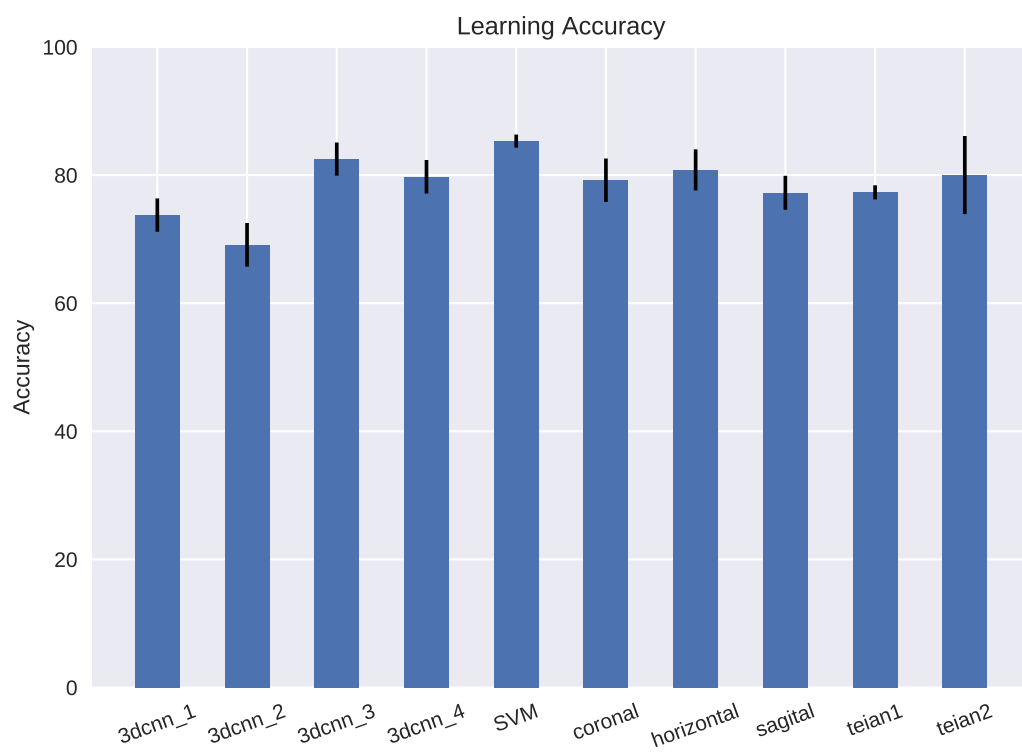


図 5.2 被験者 B の学習結果

5.8 被験者 H

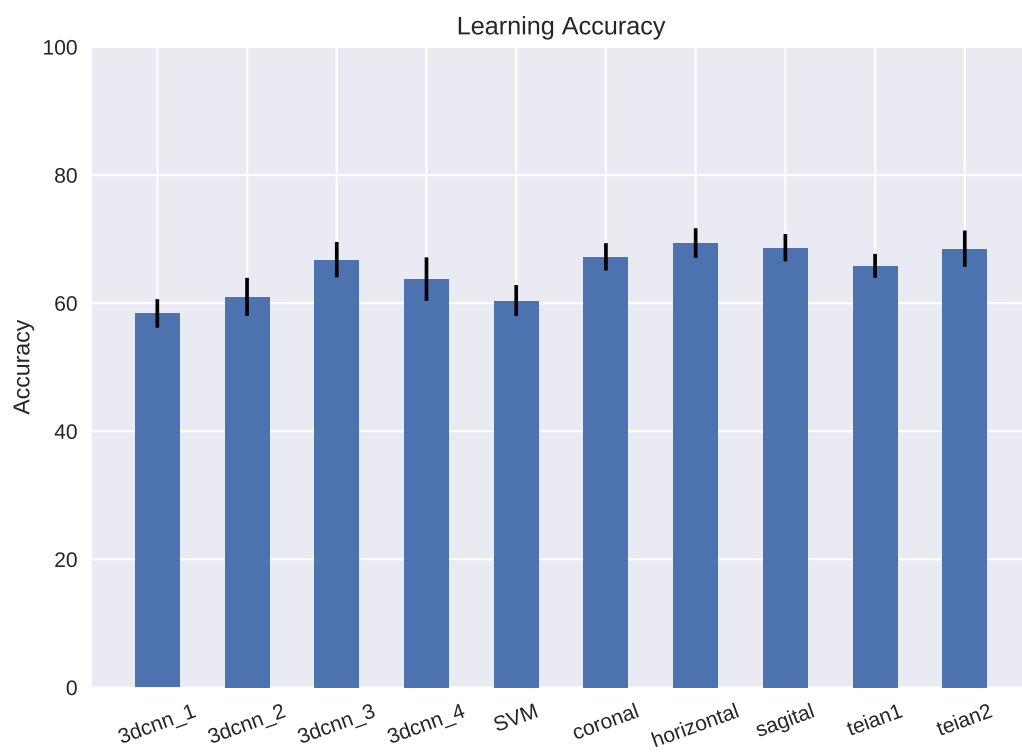


図 5.3 被験者 C の学習結果

5.8 被験者 H

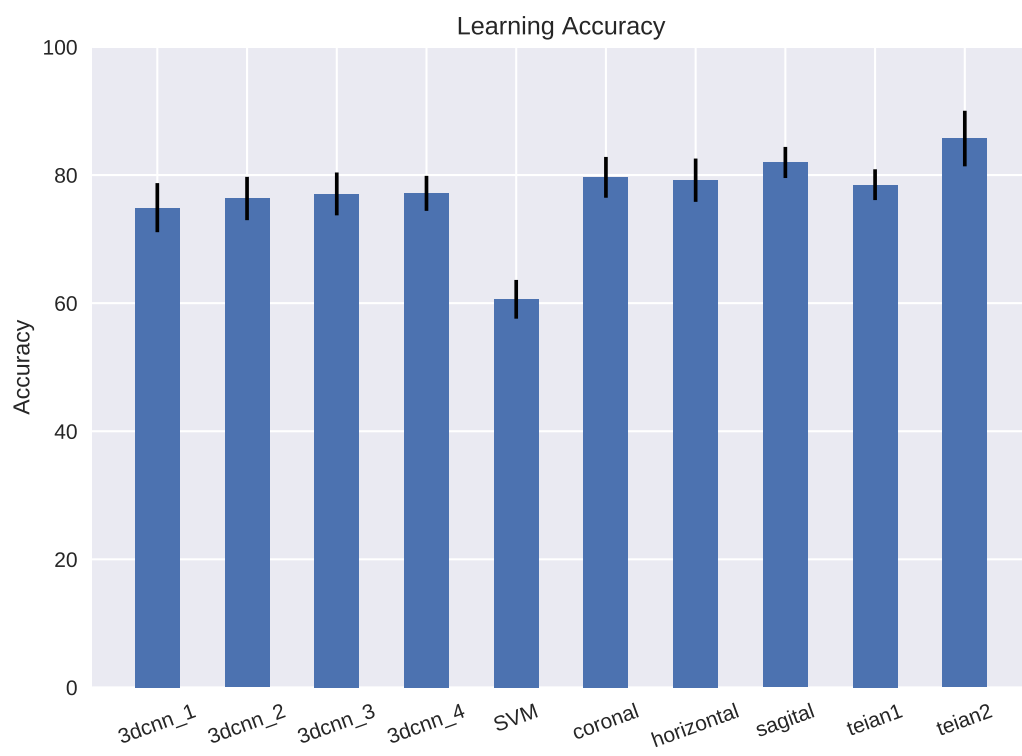


図 5.4 被験者 D の学習結果

5.8 被験者 H

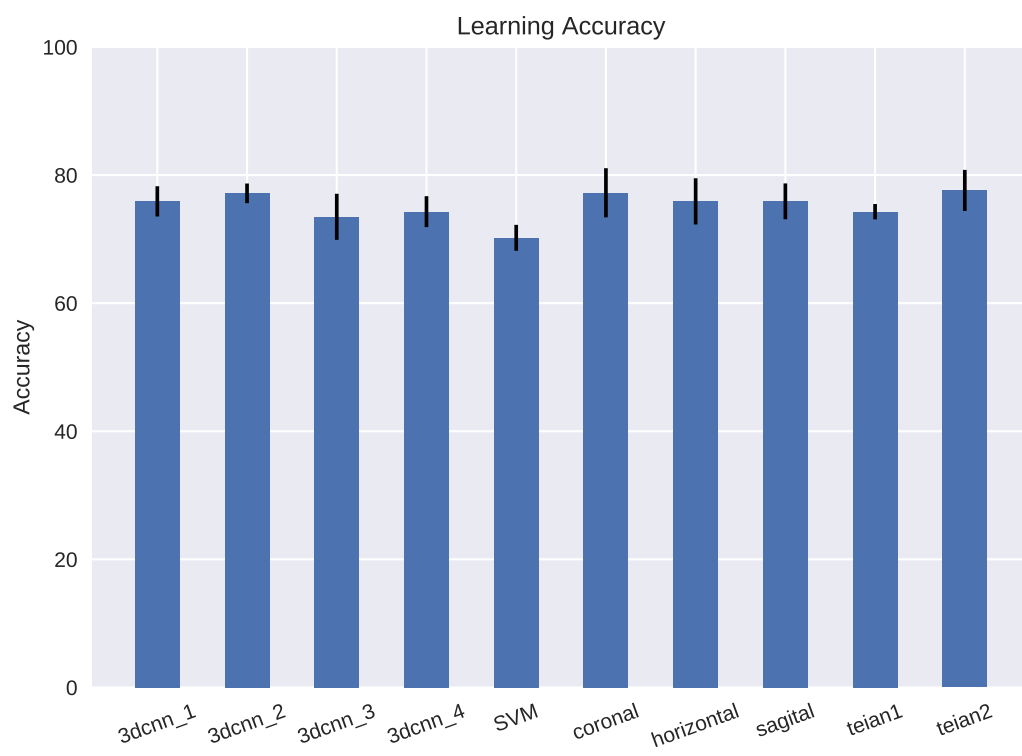


図 5.5 被験者 E の学習結果

5.8 被験者 H

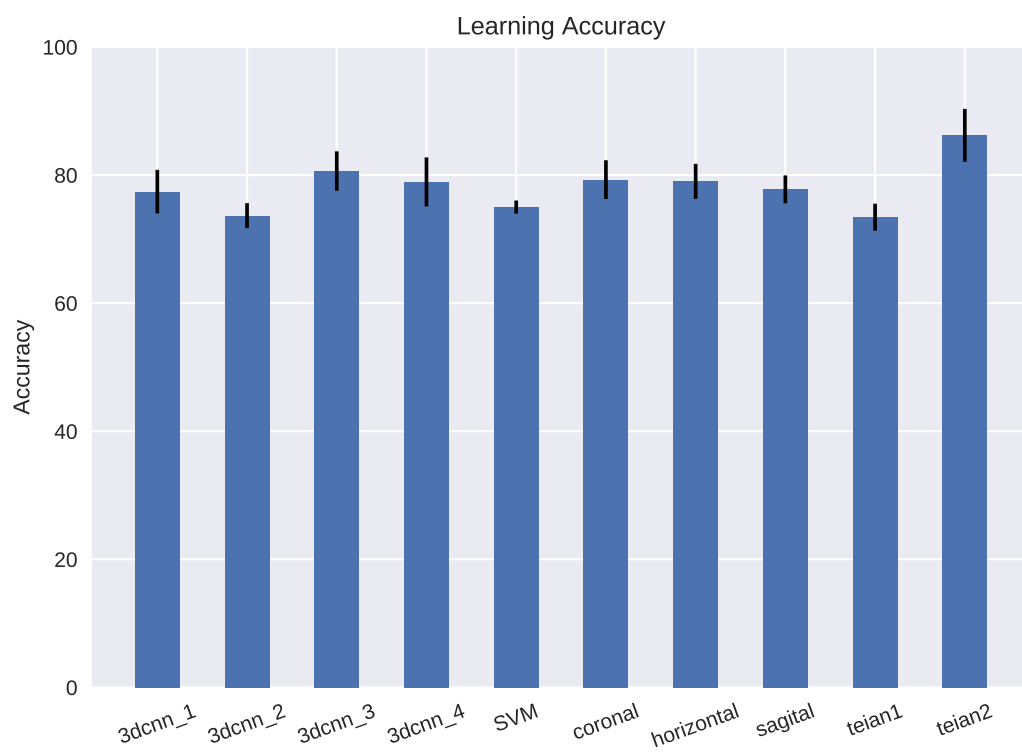


図 5.6 被験者 F の学習結果

5.8 被験者 H

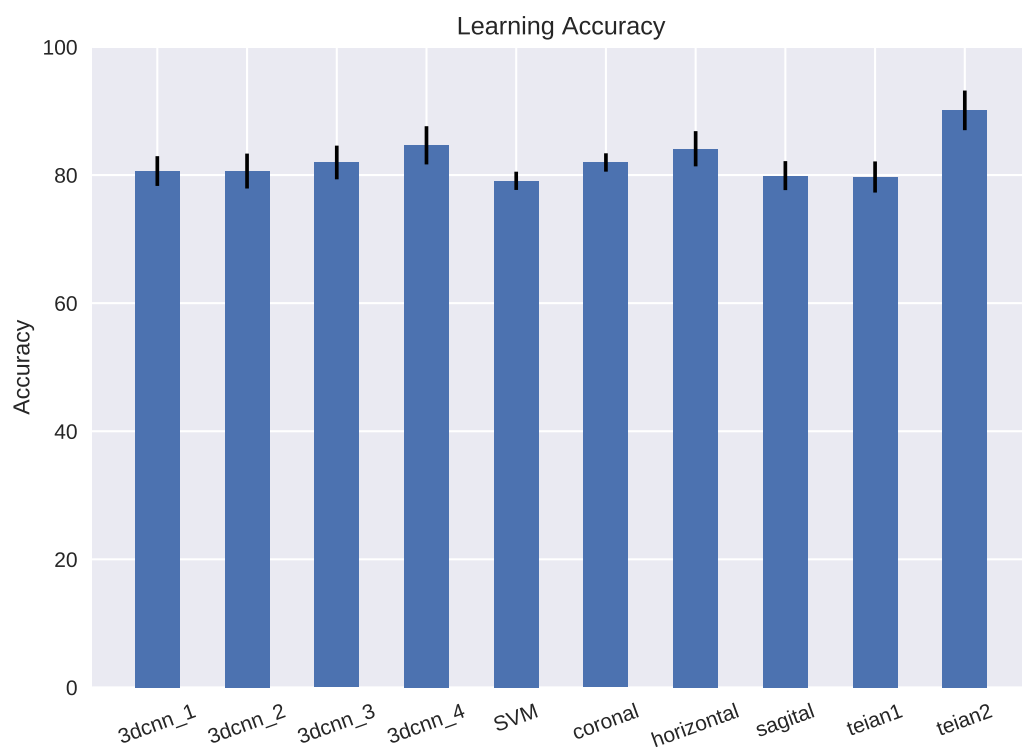


図 5.7 被験者 G の学習結果

5.8 被験者 H

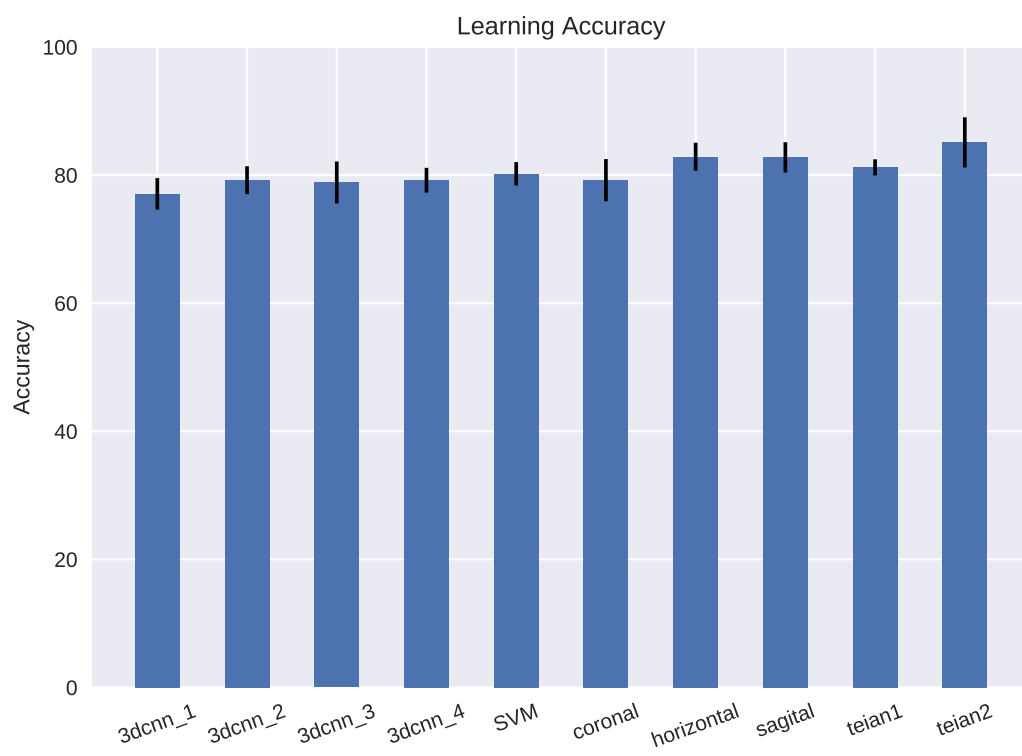


図 5.8 被験者 H の学習結果

第 6 章

考察

6.1 提案手法と従来の 2D-CNN および SVM との比較

それぞれの学習方法で算出した識別率を表 2 に示す。提案手法および、スライス画像それぞれに対して従来の 2D-CNN を適用した学習と、SVM を用いた学習を比較すると、被験者 B を除き最も高い識別率となった。このことから、CNN は人間の感情の脳情報デコーディングにおいて有効であると考えられる。提案手法 2 と、スライス画像それぞれに対して従来の 2D-CNN を適用した学習を比較すると、被験者 B を除き提案手法の方が高い識別率となった。また、提案手法 2 の場合に被験者 G で 90.1% と最も高い識別率となった。このことから、2 次元 EPI 画像のスライスをそのまま用いた場合に失われていた局所特徴が、提案手法では失われずに表現できていると考えられる。

6.2 提案手法と 3D-CNN の比較

提案手法と 3D-CNN を比較した場合、被験者 B を除き、提案手法の方が高い識別率となった。さらに、提案手法と 3D-CNN における最大 batch size と、エポックごとの学習にかかる時間を表 3 に示す。最大 batch size を比較すると、提案手法の方がより batch size を大きくすることができることがわかった。つまり、より低スペックなマシンにおいても学習を行うことができると言える。さらに、学習時間を比較すると提案手法の場合により早く学習が終わることがわかった。これらのことから、提案手法は 3 次元 fMRI 画像からの脳情報デコーディングへの 2D-CNN の適用において有効であると考えられる。3D-CNN があま

6.2 提案手法と 3D-CNN の比較

り高い識別率とならなかった理由に、2D-CNN よりも 1 次元増えたことで、学習に必要なメモリサイズが膨大になってしまうことの対策としてモデルの規模感を 2D-CNN と異なるものにしたことが理由として挙げられる。これによって適切な特徴抽出ができていない可能性が考えられる。そのため、よりハイスペックなマシンであれば学習モデルのチューニングによってより高い識別率となる可能性はある。

6.2 提案手法と 3D-CNN の比較

表 6.1 実験結果

	A	B	C	D	E	F	G	H
2D-CNN(提案手法 1)	72.9%	77.3%	65.8%	78.5%	74.3%	73.4%	79.7%	81.2%
2D-CNN(提案手法 2)	80.2%	80.0%	68.5%	85.7%	76.6%	86.2%	90.1%	85.3%
2D-CNN(horizontal)	75.5%	80.8%	69.4%	79.2%	75.9%	79.0%	84.1%	82.9%
2D-CNN(sagittal)	77.0%	77.2%	68.7%	82.0%	75.9%	77.8%	79.9%	82.8%
2D-CNN(coronal)	75.1%	79.2%	67.2%	79.6%	77.2%	79.3%	82.0%	79.2%
SVM	66.2%	85.3%	60.4%	60.6%	70.2%	75.0%	79.1%	80.2%
3D-CNN	73.6%	82.5%	66.8%	77.1%	73.5%	80.6%	82.0%	78.8%

表 6.2 実験結果

	最大 batch size	時間 (s/epoch)
提案手法	16	3.24
3D-CNN	4	4.30

第7章

結論

脳情報デコーディングを行う際の機械学習アルゴリズムに CNN を用いたものはいくつか報告されているが、これらは物体認識などの感情と比較して低次脳機能のデコーディングである。また、脳活動データは本来 3 次元であるが、多くの研究では 2 次元のスライス画像の集まりに 2D-CNN を適用している。そこで本研究では、これまで SVM で行われていた感情の脳情報デコーディングを対象として、3D-CNN の適用に加え、新たに提案する特徴選択を行い 2D-CNN を適用し、その性能を明らかにしている。

評価実験では、被験者 8 名の脳活動データに対して、SVM、従来の 2D-CNN、提案手法を用いた 2D-CNN、3D - CNN を用いて学習し、識別率を比較する。結果として提案手法が 90.1%と最も高い識別率となり、人間の感情の脳情報デコーディングにおいて提案手法が有効であることを示した。また、提案手法を用いた 2D-CNN は 3D-CNN よりも少ない計算資源で学習することができることがわかった。

しかし、学習モデルの構築に関する検討が十分にできていない。今後の課題として、CNN におけるハイパーパラメータの探索や、VGG16 などの既存の学習モデルを用いた転移学習の検討が挙げられる。

謝辞

本研究を進めるにあたり，ご指導していただきました吉田真一准教授には心から感謝致します．学部時代から計4年間に渡り，非常に丁寧かつ熱心なご指導をいただきました．私が無事，修士論文を執筆することができるのは先生のお力添えあってのことです．深く感謝致します．

そして副査を担当して頂いた岩田誠教授と，門田宏准教授に深く感謝致します．お二人からはそれぞれの専門の知見から多くのアドバイスを頂きました．本当にありがとうございました．

また，4年間共に仲間として研究活動を行ってきたM2の4名にも深く感謝致します．1人だけでは研究をここまで進めることはできなかったと思います．さらに研究活動に協力してくださった先輩や後輩の方々にも心から感謝致します．ありがとうございました．

参考文献

- [1] Orhan Firat , et al., “DEEP LEARNING FOR BRAIN DECODING,” IEEE International Conference on Image Processing , pp. 2784–2788, 2014 .
- [2] Tomoyasu Horikawa , et al., “Hierarchical Neural Representation of Dreamed Objects Revealed by Brain Decoding with Deep Neural Network Features,” Front. Comput. Neurosci., Vol. 11, Article 4 , 2017 .
- [3] Haiguang Wen , et al., “Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision,” Cerebral Cortex , pp. 1–25, 2017 .
- [4] Kamitani, Y., and Tong, F., “ Decoding the visual and subjective contents of the human brain, ”Nature Neuroscience, Vol.8, No.5, pp.679-685, 2005.
- [5] 高橋宏和, “ fMRI による脳情報デコーディングを用いた快・不快画像の識別, ”平成 26 年度高知工科大学学士学位論文 , 2015.
- [6] S. A. Huettel, A. W. Song, G. McCarthy, and 福山秀直, fMRI 原理と実践. メディカル・サイエンス・インターナショナル, 2016
- [7] Cortes, C., and Vapnik, V.N., “ Support-Vector Networks ”, Machine Learning, Vol.20, Issue 3, pp.273-297, 1995.
- [8] 平井有三, “ はじめてのパターン認識 ”, 森北出版株式会社 , 2012 .
- [9] K. Fukushima, S. Miyake, and T. Ito, “ Neocognitron: A neural network model for a mechanism of visual pattern recognition, ”IEEE transactions on systems, man, and cybernetics, no. 5, pp. 826834, 1983.
- [10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “ Gradient-based learning applied to document recognition, ”Proceedings of the IEEE, vol. 86, no. 11, pp. 22782324, 1998.
- [11] V Nair and G E Hinton, ”Rectified linear units improve restricted Boltzmann ma-

参考文献

- chines” ICML’10 Proceedings of the 27th International Conference on International Conference on Machine Learning, pp. 807-814, 2010.
- [12] 神蔦敏弘, 麻生英樹, 安田宗樹, 前田新一, and 岡野原大輔, 深層学習. 近代科学社, 2015.
- [13] S. Ji, W. Xu, M. Yang, and K. Yu, “ 3d convolutional neural networks for human action recognition, ”IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pp. 221231, 2013.