

ニューラルネットワークを用いた脳波による  
運動想起判別手法の開発  
-学習における人工脳波の有用性と CNN・  
LSTM を用いたネットワーク構造の提案-

平成 30 年度 修士論文  
横浜市立大学大学院  
生命ナノシステム科学研究科  
物質システム科学専攻  
高橋馨子

## 目次

1 序論 .....	5
1.1 背景 .....	5
1.2 課題と目的 .....	6
2 関連研究 .....	7
2.1 Brain Computer Interface .....	7
2.1.1 脳波 .....	9
2.1.2 事象関連同期・事象関連脱同期 .....	11
2.1.3 Common spatial pattern .....	13
2.2 人工ニューラルネットワーク .....	14
2.2.1 ANN の構造 .....	15
2.2.2 誤差逆伝搬法 .....	17
2.2.3 疊み込みニューラルネットワーク .....	18
2.2.4 Long short-term memory .....	21
2.3 経験的モード分解 .....	24
3 ANN による運動想起判別手法 .....	25
3.1 使用データ .....	25
3.2 人工脳波作成 .....	26
3.3 時間周波数解析 .....	28
3.4 ANN での処理 .....	29
3.4.1 CNN の構造 .....	30
3.4.2 CNN+LSTM の構造 .....	34
4 結果 .....	38
4.1 Morlet mother wavelet のパラメータ比較 .....	38
4.2 CNN および CNN+LSTM による識別結果 .....	39
5 考察 .....	43
5.1 従来の手法との比較 .....	43
5.2 CNN と CNN+LSTM の比較 .....	44
5.3 学習データ数と識別精度 .....	46
5.4 人工データの有用性 .....	49
6 結論 .....	55

7	文献 .....	56
8	謝辞 .....	58
9	研究発表実績 .....	59

ニューラルネットワークを用いた脳波による  
運動想起判別手法の開発  
-学習における人工脳波の有用性と CNN・  
LSTM を用いたネットワーク構造の提案-

高橋馨子

主指導教員 ミケレットルジェロ教授  
副指導教員 野々瀬真司准教授  
副指導教員 山田重樹准教授

# ニューラルネットワークを用いた脳波による運動想起判別手法の開発-学習における人工脳波の有用性と CNN・LSTM を用いたネットワーク構造の提案-

物質システム科学専攻 高橋馨子  
指導教員 ミケレット ルジェロ

## 【重要語句】

BCI (brain-computer interface) : 脳信号を読み取り機械に伝える技術.

NN (neural network) : 人間の脳の神経回路を模擬した情報処理システムの総称.  
複雑な特徴パターンの認識能力に優れる.

CNN (convolutional neural network) : NN の一種. データを畳み込むことでその特徴を抽出する. 特に画像認識の分野で優れた性能をもつ.

LSTM (long short-term memory) : 時系列データを処理するために作られた NN のモデル. 過去に入力された情報をネットワーク内に記憶し, 学習に利用する.  
特に自然言語処理の分野で優れた性能をもつ.

EMD (empirical mode decomposition) : 経験的モード分解. 時系列信号を固有モード関数 (IMF) と呼ばれる狭帯域かつトレンド成分をもたない信号の和に分解する手法. 非線形・非定常な信号の解析に適する.

## 【研究の背景と目的】

人間はある運動を想起しただけで, 実際にその運動をした時と同じ脳波が同じ位置から発生する. この想起時の脳波を BCI に利用することで, 高齢者や身体障害者の運動補助に役立つことが期待されている. しかし脳波特性は発現場所や発現のしやすさに個人差がある. また, 脳波は非常に微小な電気現象であるため明瞭な波形の取得が難しく, ノイズの影響も受けやすい. そのため脳波から特徴を正確に読み取る手法の開発が重要となり, そのひとつとして, パターン認識能力に優れている NN を BCI へ応用する試みがなされている.

しかし, 一般的に NN を学習させるには多量のデータが必要であり, 運動補助への BCI 利用を考えると, 必要分の脳波データを取得することはユーザー含め関係者に対し大きな負担になると懸念される.

本研究では, 限られた実測脳波から Dinares-Ferran (2018) が提案した EMD による人工脳波作成手法を用い学習データ数を増やすことで, NN の判別精度向上を図った. また, CNN と LSTM を組み合わせた NN を提案し, その判別精度を検証した.

## 【データ】

評価実験には、Phan (2010) が行った左手右手の運動想起時のデータを用いた。これは 2 名 (subject 1, subject 2) の被験者の頭に 62 個の電極をつけ、左右各 2 秒間想起した時の脳波を 100 回測定し、8~30Hz のバンドパスフィルタにかけたものである。図 1 は電極チャンネルの位置を示す。サンプリング周波数は 500Hz であるため、1 施行分のデータは  $1000$  ( $500\text{Hz} \times 2\text{秒}$ )  $\times 62$  の行列として表される。

本研究では上記 200 個のデータのうち、ANN の学習に 60 個、学習結果の検証に 140 個のデータを使用した。

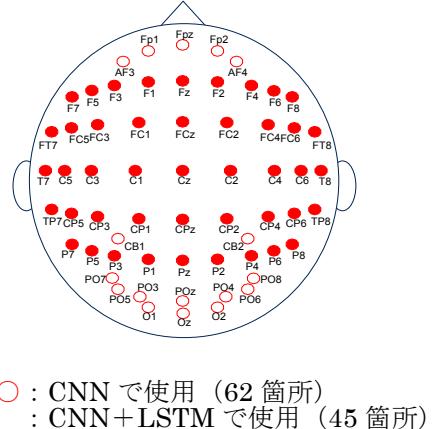


図 1. 頭上から見た脳波計電極の位置。

## 【人工脳波の作成手法】

学習用実測脳波から学習用人工脳波の作成には Dinares-Ferran (2018) の手法を用いた。まず実測脳波信号  $EEG(t)$  を EMD により固有モード関数  $IMF(t)$  とトレンド成分  $r_n(t)$  に分解し、

$$EEG(t) = \sum_{k=1}^n IMF_k(t) + r_n(t)$$

同クラス（左手もしくは右手）に属する脳波から得られた  $IMF$  をランダムに組み合わせ直すことで、人工脳波を作成した。これらと作成に用いた実測脳波とを合わせて ANN の学習に用いた。

## 【NN の構造とデータ入力方法】

本研究で検証した CNN と CNN+LSTM の構造を図 2 に示す。

CNN へデータを入力するにあたって脳波を画像化するために、まず脳波をウェーブレット変換しチャンネル毎のスカラグラム（スケール：8~30 Hz）を作成した。即ち脳波データを電極 62 チャンネル毎の 1 次元データ ( $1000 \times 62$ ) から、周波数と時刻における信号の強さを表す 62 枚の 2 次元画像 ( $23 \times 50 \times 62$ ) に変換し、CNN へ入力した。CNN では畳み込み処理とプーリング処理を 2 回ずつ行った。

CNN+LSTM では図 1 に示した 45 チャンネルのデータを使用した。チャンネルの位置情報を入力データに含めるため、データの並べ替えを行った。脳波は CNN の場合と同様ウェーブレット変換した後 ( $23 \times 50 \times 45$ )、時間毎に図 1 に示したチャンネル順に並べ替え、 $5 \times 9$  の画像を 23 枚、50 組作成した ( $5 \times 9 \times 23 \times 50$ )。CNN+LSTM では、入力データを時間毎に畳み込み処理した後、LSTM を適用した。

CNN, CNN+LSTM どちらのネットワークも最後全結合層へデータが入力され, 2つの値が出力される. 学習過程ではこの出力値と正解ラベル (左手 : [1, 0], 右手 : [0, 1]) との誤差が小さくなるよう学習データを用いてネットワーク内のパラメータを最適化した. 判別精度の検証では, 学習済みのネットワークに検証データを入力した際の出力値から入力した脳波の判別を行い, その判別精度を評価した.

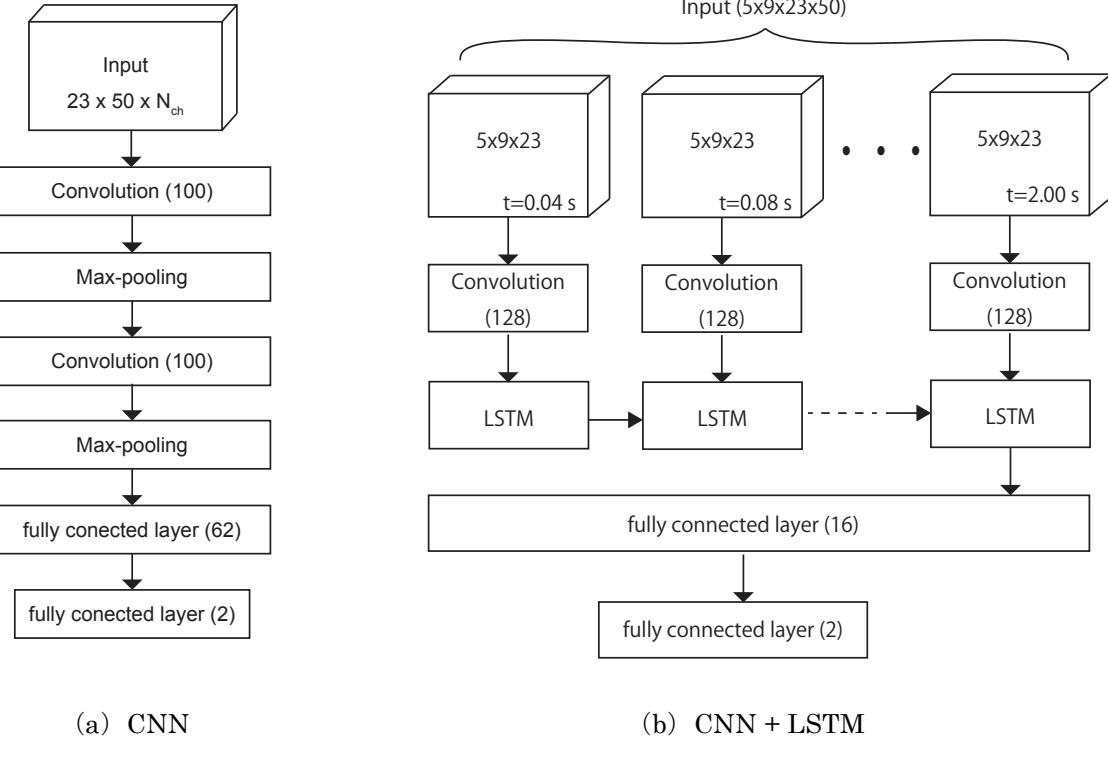


図 2. ニューラルネットワークの構造.

### 【結果と議論】

図 3 に, 左右手運動の判別について 2 種の ANN における, 学習データ数に対する判別精度を示す. 全体として CNN+LSTM は CNN と比べ高い精度が得られた. また, 実測データに人工データを加え学習することで, 判別精度向上がみられた. 特に subject1 の場合, 実測データ 30 個と人工データ 30 個で CNN+LSTM を学習させた場合と, 実測データ 60 個で学習させた場合では同程度の判別精度が得られた. このことから人工データは ANN の判別精度向上に有用であることが示唆される.

図 4 に, subject1 において実測データのみで CNN+LSTM を学習した場合の判別精度 (■) と, 人工データを加え 60 個にした場合の判別精度 (▲) の比較を示す.

(a) より, 判別精度 (■) は実測データ数の増加に従い向上するものの, 向上率は徐々に低下していくことがわかる. また, 実測データ 12 個以上の判別精度 (▲) は, 実測データだけの判別精度 (■) より向上するものの, 実測データ 6 個を 60 個に増

やした場合の精度は、実測データ 6 個のみの精度より低下している。この結果から、人工データ数に対して実測データ数が少なすぎると判別精度が低下することがわかった。(b) では (a) の x 軸を学習に用いた実測データ数の逆数とした。この逆数と判別精度 (■) はほぼ比例関係にある。これは学習データ数を増やし続けたとしても判別精度に上限があることを示す。更に、実測データ数の逆数と判別精度 (▲) との関係と、実測データ数の逆数と判別精度 (■) との関係を比較すると、交点があり、この交点 K の x 座標値が 0.11027 であることから、 $1/0.11027 \approx 9.06$ 、即ち 9 個以下の実測データを人工データで 60 個に増加し ANN 学習させても、判別精度向上には寄与しない、もしくは逆効果であることがわかる。以上より ANN 学習において実測データから人工データへの増加率に限界値があることが示唆される。

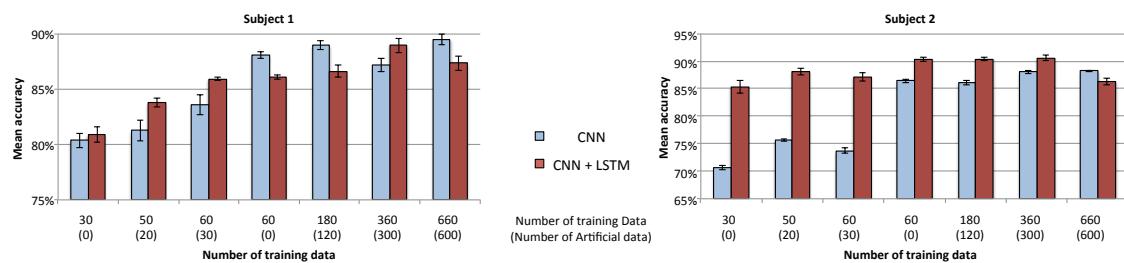


図 3. 人工脳波追加に対する判別能力の性能。

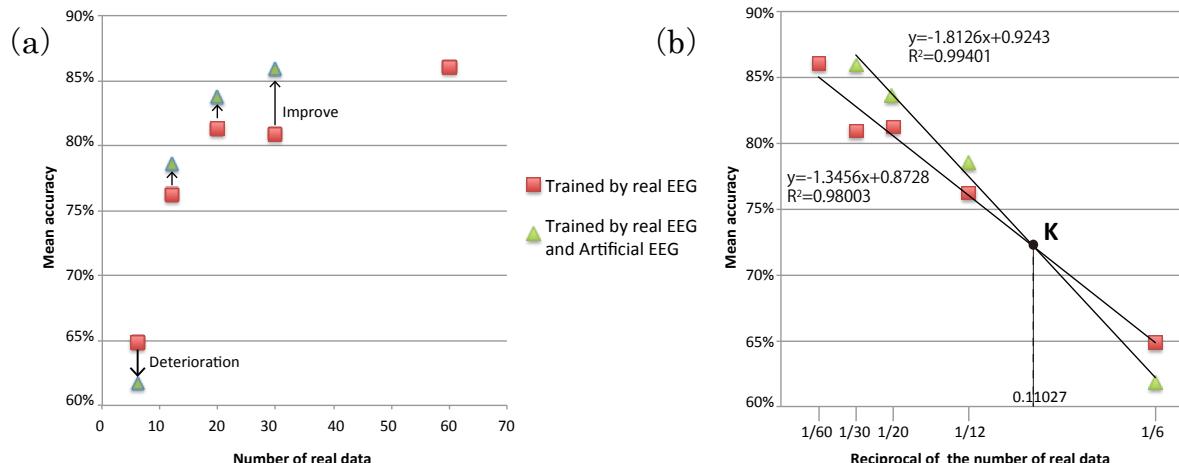


図 4. subject 1 における実測データ数に対する判別精度。

## 【参考文献】

1. Dinares-Ferran, J., R. Ortner, C. Guger, and J. Sole-Casals. (2018) A new method to generate artificial frames using the empirical mode decomposition for an EEG-based motor imagery BCI, *Frontiers Neuroscience*. 12:308.
2. Phan, A. H., and A. Cichocki. (2010) Tensor decompositions for feature extraction and classification of high dimensional datasets, *Nonlinear Theory and Its Applications, IEICE*, Volume 1, Issue 1, Pages 37-68

# 1 序論

## 1.1 背景

Brain computer interface (BCI) とは脳信号を読み取り機械に伝えることで、言語や動作を媒介せずに機械制御を行う技術の総称である (Wolpaw et al., 2002). この技術は身体障害を伴う高齢者や難病患者の運動補助や円滑なコミュニケーションなどに役立つとして、近年研究が活発に進められている (山ノ井, 2016). また、最近では健常者の思考を読み取るニューロマーケティングへの応用や、ゲームのコントローラーとしての応用なども見込まれており (Lance et al., 2012; Erp et al., 2012), 今後さらに実社会への適応が進められていくと考えられる。

BCI を使用するにあたり、脳内情報を命令に変換するためには脳波を正しく解読する必要があり、その性能は脳波のパターン認識技術に依存する。したがって、脳波を正確に識別するアルゴリズムの開発が重要となる。これまで利用してきた識別アルゴリズムには、線形分離やサポートベクターマシン(SVM)などがある。(Lotte et al., 2018). また、ニューラルネットワーク (ANN) による識別手法では、学習的に任意の非線形写像が得られることから、脳波特徴の個人差にも対応できると考えられ、注目されている (Zhang et al., 2018; Li et al., 2017) .

## 1.2 課題と目的

脳波は個人差が大きく、再現性も低い（伊藤・他, 2004）。とくに高齢者や患者の脳波はBCI使用時の疲労具合や投薬状況、モチベーションなどにも影響されるため、その状態は極めて多様である。また、脳波は非常に微小な電気現象であるため明瞭な波形の取得が難しく、ノイズの影響を受けやすい。そのため、統一された判断基準やパラメータを決定して汎用のシステムを構築することは非常に難しく、ユーザーの特性に合わせて個別に基準やパラメータを調整する必要がある。また、より正確な脳波解読を目指すのであれば、使用時ごとにキャリブレーションすることが好ましい。

識別器の精度は、キャリブレーション用として事前に取得された脳波データの数（測定試行回数）に比例する（Ramoser et al., 2000）。また、ANNを使用する場合、その学習には一般的に多量のデータが必要であるため、キャリブレーションには事前に十分な量の脳波データを取得する必要がある。しかし、長時間における事前データ取得の作業は、高齢者や患者といったユーザーを含め、関係者にとって大きな負担となることが懸念される。

本研究では、ANNを用いた脳波による運動想起判別アルゴリズムを開発することが目的であり、キャリブレーションにおける負担削減を目指す。そこで、限られた実測脳波データから Dinares-Ferran (2018) が提案した人工脳波作成手法を用いて学習データ数を増やすことで、ANNの判別精度向上を図った (Dinares-Ferran et al., 2018)。また、CNNとLSTMを組み合わせたANNを提案し、その判別精度を検証した。

## 2 関連研究

### 2.1 Brain Computer Interface

Brain computer interface (BCI) の基本的な構造を図 2-1 に示す。BCI では主に 5 つのステップを踏む。まず (1) 外部刺激やメンタルタスクによって脳活動を誘発させる。BCI はその活動を (2) 計測し、得られた信号から (3) 特徴抽出がなされ、(4) パターン認識などにより信号の判別が行われる。その後 (5) コマンドに変換され、機械を制御する。以下に (1) ~ (4) ステップの詳細を述べる。

#### (1) 脳活動の誘発

脳信号には知覚や認知処理、運動、運動想起などの特定の事象に関連して発生するものがある。脳波においてはこれを事象関連電位 (ERP) と呼ぶ。また、ある事象に対して脳波の特定の帯域のパワーが増幅/減少する現象を事象関連同期 (ERS) / 脱同期 (ERD) と呼ぶ。これらを意図的に誘発させることでその人の意思や考えを BCI へ伝える試みが近年多くなされている (開・金山, 2016)。ERS/ERD については 2.1.2 章で述べる。

#### (2) 脳活動の測定

脳活動を測定する手法は、大きくわけて侵襲脳計測法と非侵襲脳計測法のふたつがある。侵襲脳計測法では脳内に直接電極を埋め込むことで神経細胞のスパイク活動や皮質脳波活動を計測する手法である。この方法では高い計測精度が保たれるが、外科手術が必要なうえ、脳表面を傷つけてしまう可能性も高い。そのため現在は非侵襲脳計測を用いる BCI の研究が非常に多い。

非侵襲脳計測法には脳磁図(MEG)や核磁気共鳴画像装置(MRI)を用いた機能的核磁気共鳴画像法(fMRI)がある。しかし、これらの手法は大型で高価な装置を必要とするため、実社会に適用するのは難しい。一方、脳波(EEG)は計測装置が比較的小型で安価であるため、応用しやすい手法といえる。

#### (3) 特徴抽出

脳信号には、得たい特性を表す成分の他に生体現象や実験環境などに起因したノイズが含まれる。特徴抽出では、これらの不要な成分を除去し所望の成分のみ抽出する。また、脳波は時間とともに周波数成分やその出現領域が変化するため、時間周波数解析を行うことで特徴を抽出する方法もある。脳波については 2.1.1 章にて、また時間周波数解析については 3.3 章にて述べる。

#### (4) 脳信号の判別

観測された脳信号から被験者がどのような外部刺激をうけたか、もしくはメンタルタスクを行ったかを判別する。判別手法には特徴ベクトルから経験的に閾値を設けて識別する方法と、機械学習を用いる方法がある。前者の方法では、観測環境や被験者によって脳信号の特徴に揺らぎがあるため、安定した識別精度を得るのは難しい。一方後者の方法では、識別器のパラメータを識別クラスが既知の学習データから学習するため、より有効であると言われている。機械学習には線形法と非線形法がある。線形法にはフィッシャーの線形判別（LDA）や線形サポートベクターマシン（SVM）がある。また非線形法では非線形 SVM やニューラルネットワーク（ANN）を用いたディープラーニングがある。複雑なデータセットや、非常に大きなデータセットを使う場合、非線形手法のほうが優れた結果が得られることがわかっている（Muller et al., 2003）。また、一般的な機械学習ではデータからマニュアルで特徴量を抽出し、適切な分類器を作成するのに対し、ディープラーニングの場合は特徴抽出から分類モデルの作成までの処理方法を自動的に学習できる点から、近年多く研究がなされている。ANNについて2.2章で述べる。

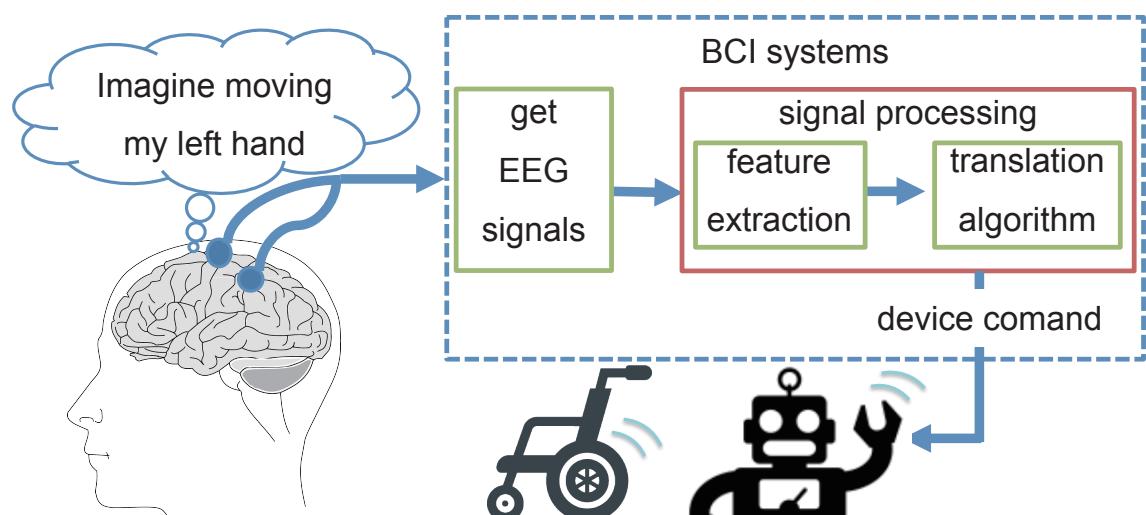


図 2-1 BCI システムの構成。

### 2.1.1 脳波

脳波は、脳内のニューロンが活動した際に生じる電流を、脳組織や頭蓋骨を経て頭皮上で観察したものである。脳波の電位は数十マイクロボルトとわずかである上に、測定環境などのノイズを含め複数の要因が複雑に混ざった状態で観測されるため、その特徴を取得するのは難しい。また、計測部位によっても現れる脳波は異なり、出現する特徴波には個人差がある（伊藤・他, 2004）。

脳波スペクトルには帯域ごとに名前が付けられている（Nunez and Srinivasan, 2007）。表 2-1 にそれぞれの名称とその周波数帯を示す。これらは違った生理学的な意義を反映することがわかっている。例えば、 $\gamma$  波は睡眠時に観測され、 $\alpha$  波は開眼時、眠い時、安静時などに後頭部中心に現れる。（開・金山, 2016）また、運動に関する脳活動としては、運動前に電位がゆっくり減少する準備電位や、感覚運動野において  $\mu$  波や  $\beta$  波の振幅が減少する現象が報告されている。（Pfurtscheller and Lopes da Silva, 1999）

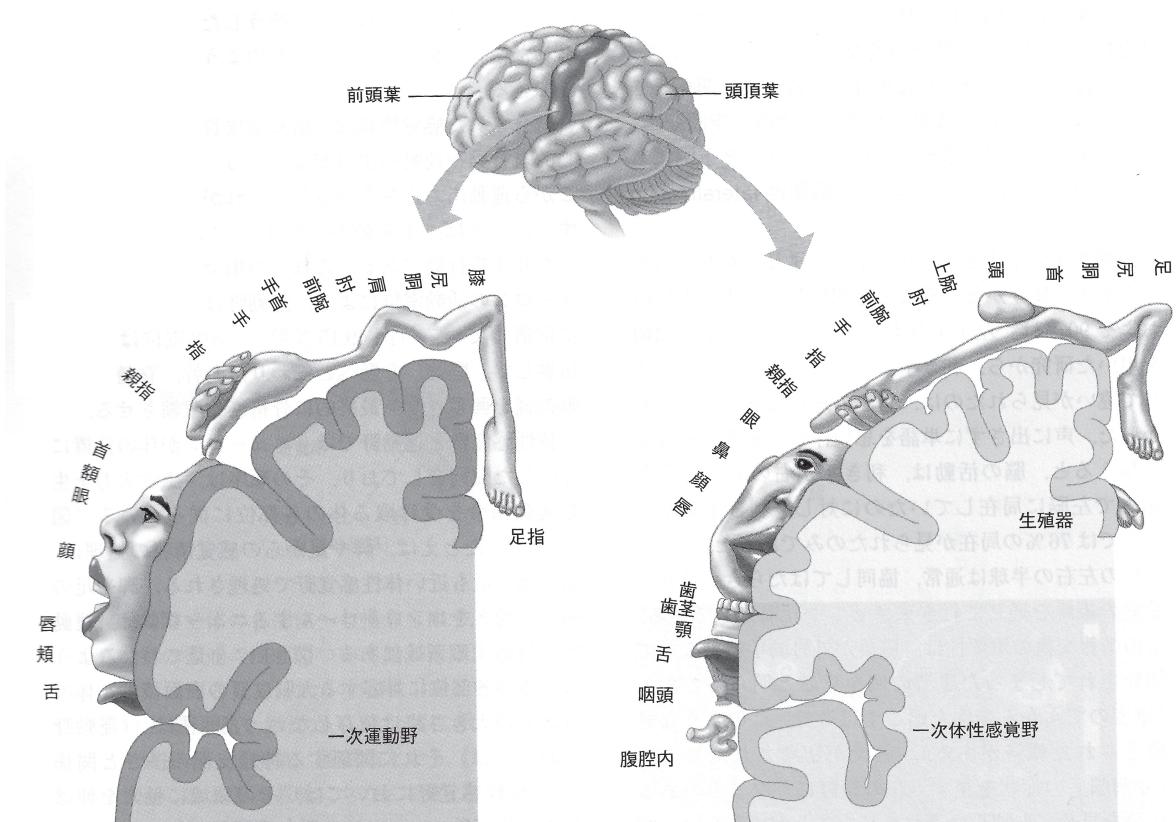
表 2-1 脳波の分類.

脳波の種類 (名称)	主な周波数範囲
δ 波	1～4Hz
θ 波	4～8Hz
μ 波	7～13Hz
α 波	8～13Hz
β 波	14～30Hz
γ 波	30Hz～

### 2.1.2 事象関連同期・事象関連脱同期

ある事象に対して脳波の特定の帯域のパワーが増幅/減少する現象を、事象関連同期（ERS）/脱同期（ERD）という。運動または運動想起時には事象関連脱同期が生じ、 $\mu$  帯域や $\beta$  帯域の活動が平常時より減衰することがわかっている。この現象は、その運動に対応する大脳皮質運動野にて発生する。（Reece et al., 2011）

図 2-2 に各体の部位に対応する運動野と体性感覚野の位置を示す。この図では、身体部位の面積の大きさが皮質表面の相当領域の面積に対応するよう書かれている。図より、手に対応する運動野の面積は広いことがわかる。また体の左右の運動は、脳の異なる半球で観測される。よって、左右の手の運動想起は ERD の観測が比較的容易であるため、運動想起識別の対象として用いられることが多い。



(Reece et al., 2011)

図 2-2 運動野と体性感覚野の体部位局在. 上中央図は運動野及び体性感覚野の位置を示す. また下図はそれぞれの皮質の断面図を示す.

### 2.1.3 Common spatial pattern

脳波には様々なノイズが含まれる。ノイズは BCI の識別精度を低下させる要因となるため、その影響を低減させるフィルタ開発が進められている。その中でも空間フィルタは、脳波から抽出する情報を制限することで特徴ベクトルに用いる次元を削減可能であるという長所を持っている。

Common spatial pattern (CSP) とは、空間フィルタの一種であり、現在 BCI での運動想起特徴抽出に最も広く適応されている手法である。CSP ではバンドパスフィルタ処理された信号において、あるクラスからの投影信号の分散を最大にしつつ、別のクラスでの分散は最小になるような空間フィルタを算出する。(Ramoser et al., 2000) ただし、事前処理にてバンドパスフィルタをかける際、運動想起に関連する周波数成分に対応した適切な周波数帯域を選ぶ必要があり、未だ最適な帯域決定方法は画一されてない。

## 2.2 人工ニューラルネットワーク

人工ニューラルネットワーク（以下、ANN）とは人間の脳の神経回路を模擬した情報処理システムの総称であり、単純な基準では判別しにくい複雑なパターンの認識機能に長けている。近年 ANN は多くの分野にてその適用性が研究されており、BCI へ応用する試みも多くなされている。本章では ANN の基本的な性質と仕組みについて説明する。

生物でのニューロンは、電気的な刺激をやりとりしながら情報を伝達している。ニューロンは他のニューロンから刺激を入力された時、その刺激が一定以上であれば次のニューロンに刺激を伝えるという仕組みを持つ。刺激のやりとりが多いニューロン同士はその結合も強くなり、情報伝達の優先度も高くなる。ANN ではこの仕組みを数学的モデルに置き換えたものである。

### 2.2.1 ANN の構造

ANN は構造によって幾つか種類があり、本研究では順伝播型全結合ニューラルネットワーク、畳み込みニューラルネットワーク (CNN)、再帰型ニューラルネットワークのひとつである Long short-term memory (LSTM) を用いた。

ここでは順伝播型全結合ニューラルネットワークを例に ANN の基本的な仕組みを説明する (CNN, LSTM については次章にて詳しく説明する)。図 2-3 に順伝播型全結合ニューラルネットワークを示す。ネットワークは入力層、中間層、出力層からなり、各層におかれた複数のユニットは前後の層のユニットと接続している。それぞれのユニットは図 2-4 に示すように前層の各ユニットから重みつき入力をうけ、その総和から閾値を引いた値を応答関数に代入することで出力値を決定する。つまり、 $x$  層第  $j$  ユニットへの入力の総和を  $I_j^x$ 、出力を  $O_j^x$ 、 $x-1$  層第  $i$  ユニットから  $x$  層第  $j$  ユニットへの結合の重みを  $W_{ij}^{x-1,x}$ 、 $x$  層第  $j$  ユニットの閾値を  $\theta_j^x$  とし、応答関数を  $f$  とおくと、 $I_j^x$  と  $O_j^x$  は (1), (2) 式で表せる。

$$I_j^x = \left( \sum_{i=0}^N W_{ij}^{x-1,x} O_i^x \right) - \theta_j^x \quad (1)$$

$$O_j^x = f(I_j^x) \quad (2)$$

ここで  $N$  は  $x-1$  層のユニット数とする。本研究では応答関数として (3) 式に定義されるシグモイド関数を用いた。

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (3)$$

ANN を特徴付けるパラメータは重みと閾値である。これらパラメータは、あらかじめ用意した入力層への入力データと目標となる出力層からの出力値からなるデータセットを用い、学習することで決定される。つまり学習作業では、ある入力に対してるべき出力パターンを再構成できるように ANN のパラメータを決めていく。

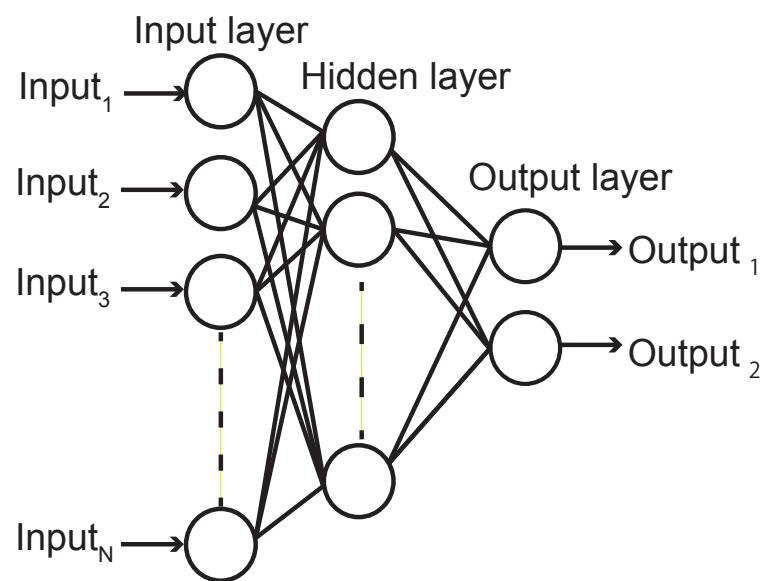


図 2-3 順伝播型全結合ニューラルネットワーク.

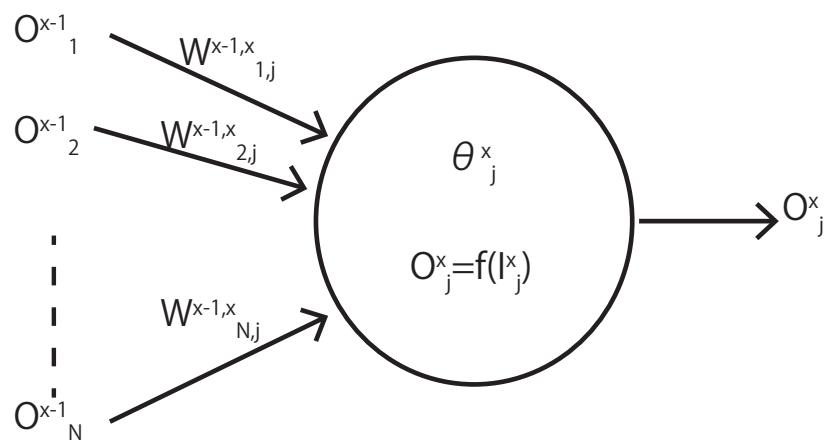


図 2-4 ANN 内のユニットにおける出入力.

## 2.2.2 誤差逆伝搬法

ANN の学習には誤差逆伝搬法（以下、BP）を用いた。BP では、まず ANN のパラメータの初期値をランダムに仮定し、学習データの入力に対する出力層からの出力  $O_j^x$  と教師信号  $Y_j^x$  との誤差評価尺度 E を計算する。本研究では誤差評価尺度式に交差エントロピー（(4) 式）を用いた。ここで  $N_x$  は  $x$  層におけるニューロン数を表す。

$$E = - \sum_{n=1}^{N_x} [Y_n^x \ln O_n^x + (1 - Y_n^x) \ln (1 - O_n^x)] \quad (4)$$

次にこの E の値が最小になるように勾配法を用いてパラメータの修正を行う。パラメータの修正方法についてユニットの重み  $W_{ij}^{k-1}$  の修正を例にして述べる。

E の  $W_{ij}^{x-1}$  における偏微分は

$$\frac{\partial E}{\partial W_{ij}^{x-1}} = \frac{\partial E}{\partial O_j^x} \cdot \frac{\partial O_j^x}{\partial I_j^x} \cdot \frac{\partial I_j^x}{\partial W_{ij}^{x-1}} \quad (5)$$

と表される。ここで

$$\frac{\partial E}{\partial O_j^x} = -\frac{Y_j^x}{O_j^x} + \frac{(1 - Y_j^x)}{(1 - O_j^x)} \quad (6)$$

$$\frac{\partial O_j^x}{\partial I_j^x} = \frac{\partial \sigma(I_j^x)}{\partial I_j^x} = \frac{\partial}{\partial I_j^x} \left\{ \frac{1}{1 + \exp(-I_j^x)} \right\} = \sigma(I_j^x) \{1 - \sigma(I_j^x)\} = O_j^x (1 - O_j^x) \quad (7)$$

$$\frac{\partial I_n}{\partial W_{ij}^{x-1}} = \frac{\partial}{\partial W_{ij}^{x-1}} \left\{ \left( \sum_{i=0}^N W_{ij}^{x-1,x} O_i^{x-1} \right) - \theta_j^x \right\} = O_j^{x-1} \quad (8)$$

より (5) 式は

$$\frac{\partial E}{\partial W_{ij}^{x-1}} = O_j^{x-1} (O_j^x - Y_j^x) \quad (9)$$

となる。すなわちユニットの重み  $W_{ij}^{x-1}$  の修正量は (10) 式で計算される。

$$\Delta W_{ij}^{x-1} = -\varepsilon (O_j^x - Y_j^x) O_j^x (1 - O_j^x) O_j^{x-1} \quad (10)$$

ここで  $\varepsilon$  は学習率であり、小さな正の数とする。また、収束特性を改善するためにモーメンタム確率的勾配降下法を用い、もう一つのパラメータ、すなわちモーメンタム  $\alpha$  と学習率減衰  $\beta$  を導入した。これより学習  $t+1$  回目の重み  $W^{t+1}$  は

$$W^{t+1} = W^t - \varepsilon (O_j^x - Y_j^x) O_j^x (1 - O_j^x) O_j^{x-1} + \alpha W^t \quad (11)$$

$$\varepsilon = \varepsilon_0 \frac{1}{(1 + \beta \cdot N_{epoch})} \quad (12)$$

で更新される。ここで  $\varepsilon_0$  は学習率の初期値、 $N_{epoch}$  はエポック数を示す。以上の計算を繰り返すことでパラメータの最適化を行った。

### 2.2.3 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク（以下、CNN）とは、全結合層の他に畳み込み層とプーリング層という特徴的な構造をもつANNの一種である。現在幅広い分野で活用されており、特に画像認識の分野で優れた性能を発揮している。

畳み込み層では入力データにフィルタを適用し、畳み込み演算をする。図2-5に3次元の入力画像 $I$ に対してフィルタ処理 $W$ を適用し、2次元畳み込みをした場合を例として示す。入力画像の大きさを $N \times M \times L$  ( $N \times M$ の画像 $L$ 枚)、フィルタサイズを $X \times Y \times L$ とし、 $I, W$ の各要素を $I_{nml}, W_{xyl}$ とする。畳み込み演算では、入力画像 $I$ の左上から右下へ順にフィルタ $W$ との内積をとる。ここで出力を $O$ とすると、その要素 $O_{ij}$ は(13)式で表される。

$$O_{ij} = \sum_{x=1}^X \sum_{y=1}^Y \sum_{l=1}^L W_{xyl} I_{(i+x-1)(j+y-1)l} + b \quad (13)$$

計算された出力 $O$ は特徴マップと呼ばれる。CNNは学習過程において、入力データを識別するのに効果的な特徴マップを出力できるフィルタ $W$ を探す。

プーリング層では、畳み込み層により得られた特徴マップの圧縮を行う。プーリング方法には幾つか種類があるが、本研究ではmaxプーリングを用いた。maxプーリングは区域内の最大値をその区域の代表値として取り出すことで特徴の圧縮を行う。

(図2-6)

CNNでは活性化関数に正規化線形関数(ReLU)を用いた。((14)式、図2-7) LeLUは $x$ が正のとき微分が常に1になるため、勾配消失が起きにくく、また計算が簡単で高速であることが特徴である。

$$ReLU(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases} \quad (14)$$

ANNを用いたクラス分類において、その精度を向上させるには、正確にタグ付けされた優れた学習データセットを各クラス十分な量用意する必要がある。しかし、多量のデータを集め、全てをマニュアルで識別しタグ付けしていく作業は大きな負担となる。CNNでの画像分類においては、ある程度タグ付けされたデータを元にデータを水増し(data augmentation)することで、識別精度向上に成功した例が多々ある。データの水増し手法としては、画像の左右を反転、回転、トリミング、コントラストを変化させる、ノイズを増やす、などといった方法があげられる。一方で水増しに使用した元データ数が少なすぎた場合には過学習が生じたり、そもそも元データの品質が悪かった場合には識別精度が低下してしまったりする場合もあり、注意が必要である。

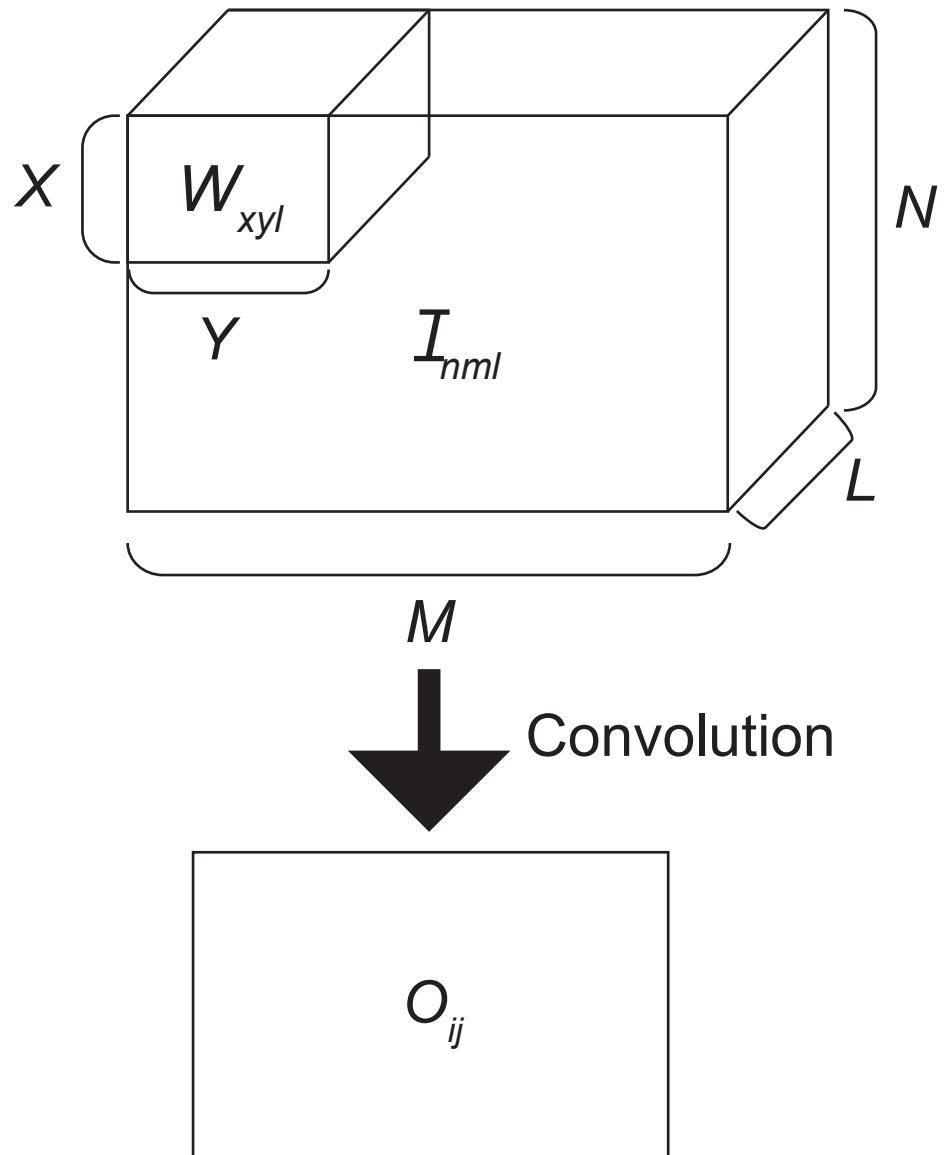


図 2-5 3 次元の入力画像に対する畠み込み処理.

## Max pooling

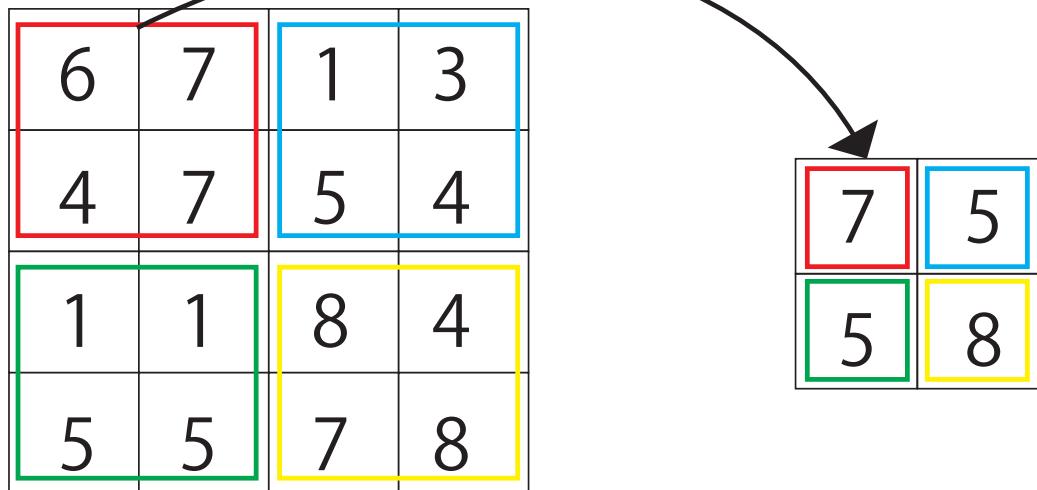


図 2-6 max プーリング処理.

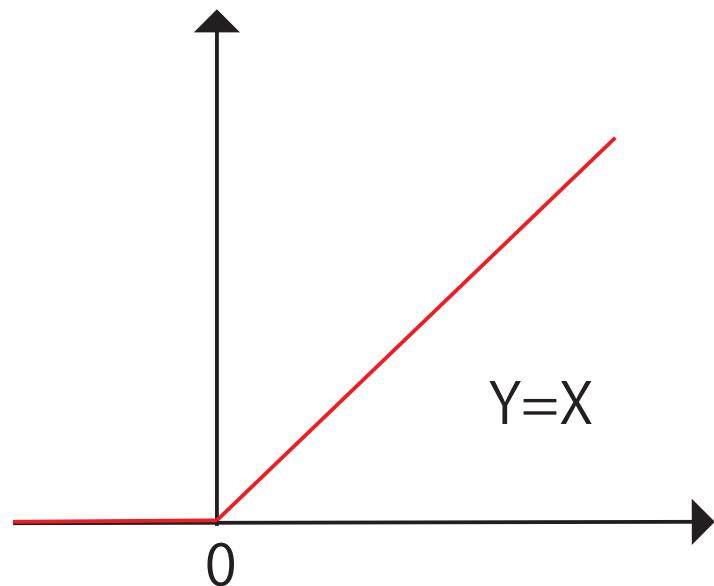


図 2-7 正規化線形関数 (ReLU).

## 2.2.4 Long short-term memory

Long short-term memory (LSTM) とは, Hochreiter (1997) が提案した, 時系列データを処理するために作られた再帰型ニューラルネットワーク (以下, RNN) モデルの一種である. 時系列データとは時間的変化をもつデータのこと, 脳波もその一種に属する. RNN は過去に入力された情報をネットワーク内に記憶し, 学習に利用することが可能である. LSTM は従来の RNN に比べ, 長期の時系列データでも学習可能である点が特徴であり, 近年自然言語処理や音声認識処理などの分野で優れた性能を発揮している. 本研究では Gers (2000) が提案した forget gate を導入した LSTM を用いた.

LSTM でのデータ処理の流れを図 2-8 に, また, それを時間経過で展開したもの を図 2-9 に示す. 時間  $t$  におけるデータを  $x_t$ , 入力に対する出力を  $h_t$  すると,  $h_t$  の算出過程において, 入力  $x_t$  の他に  $t-1$  での出力  $h_{t-1}$  も寄与している. これは, RNN に共通する仕組みであるが, LSTM では LSTM ブロック内に幾つかのゲートを持つことで長期記憶を可能としている. ゲートではニューロン内の情報を選択的に通す機能を有する.

LSTM ブロック内の構造を図 2-10 に示す. まず出力ゲートでは原始的な RNN の機能をもっており, 出力値  $o_t$  は (15) 式で表される.

$$o_t = \sigma(W_o x_t + R_o h_{t-1} + b_o) \quad (15)$$

$W$  は重み係数,  $b$  は閾値,  $\sigma(x)$  はシグモイド関数を表す. また, 忘却ゲート  $f_t$  及び出力ゲート  $i_t$  でも同様の処理がなされる.

$$f_t = \sigma(W_f x_t + R_f h_{t-1} + b_f) \quad (16)$$

$$i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i) \quad (17)$$

入力ゲートでは式(17)と並行して, 式(18)の処理もなされる.

$$z_t = \tanh(W_z x_t + R_z h_{t-1} + b_z) \quad (18)$$

ここで, 忘却ゲートと入力ゲートは, メモリセルを通じて互いに影響し合う関係にあり,  $f_t$ ,  $i_t$ ,  $z_t$  を用いて以下の処理が行われる.

$$c_t = i_t \times z_t + c_{t-1} \times f_t \quad (19)$$

ここでメモリセルは,  $c_t$  という値を次の計算に渡す, もしくは  $c_{t-1}$  という値を保持するという役割を持っているだけで, 特別な計算は行わない. 最後に LSTM ブロックから得られる出力  $h_t$  は (20) 式より得られる.

$$h_t = o_t \times \tanh(c_t) \quad (20)$$

この値は次のデータが入力された時の計算にも用いられる.

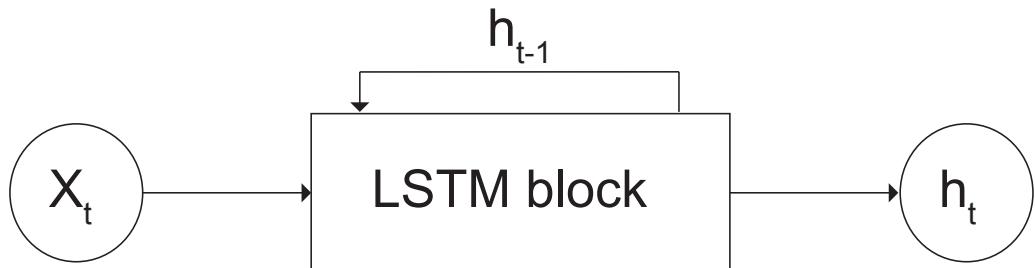


図 2-8 LSTM におけるデータ処理の流れ.

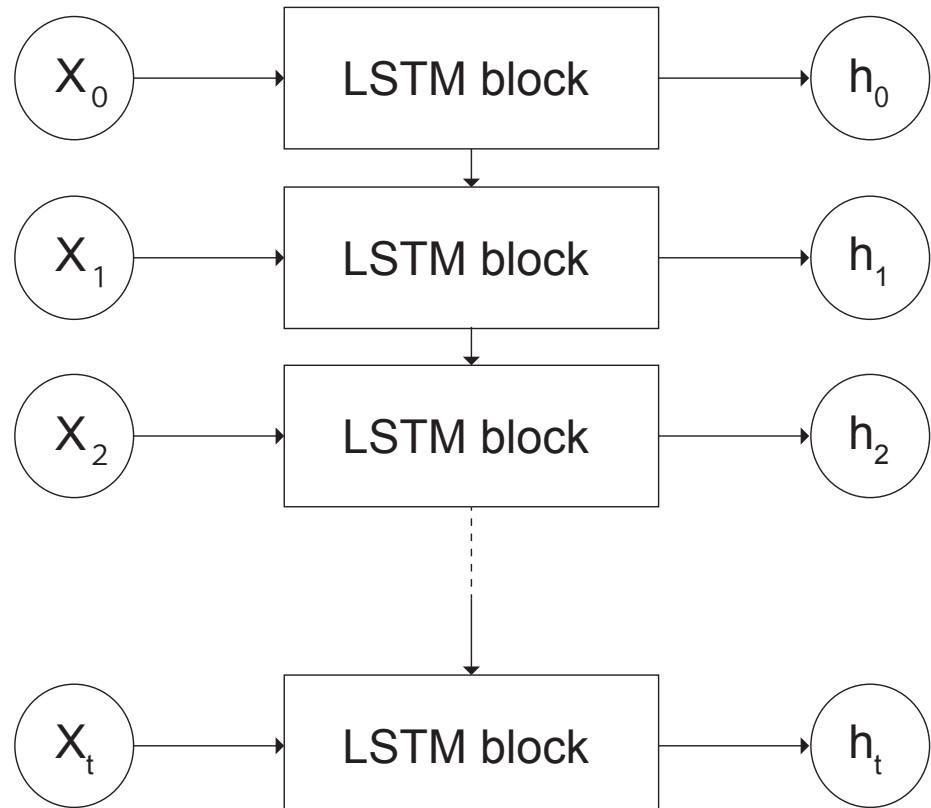


図 2-9 LSTM における、各時間の入力に対するデータ処理の流れ.

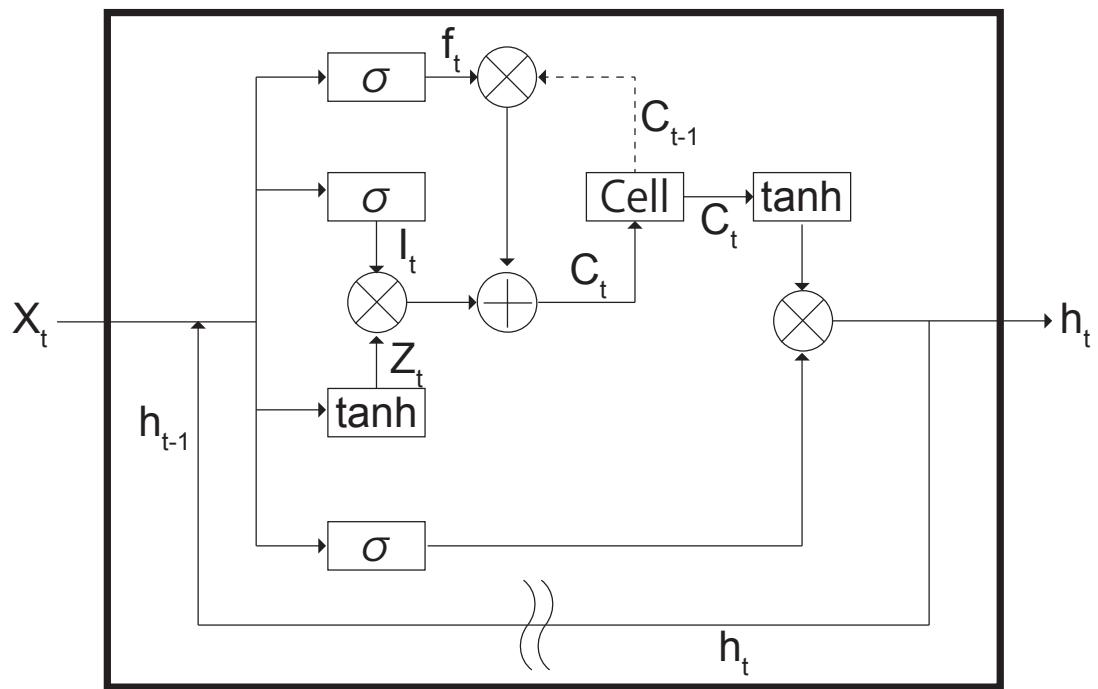


図 2-10 LSTM ブロック内の構造.

### 2.3 経験的モード分解

脳波は非線形かつ非定常な波形であるため、フーリエ変換や wavelet 変換といった手法では、脳波を分解し表現することに限界がある。本研究では人工脳波作成の際、経験的モード分解 (EMD) を用いた。EMD とは時系列信号を固有モード関数 (IMF) の和に分解する手法のことである (Huang et al., 1998)。つまり、固有モード関数を  $IMF(t)$ 、トレンド成分  $r_n(t)$  をすると、時系列信号  $x(t)$  は (21) 式として表される。

$$x(t) = \sum_{k=1}^n IMF_k(t) + r_n(t) \quad (21)$$

IMF は (1) 信号の極大値の数とゼロ交差点の数の差が 1 以下であり、(2) 極大値の包絡線と極小値の包絡線の平均がゼロである、という条件をもった関数である。

### 3 ANNによる運動想起判別手法

#### 3.1 使用データ

本研究では、Phan (2010) が行った左手右手の運動想起時のデータを用いた。これは2名 (subject 1, subject 2) の被験者の頭に62個の電極をつけ、左右各2秒間想起した時の脳波を100回測定し、8-30Hzのバンドパスフィルタにかけたものである。図3-1は電極チャンネルの位置を示す。サンプリング周波数は500Hzであるため、1施行分のデータは $1000\ (500\text{Hz} \times 2\text{秒}) \times 62$ の行列として表される。

本研究では上記200個のデータのうち、ANNの学習に60個、学習結果の検証に140個のデータを使用した。これらの脳波は、以降作成する人工脳波と区別するため、実測脳波と呼ぶことにする。

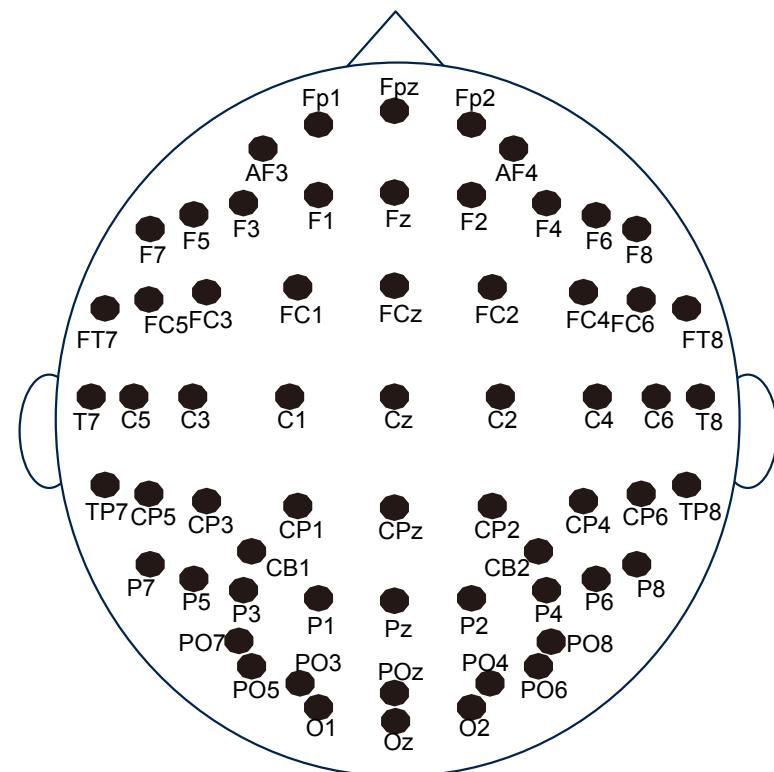


図 3-1 電極チャンネル位置。

### 3.2 人工脳波作成

実測脳波から人工脳波の作成には Dinares-Ferran (2018) の手法を用いた。まず実測脳波信号 $EEG(t)$ を EMD により固有モード関数 $IMF(t)$ とトレンド成分 $r_n(t)$ に分解する。信号の分解方法を以下に示す。

- (1)  $s(t) = r_{i-1}(t)$ とする。ここで  $i$  の初期値は 1 であり,  $r_0(t) = EEG(t)$ とする。
- (2)  $s(t)$ において、全ての極大値を内挿する上側包絡線と全ての極小値を内挿する下側包絡線を、極大値と極小値のスプライン補間から算出する。
- (3) 上側包絡線と下側包絡線の平均値を $m(t)$ とする。
- (4)  $h(t) = s(t) - m(t)$ から IMF の候補を得る。
- (5)  $h(t)$ が IMF の条件を満たしていない場合、 $s(t) = h(t)$ とし、(2) へ戻る。
- (6)  $h(t)$ が IMF の条件を満たしている場合、 $IMF_i(t) = h(t)$ を得る。
- (7)  $r_i(t) = r_{i-1}(t) - IMF_i(t)$ とする。
- (8)  $r_i(t)$ が単調関数、もしくは包絡線を計算するのに十分な極値を持たない場合、分解を終了する。
- (9) (8) 以外の場合、 $s(t) = r_i(t)$ とし (2) へ戻り、ループを繰り返すことで $IMF_{i+1}(t)$ を得る。

本研究では、(5)において、 $10 \times \log_{10} \left[ \frac{s(t) \times s(t)}{m(t) \times m(t)} \right]$  が 40dBに達するまでループを繰り

返した。また (8) での分解終了条件は、 $10 \times \log_{10} \left[ \frac{r_{i-1}(t) \times r_{i-1}(t)}{r_i(t) \times r_i(t)} \right] < 50$ dBかつ極大値と極小値の数が 2 以下、とした。これより、 $EEG(t)$ は (22) 式と表される。

$$EEG(t) = \sum_{k=1}^n IMF_k(t) + r_n(t) \quad (22)$$

同クラス（左手もしくは右手）に属する脳波から得られた IMF をランダムに組み合わせ直すことで、人工脳波を作成した。図 3-2 に組み合わせ方を示す。脳波データ $EEG_1$ を EMD したときに得られた IMF を $IMF_{1-k}$ とすると、j 番目の IMF である $IMF_{1-j}$ は他の脳波データ $EEG_x$ から分解された j 番目の IMF,  $IMF_{x-j}$ に入れ替えることで IMF の組み替えがなされる。脳波は約 8~10 個の IMF に分解できるため、本研究では  $k=2 \sim 5$  の IMF を組み合わせ直し、人工脳波を作成した。

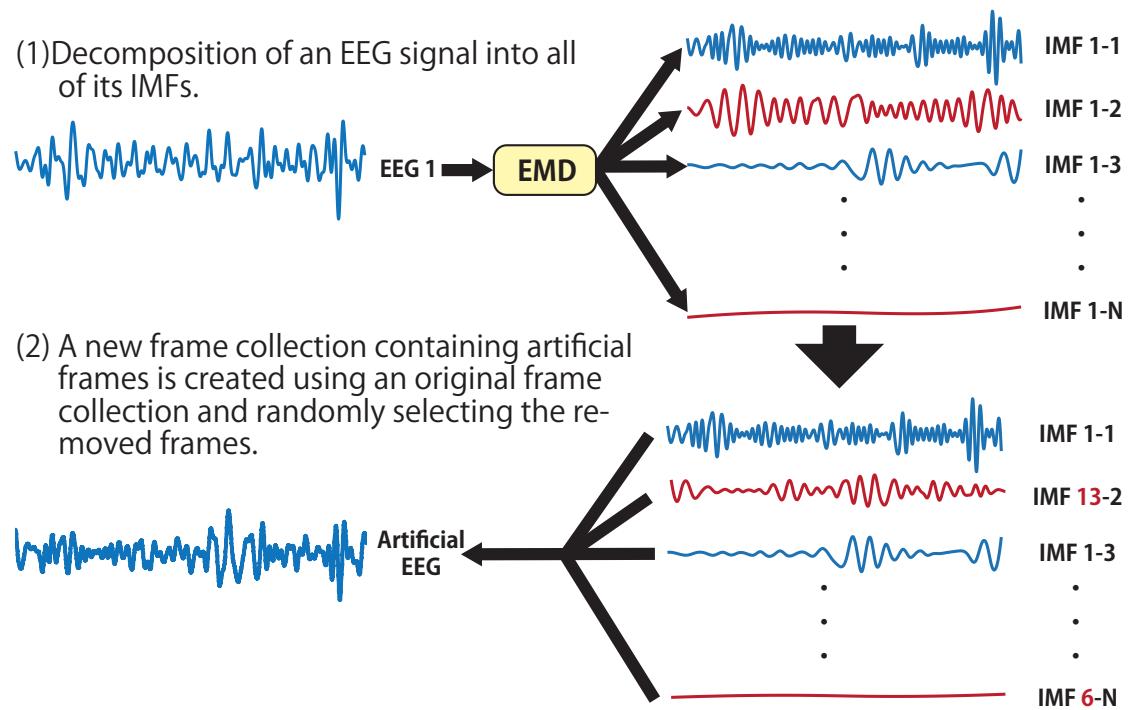


図 3-2 人工脳波作成時における、IMF の組み合わせ方。

### 3.3 時間周波数解析

脳波は周波数帯域や出現部位の違いにより脳活動の種類を識別することができるところから、時間周波数解析は重要な脳波解析手段のひとつである。時間周波数解析にはおもに短時間フーリエ変換や wavelet 変換が用いられる。

短時間フーリエ変換では、時系列データから窓関数を用いて時間をずらしながら一定の区間を切り出し、フーリエ変換を行うことで、周波数スペクトルの時間的変化を見る手法である。しかし不確定性の原理より、周波数分解能は窓関数のサイズに反比例するため、周波数と時間の両者の分解能を両立することはできない。

そこで 1980 年代初頭に Morlet は Wavelet 変換を考案した。wavelet 変換では mother wavelet と呼ばれる基本関数のサイズを、低周波数解析に対しては拡大、高周波数解析に対しては縮小させることで、低周波領域では周波数の分解能を優先し、高周波領域では時間の分解能を優先するように変換可能とした。よって広い周波数領域において時間と周波数にかかる信号情報を同時に抽出することが可能となった。wavelet 変換の式を (23) 式に示す。

$$W_\psi(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t) \cdot \psi^* \left( \frac{t-b}{a} \right) dt \quad (23)$$

$b$  は中心座標,  $a$  はスケールに関連したパラメータである。mother wavelet  $\psi$  には Morlet wavelet を用いた。((24) 式)

$$\psi_{Morlet} = \frac{1}{\sqrt{\pi f_b}} \exp(2\pi i f_c t) \cdot \exp \left( \frac{-t^2}{f_b} \right) \quad (24)$$

$f_c$  は中央周波数、 $f_b$  はバンドワイドパラメータである。Morlet wavelet の波形は滑らかで生体反応として現れる波形に似ていることから、脳波処理時によく利用される。

### 3.4 ANN での処理

本研究では、2種類の ANN を構築した。ひとつは CNN を用いて、脳波のスカラグラムを画像認識技術で判別する手法である。もう一方は、脳波のスペクトル強度の時間的变化の特徴を LSTM で判別する手法である。プログラミングには Python3 を使用し、Keras ライブラリーを用いた。以下にそれぞれの詳細を述べる。

### 3.4.1 CNN の構造

本研究で検証した CNN の構造を図 3-3 に示す。この構造は Kwon (2018) の CNN を用いた EEG による感情識別にて用いられた CNN のモデルを参考に作成した。

CNN へデータを入力するために、短時間フーリエ変換もしくは wavelet 変換を用いて脳波を時間周波数解析し画像化した。図 3-4 に脳波を変換した例を示す。これより、脳波はチャンネル毎にスケール 8~30 Hz, 時間 0~2 秒のスペクトル画像となる。即ち脳波データを電極 62 チャンネル毎の 1 次元データ ( $1000 \times 62$ ) から、周波数と時刻における信号の強さを表す 62 枚の 2 次元画像 ( $23 \times 50 \times 62$ ) に変換し、CNN へ入力した。CNN の各層における出力データのサイズとパラメータ数を表 3-1 に示す。CNN では畳み込み層、ReLU 層、Max プーリング層のセットが 2 セット組み込まれている。Kwon (2018) の CNN では 4 層の畳み込み層が組み込まれていたが、本研究では学習に用いるデータ数の少なさと、計算コストを考慮し、より単純なネットワークを構築した。畳み込み層でのカーネルサイズは  $3 \times 5$ , ストライドは  $1 \times 1$  とし、フィルタの枚数は 100 枚とした。またプーリング層ではカーネルサイズは  $3 \times 3$ , ストライドは  $2 \times 2$  とした。

2 回目のプーリング層から出力された配列は 1 次元に変換され、2 層の全結合層へ入力されたのち、最終的に 2 つの出力値が得られる。学習過程ではこの出力値と正解ラベル (左手 : [1, 0], 右手 : [0, 1]) との誤差が小さくなるよう学習データを用いてネットワーク内のパラメータを最適化した。判別精度の検証では、学習済みのネットワークに検証データを入力した際の出力値から、入力した脳波の判別を行い、その精度を評価した。

CNN の判別能力検証ではまず、60 個の実測脳波を用いて CNN の学習を行なった。ここでの脳波画像化手法には、短時間フーリエ変換および、wavelet 変換を用いた。wavelet 変換では、mother wavelet のバンドワイドパラメータ  $b=1, 2, \dots, 6$  の計 6 種類を試した。

次に実測脳波 30 個から人工脳波を 20 個、30 個作成し、学習データ数を合計 50, 60 個とさせ、CNN の学習を行った。

最後に、実測脳波 60 個から人工脳波を 120, 300, 600 個作成し、学習データ数を合計 180, 360, 660 個とさせ、CNN の学習を行った。

学習し終わった CNN を用いて検証用脳波 140 個の判別を行い、その正答率を比較すること判別能力を検証した。これら検証は subject1,2 それぞれで行った。

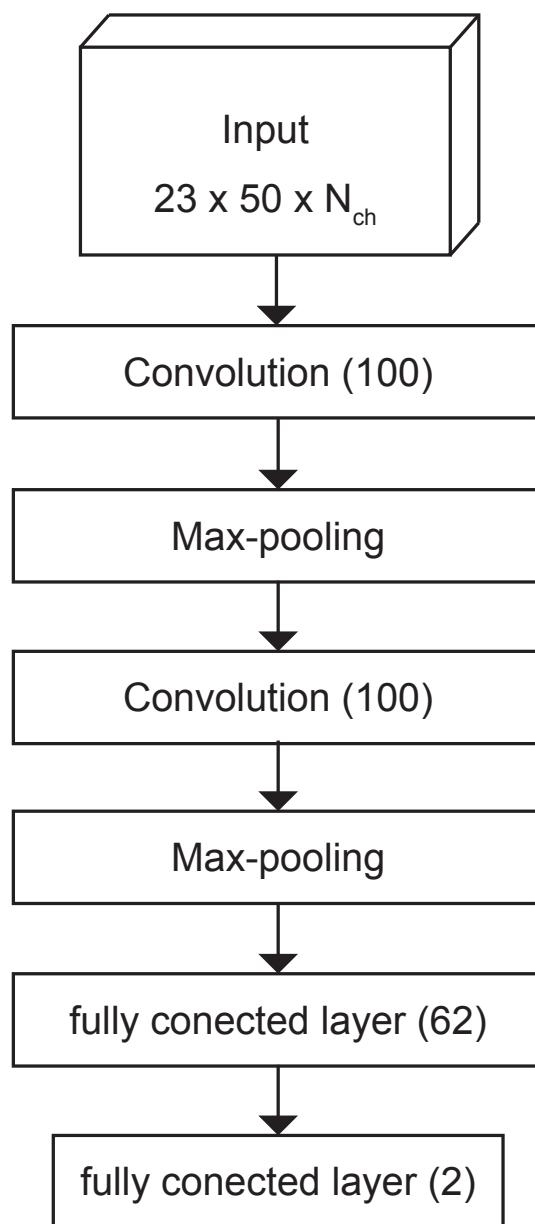


図 3-3 本研究で用いた CNN の構造.

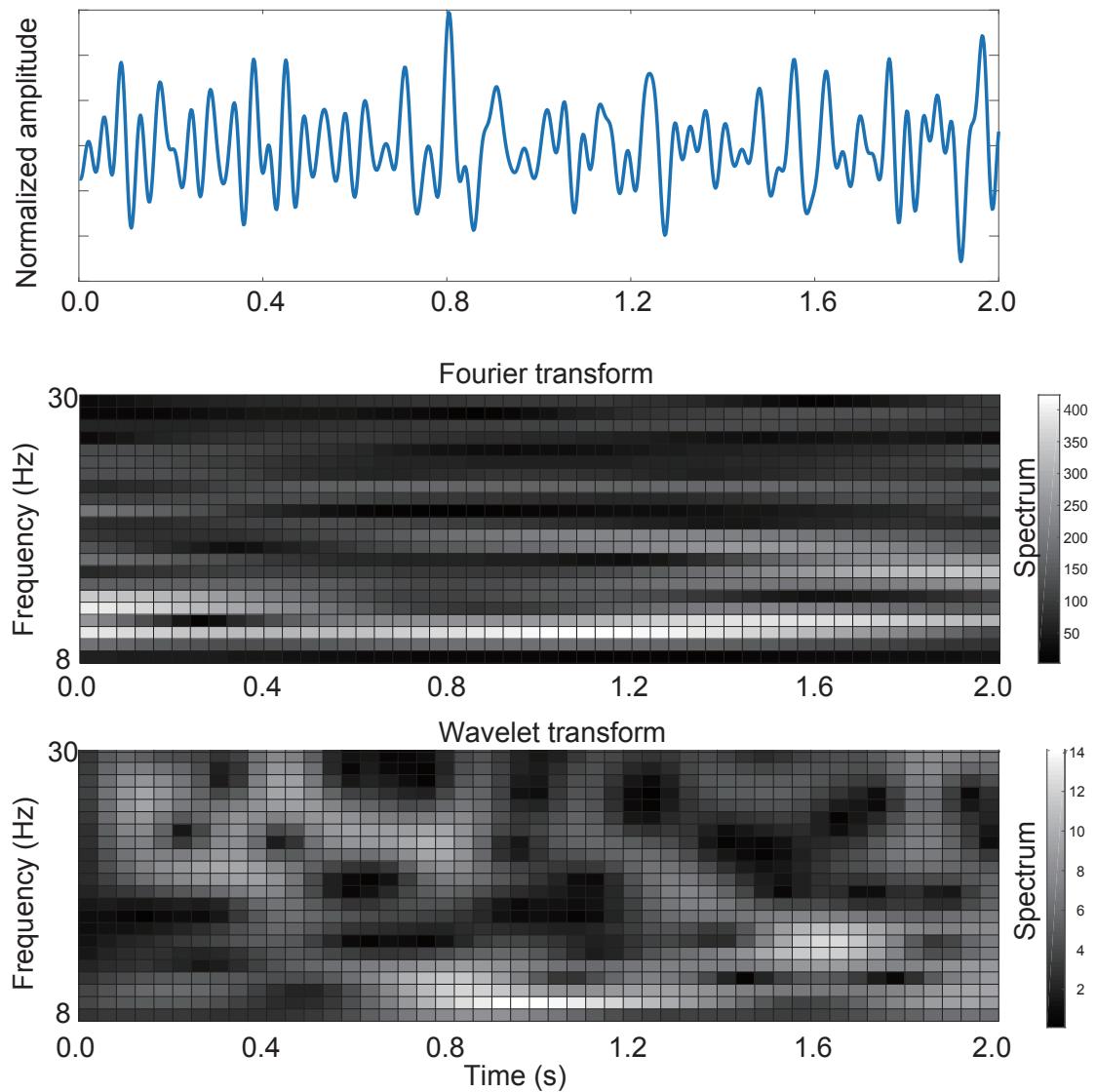


図 3-4 脳波のフーリエ変換及び wavelet 変換例。

表 3-1 CNN の各層における出力データのサイズとパラメータ数.

Layer (type)	Output Shape	Number of parameter
conv2d_1 (Conv2D)	(23, 50, 100)	93100
activation_1 (Activation)	(23, 50, 100)	0
max_pooling2d_1 (MaxPooling2)	(11, 24, 100)	0
conv2d_2 (Conv2D)	(9, 20, 100)	150100
activation_2 (Activation)	(9, 20, 100)	0
max_pooling2d_2 (MaxPooling2)	(4, 9, 100)	0
flatten_1 (Flatten)	(3600)	0
dense_1 (Dense)	(62)	223262
activation_3 (Activation)	(62)	0
dense_2 (Dense)	(2)	126

### 3.4.2 CNN+LSTM の構造

脳波から運動想起を識別する際、脳波の時間変化と、その変化が生じる場所が重要な情報となる。本研究では、CNNで位置に関する特徴を抽出し、LSTMで時間に関する特徴を認識できないかと考えた。そこで図 3-5 に示した CNN+LSTM を構築した。

CNN+LSTM では図 3-6 に示した 45 チャンネルのデータを使用した。チャンネルの位置情報を入力データに含めるため、データの並べ替えを行った。脳波は CNN の場合と同様 wavelet 変換した後 ( $23 \times 50 \times 45$ )、時間毎に図 3-6 に示したチャンネル順に並べ替え、 $5 \times 9$  の画像を 23 枚、50 組作成した ( $5 \times 9 \times 23 \times 50$ )。 $5 \times 9$  の画像の例を図 3-7 に示す。CNN+LSTM では、入力データを時間毎に畳み込み処理した後、LSTM を適用した。その後、全結合層へデータが入力され、CNN の場合と同様、最終的に出力された 2 つの値と正解ラベルとの誤差が小さくなるよう学習データを用いてネットワーク内のパラメータを最適化した。判別精度の検証では、学習済みのネットワークに検証データを入力し、その際の出力値から入力した脳波の判別を行い、その精度を評価した。

CNN+LSTM の判別能力検証では、まず実測脳波 60 個から人工脳波を 120, 300, 600 個作成し、学習データ数を合計 180, 360、660 個とさせ、CNN+LSTM の学習を行った。次に実測データ 6, 12, 30 個から学習データ数が合計 60 個になるよう、人工脳波を作成し、CNN+LSTM の学習を行った。これら検証は subject1,2 それぞれで行った。

また、subject1においては、実測データ 6, 12, 30 個から学習データ数が合計 50 個になるよう、人工脳波を作成し CNN+LSTM の学習を行った。

さらに subject2 においては実測脳波 36 個から学習データ数が合計 60 個になるよう、人工脳波を作成し、CNN+LSTM を学習させた場合も検証した。

CNN の場合と同様に、学習し終わった CNN +LSTM を用いて検証用脳波 140 個の識別を行い、その正答率を比較すること識別能力を検証した。

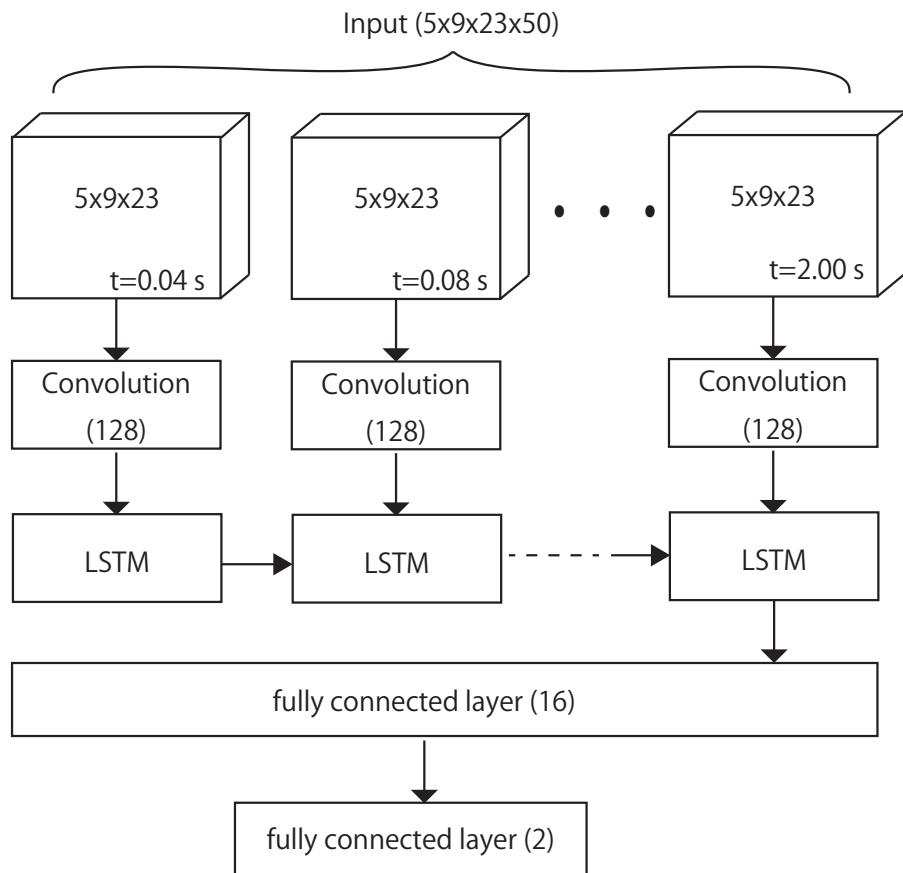


図 3-5 本研究で用いた CNN+LSTM の構造.

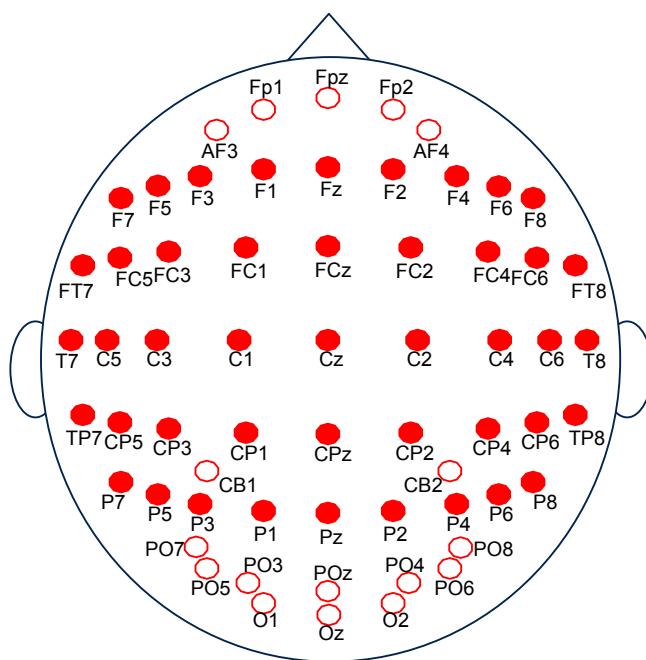


図 3-6 CNN+LSTM にて使用したデータのチャンネル位置. ●で示された 5 行 9 列, 計 45 個のチャンネルから得られた脳波データを使用した.

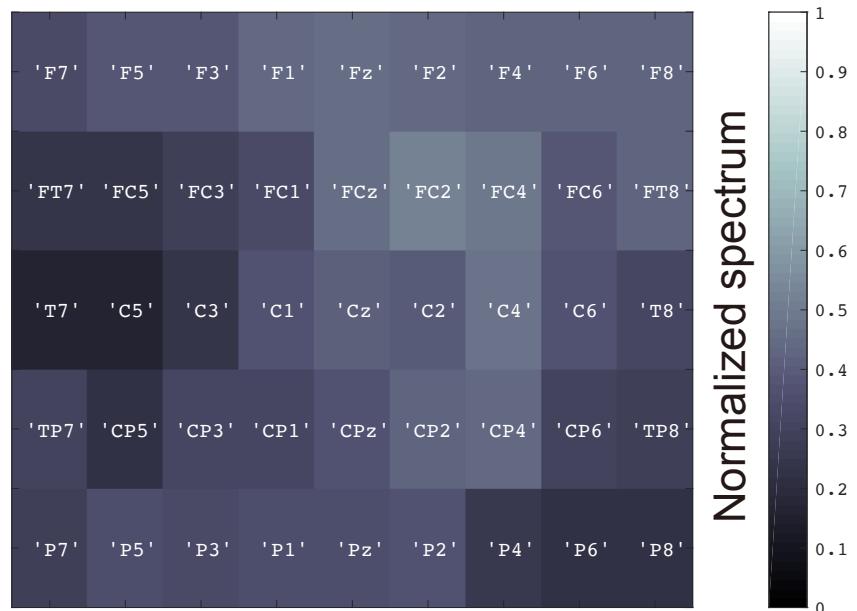


図 3-7 CNN+LSTM での入力画像の例. 白文字は対応するチャンネル名を示す.

表 3-2 CNN の各層における出力データのサイズとパラメータ数.

Layer (type)	Output Shape	Number of parameter
time_distributed_1 (Time Distributed)	(50, 7, 3, 128)	26624
time_distributed_2 (Time Distributed)	(50, 2688)	0
lstm_1 (LSTM)	(256)	3015680
dense_1 (Dense)	(16)	4112
activation_1 (Activation)	(16)	0
dense_2 (Dense)	(2)	34
activation_2 (Activation)	(2)	0

## 4 結果

### 4.1 Morlet mother wavelet のパラメータ比較

入力データに短時間フーリエ変換をかけた場合と,  $fb=1\sim 6$  の wavelet 変換をかけた場合の CNN による識別精度を比較する. ANN による学習は条件ごとに 3 回ずつ行い, 評価にはその平均値を用いた. 表 4-1, 表 4-2 にその結果を示す. まず, 短時間フーリエ変換と wavelet 変換での識別精度を比較すると, wavelet 変換の方が subject1 で最大 5%, subject 2 で 1%高い. 故に wavelet 変換は短時間フーリエ変換より脳波のスペクトル解析に有用と考えられる. 以降の解析では入力データの事前処理に wavelet 変換を用いた. また, CNN の識別制度はパラメータ  $fb$  値にほとんど依存せず, 一定の識別精度が得られることがわかった. 以降, パラメータ  $fb$  は 6 として解析した.

表 4-1 wavelet 変換でのパラメータ  $fb$  に対する判別精度. (Subject 1)

	Accuracy of each trial			Average
	#1	#2	#3	
$fb=1$	87.9%	88.6%	88.6%	88.4%
$fb=2$	88.6%	89.3%	88.6%	88.8%
$fb=3$	88.6%	86.4%	88.6%	87.9%
$fb=4$	90.0%	88.6%	85.0%	87.9%
$fb=5$	90.0%	89.3%	87.1%	88.8%
$fb=6$	88.6%	87.1%	88.6%	88.1%
STFT	81.4%	83.6%	86.4%	83.8%

表 4-2 wavelet 変換でのパラメータ  $fb$  に対する判別精度. (Subject 2)

	Accuracy of each trial			Average
	#1	#2	#3	
$fb=1$	85.0	88.6	85.7	86.4
$fb=2$	85.7	87.1	85.0	85.9
$fb=3$	87.1	85.0	87.9	86.7
$fb=4$	87.9	88.6	85.7	87.4
$fb=5$	86.4	88.6	85.7	86.9
$fb=6$	85.7	87.1	87.1	86.7
STFT	86.4	85.7	87.1	86.4

## 4.2 CNN および CNN+LSTM による識別結果

以下に実測データ及び人口データでNNを学習した時の判別精度を示す。表 4-3, 表 4-4 では CNN による判別, 表 4-5, 表 4-6 には CNN+LSTM による判別結果を示す。各表には学習で用いた実測脳波データ数とそれらから作成された人工脳波データ数に対する識別結果が示されたおり, 各条件に対して ANN による学習は 4 回ずつ行った。また 4 回の結果からその平均, 標準偏差, 標準誤差を求め, 評価には平均値を用いた。

表 4-3 CNN による運動想起識別結果 (Subject 1).

Number of training data			Accuracy of each trial				Average	SD	SE
Real data	Artificial data	Total	#1	#2	#3	#4			
30	0	30	80.00%	80.71%	82.14%	78.57%	80.36%	1.29%	0.64%
30	20	50	83.57%	78.57%	82.14%	80.71%	81.25%	1.85%	0.92%
30	30	60	84.29%	85.71%	80.71%	83.57%	83.57%	1.82%	0.91%
60	0	60	88.57%	87.14%	88.57%	87.86%	88.04%	0.59%	0.30%
60	120	180	87.86%	89.29%	90.00%	88.57%	88.93%	0.80%	0.40%
60	300	360	87.86%	87.86%	87.86%	85.00%	87.14%	1.24%	0.62%
60	600	660	90.00%	90.00%	90.00%	87.86%	89.46%	0.93%	0.46%

表 4-4 CNN による運動想起識別結果 (Subject2).

Number of training data			Accuracy of each trial				Average	SD	SE
Real data	Artificial data	Total	#1	#2	#3	#4			
30	0	30	72.14%	70.71%	70.00%	70.00%	70.71%	0.87%	0.44%
30	20	50	76.43%	75.51%	75.71%	75.00%	75.66%	0.51%	0.26%
30	30	60	75.00%	72.86%	74.29%	72.86%	73.75%	0.93%	0.46%
60	0	60	85.71%	87.14%	87.14%	85.71%	86.43%	0.71%	0.36%
60	120	180	85.00%	87.14%	86.43%	85.71%	86.07%	0.80%	0.40%
60	300	360	87.14%	88.57%	87.86%	88.57%	88.04%	0.59%	0.30%
60	600	660	87.86%	88.57%	87.86%	88.57%	88.21% (a)	0.36%	0.18%

表 4-5 CNN+LSTM による運動想起識別結果 (Subject1).

Number of training data			Accuracy of each trial				Average	SD	SE
Real data	Artificial data	Total	#1	#2	#3	#4			
6	0	6	65.71%	67.14%	62.14%	64.29%	64.82%	1.85%	0.92%
12	0	12	75.00%	74.29%	78.57%	77.14%	76.25%	1.70%	0.85%
20	0	20	81.43%	82.14%	82.85%	78.57%	81.25%	1.63%	0.81%
30	0	30	81.43%	78.57%	81.43%	82.14%	80.89%	1.37%	0.69%
40	0	40	83.57%	83.57%	85.00%	85.00%	84.29%	0.71%	0.36%
6	44	50	65.71%	68.57%	65.71%	56.43%	64.11%	4.58%	2.29%
12	38	50	75.71%	78.57%	76.43%	77.14%	76.96%	1.06%	0.53%
30	20	50	85.00%	83.57%	83.57%	82.86%	83.75%	0.78%	0.39%
6	56	60	67.14%	63.57%	57.14%	59.29%	61.79%	3.86%	1.93%
12	48	60	78.57%	76.43%	78.57%	80.71%	78.57%	1.52%	0.76%
20	40	60	84.29%	82.86%	84.29%	80.71%	83.04%	1.46%	0.73%
30	30	60	85.71%	85.71%	86.43%	85.71%	85.89%	0.31%	0.15%
40	20	60	86.43%	86.43%	85.71%	86.43%	86.25%	0.31%	0.15%
60	0	60	86.43%	86.43%	85.71%	85.71%	86.07%	0.36%	0.18%
60	120	180	85.00%	87.86%	87.14%	86.43%	86.61%	1.06%	0.53%
60	300	360	90.71%	89.29%	88.57%	87.14%	88.93%	1.29%	0.64%
60	600	660	86.43%	85.71%	88.57%	88.57%	87.32%	1.28%	0.64%

表 4-6 CNN+LSTM による運動想起識別結果 (Subject 2).

Number of training data			Accuracy of each trial				Average	SD	SE
Real data	Artificial data	Total	#1	#2	#3	#4			
6	0	6	83.57%	83.57%	86.43%	86.43%	85.00%	1.43%	0.71%
12	0	12	85.71%	85.00%	86.43%	83.57%	85.18%	1.06%	0.53%
20	0	20	84.29%	82.14%	86.43%	82.14%	83.75%	1.78%	0.89%
30	0	30	87.86%	86.43%	85.71%	86.43%	86.61%	0.78%	0.39%
36	0	36	90.71%	88.57%	87.86%	88.57%	88.93%	1.07%	0.54%
40	0	40	89.29%	88.57%	90.00%	90.00%	89.46%	0.59%	0.30%
30	20	50	87.14%	87.86%	87.14%	90.00%	88.04%	1.17%	0.59%
6	54	60	83.57%	78.57%	85.71%	85.00%	83.21%	2.79%	1.39%
12	48	60	90.00%	86.43%	87.14%	87.14%	87.68%	1.37%	0.69%
20	40	60	85.71%	87.14%	85.00%	83.57%	85.36%	1.29%	0.64%
30	30	60	88.57%	85.00%	86.43%	88.57%	87.14%	1.52%	0.76%
36	24	60	87.14%	91.43%	90.00%	89.29%	89.46%	1.55%	0.77%
40	20	60	89.29%	90.00%	90.00%	89.29%	89.64%	0.36%	0.18%
60	0	60	89.29%	90.71%	91.43%	90.00%	90.36%	0.80%	0.40%
60	120	180	90.71%	89.29%	90.71%	90.71%	90.36%	0.62%	0.31%
60	300	360	90.00%	89.29%	90.71%	92.14%	90.54%	1.06%	0.53%
60	600	660	87.86%	84.29%	86.43%	86.43%	86.25%	1.28%	0.64%

## 5 考察

### 5.1 従来の手法との比較

60 個の実測脳波データセットを用いた CNN 及び CNN+LSTM による運動想起判別の結果と、従来手法である CSP による判別結果 (Phan and Cichocki , 2010) の比較を表 5-1 に示す。その結果 CNNにおいては、subject1 では 5.18%の精度向上がみられた一方、subject 2 では 3.57%精度が低下した。しかし表 4-4 (a) より実測データ 60 個と人工データ 600 個を組み合わせて CNN を学習させた場合、識別精度が 88.21%と向上することから、subject2においても人工データを用いることで従来手法とほぼ同程度の識別能力が得られる。また CNN+LSTM においては、subject1 では 3.21%，subject2 では 0.36%の精度向上が見られた。以上より、本研究が提案する CNN と CNN+LSTM による運動想起識手法は従来手法と同程度もしくはより高い識別精度が得られるが、精度の向上率は被験者によって異なることがわかった。

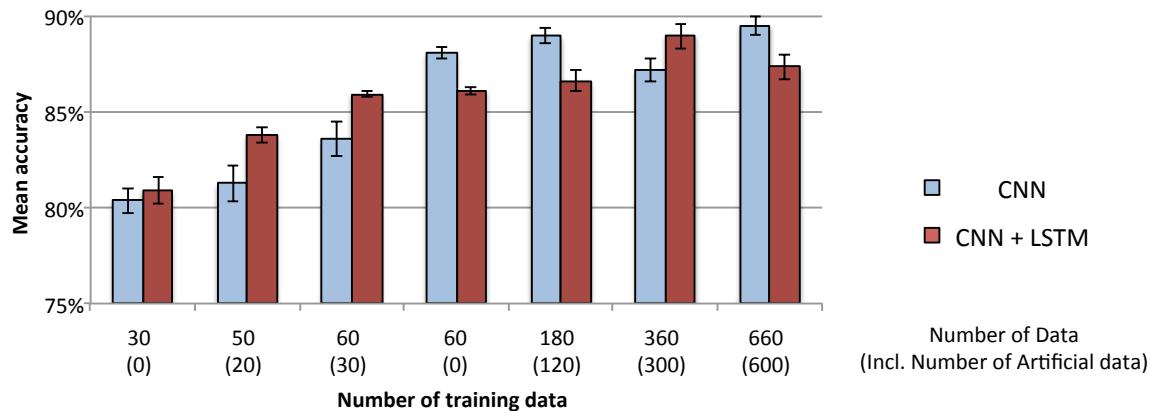
表 5-1 従来手法 (CSP) との比較。

Method	Accuracy	
	Subject 1	Subject 2
CSP	82.86%	90.00%
CNN	88.04%	86.43%
CNN+LSTM	86.07%	90.36%

## 5.2 CNN と CNN+LSTM の比較

図 5-1 に subject1 及び subject2 における, CNN と CNN+LSTM の学習データ条件が同じであるときの判別精度の比較を示す. subject1において, 学習で用いる実測データ数が 30 個の時は CNN より CNN+LSTM の方が高い判別精度示した. 一方, 学習に用いる実測データ数 60 個の時は, CNN は CNN+LSTM より精度が高い, もしくはほぼ同等と言える. subject2 においても, 学習で用いる実測データ数が 30 個の時は CNN より CNN+LSTM の方が明らかに高い判別精度示したが, 学習に用いる実測データ数が 60 個の時は, その差は小さくなり, CNN と CNN+LSTM の精度はほぼ同等と言える. よって学習データ数が少ない時は, CNN+LSTM は CNN より識別能力が高く, 学習データ数が十分にある時は CNN の精度は CNN+LSTM と同程度, もしくはそれ以上の識別精度をもつと考えられる. LSTM は, 過去の情報が記憶可能である点を特徴とするため, CNN+LSTM の手法では脳波特性の時間変化をよく認識できたために, 少ない学習データからも高い判別能力が得られたと考えられる. 本研究では, 学習に用いる実測データ数が少ないと前提としたため, 以降 CNN+LSTM の性能について検証を行った.

### Subject 1



### Subject 2

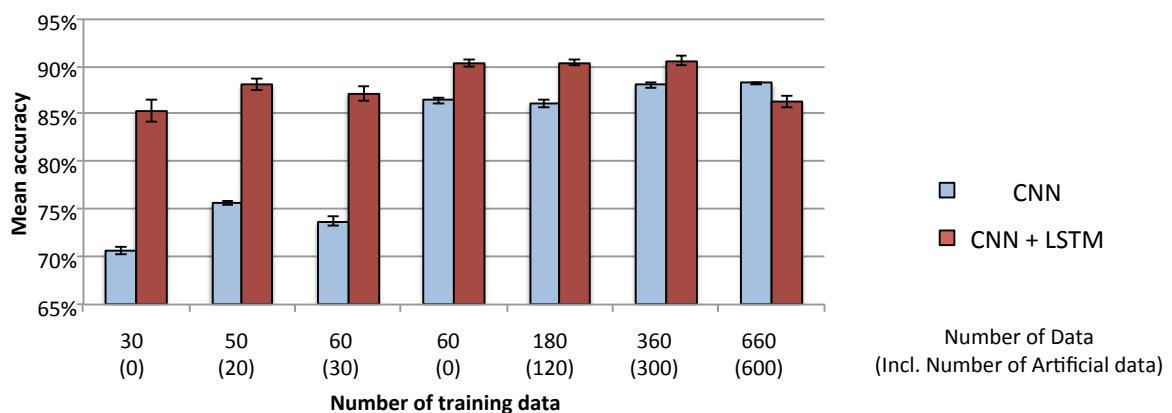


図 5-1 CNN と CNN+LSTM の判別精度の比較. 上段に subject1, 下段に subject2 の結果を示す.

### 5.3 学習データ数と識別精度

図 5-2 に subject1 における、学習に用いた実測データ数とその時の識別精度を示した。データ数が増えるに従い識別精度は高くなるが、その増加率は徐々に低下していることがわかる。そこで、横軸に学習で用いた実測データ数の逆数をとり再プロットした（図 5-3）。識別精度のグラフから最小二乗法により近似直線を求めるとき、その  $R^2$  値は 0.98 となることから、識別精度のグラフは直線によく近似可能であることがわかる。よって判別精度は、学習に用いる実測データ数に対して比例定数が負の反比例に従うことが示された。ここで近似直線の  $y$  切片が 0.8739 であることから、実測データを増やしたとき識別精度は約 87% に収束すると考えられる。この値は、subject1 に CNN + LSTM を用いたときに得られる精度上限を示していると言える。つまり CNN + LSTM を用いた脳波判別は、どれだけ多くの実測データを用いて学習してもその精度に限界点があることが示唆される。

subject2 においても subject1 と同様に学習に用いた実測データ数と識別精度を比較した（図 5-4）。学習データ数が 6~20 のときは識別精度が安定せず、subject1 のように学習データ数に対して徐々に精度が向上する様子は見られなかった。表 4-6 の標準偏差を見ると、学習データ数が少ない時は結果にばらつきが大きいことから、これは過学習が起きたために安定した精度向上が見られなかつたと考えられる。しかし、学習データ数が 30 以上になると、subject1 同様に、判別精度は学習データ数増加に従い識別精度が高くなることがわかる。図 5-5 に横軸を学習に用いた実測データ数の逆数としたときの精度グラフを示す。ここで学習データ数が 30 個以上のときの結果に対して近似直線を求めた。その結果、近似直線の  $R^2$  値は 0.86 となり、subject 2 においても学習に用いる実測データ数が 30 以上であれば、subject1 同様直線によく近似可能であることがわかる。また  $y$  切片の値より、約 94% で精度に頭打ちが生じることが示唆される。

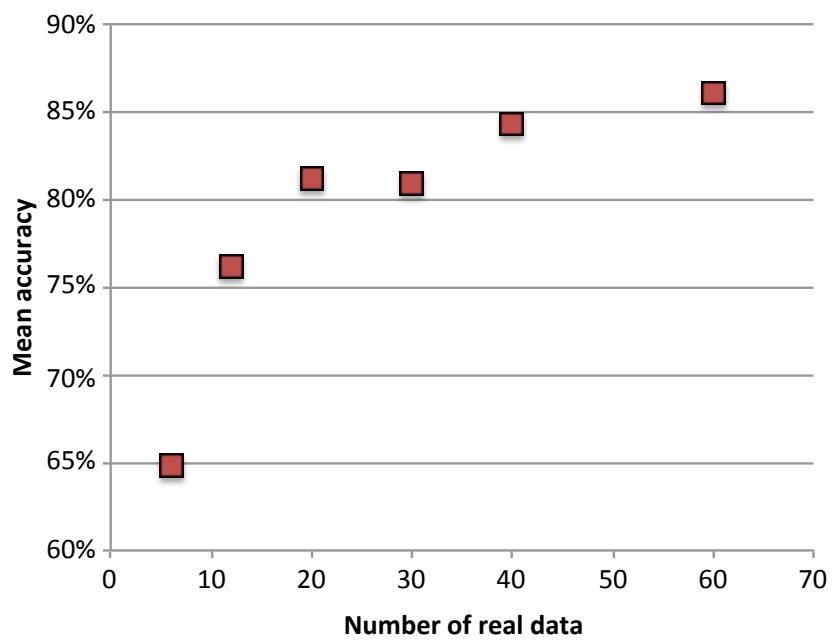


図 5-2 subject 1 における学習に用いた実測データ数に対する識別精度の変化.

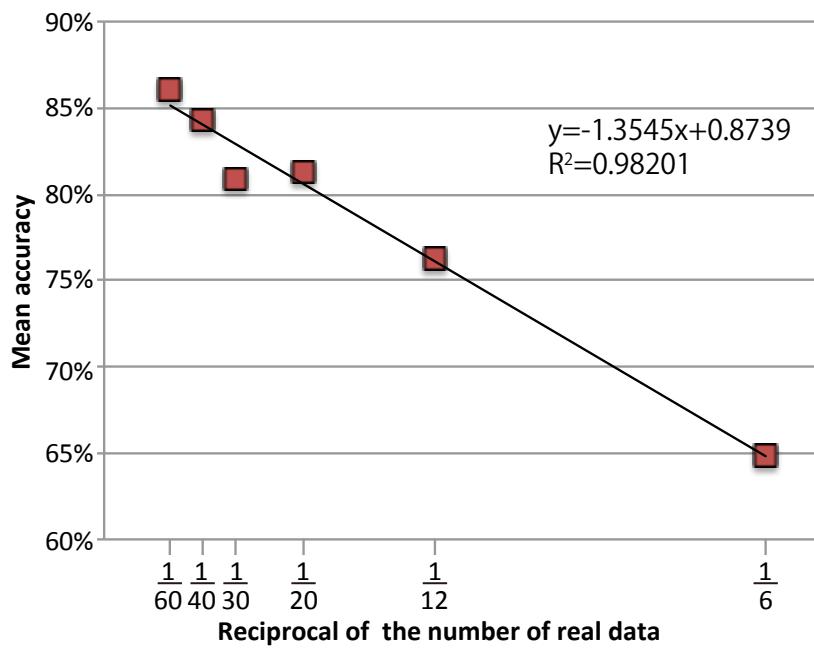


図 5-3 subject 1 における学習に用いた実測データ数に対する識別精度の変化. 横軸を学習データ数の逆数とする. 黒線はプロットに対する近似直線を示す.

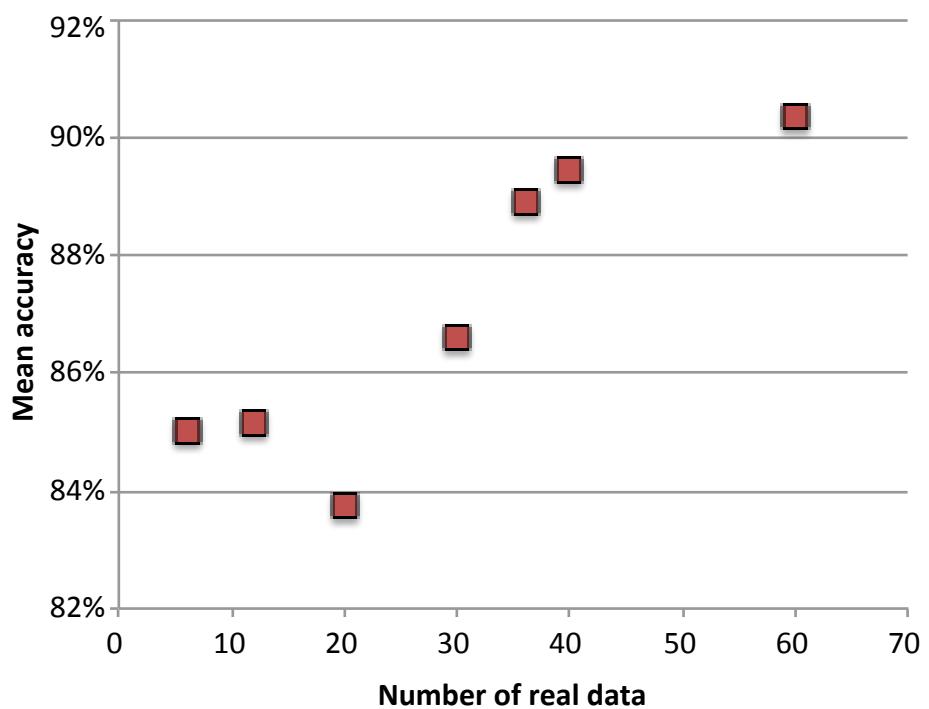


図 5-4 subject 2 における学習に用いた実測データ数に対する判別精度の変化.

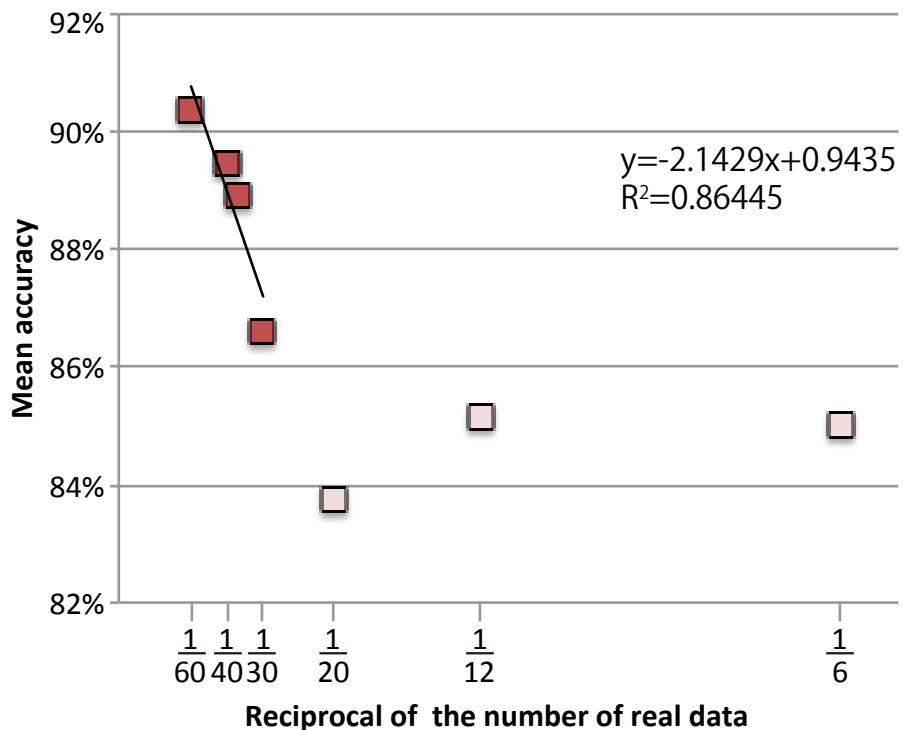


図 5-5 subject 2 における学習に用いた実測データ数に対する判別精度の変化. 横軸を学習データ数の逆数とする. 黒線はプロットに対する近似直線を示す. 濃赤色は近似直線を求めた際に用いたプロットを示す.

## 5.4 人工データの有用性

図 5-6 に, subject1において実測データのみで CNN+LSTM を学習した場合の判別精度 (■) と, 人工データを加え 60 個にした場合の判別精度 (▲) の比較を示す. 実測データ 12 個以上での判別精度 (▲) は, 実測データだけの判別精度 (■) より向上した. また, 実測データ数 30 個以上で学習した時の判別精度 (▲) は 85.89% であり, これは実測データ数 60 個で学習した時の判別精度 (■) 86.07% とほぼ同程度である. ゆえに, 人工データは NN の判別精度向上に有用であり, 人工データにより実測データを補えることが示唆された. 一方, 実測データ 6 個を人工データにより 60 個に増やした場合の精度は, 実測データ 6 個のみの精度より低下している. この結果から, 人工データ数に対して実測データ数が少なすぎると判別精度が低下することが示唆された.

これより, 人工データ作成に最低限必要な実測データ数について議論する. 図 5-7 では図 5-6 の x 軸を学習に用いた実測データ数の逆数とした. この逆数と判別精度 (■) は, 5.3 章で述べた通り, ほぼ比例関係にあるが, 実測データ数の逆数と判別精度 (▲) との関係もまた, ほぼ比例関係にある. これら二つの近似直線には交点があり, この交点 K の x 座標値が 0.10716 であることから,  $1/0.10716 \approx 9.33$ , 即ち 9 個以下の実測データを人工データで 60 個に増加し NN 学習させても, 判別精度向上には寄与しない, もしくは逆効果であることがわかる. 以上より NN 学習において実測データから人工データを作成する際の増加率に限界値があり, subject1 の場合  $60 \div 9 \approx 6.7$  より約 6~7 倍が限界であると考えられる. この限界値を踏まえた上で, 表 4-5 の実測データ 60 個に人工データを加えて学習データを 180 個 (3 倍), 360 個 (6 倍), 660 個 (11 倍) としたときの判別精度を見ると, 3 倍, 6 倍のときは精度が向上しているが, 11 倍になると精度低下が見られる. この低下は学習に用いる人工データ数の割合が大きすぎたためだと考えられる.

subject2においても上記と同様の検証を行った. 図 5-8 に subject2 における実測データのみで CNN+LSTM を学習した場合の判別精度 (■) と, 人工データを加え 60 個にした場合の判別精度 (▲) の比較を示す. また, 図 5-9 には図 5-8 の横軸を学習に用いた実測データ数の逆数とした場合を示す. 実測データ数が 20 以下の時は, 判別精度にばらつきがあるため, 図 5-9 では実測データ数 30 個以上の場合で直線近似を求めた. 結果, 判別精度 (■) 及び (▲) の近似直線について, その交点の x 座標は  $\frac{1}{29.38}$  となっ

た. よって subject2 では NN 学習において実測データから人工データへの増加率に限界値は  $60 \div 29 \approx 2.1$  より約 2 倍であると推測される. ただし, 直線近似式を求める際に用いたプロット数は 3 点と少ないことから, 正確な評価をするためには十分な量の実測データのもと検証する必要がある.

本研究で算出した限界値は subject1, 2 で異なる値を示した。よって人工データが ANN の学習に及ぼす効果は、被験者によって異なることがわかった。

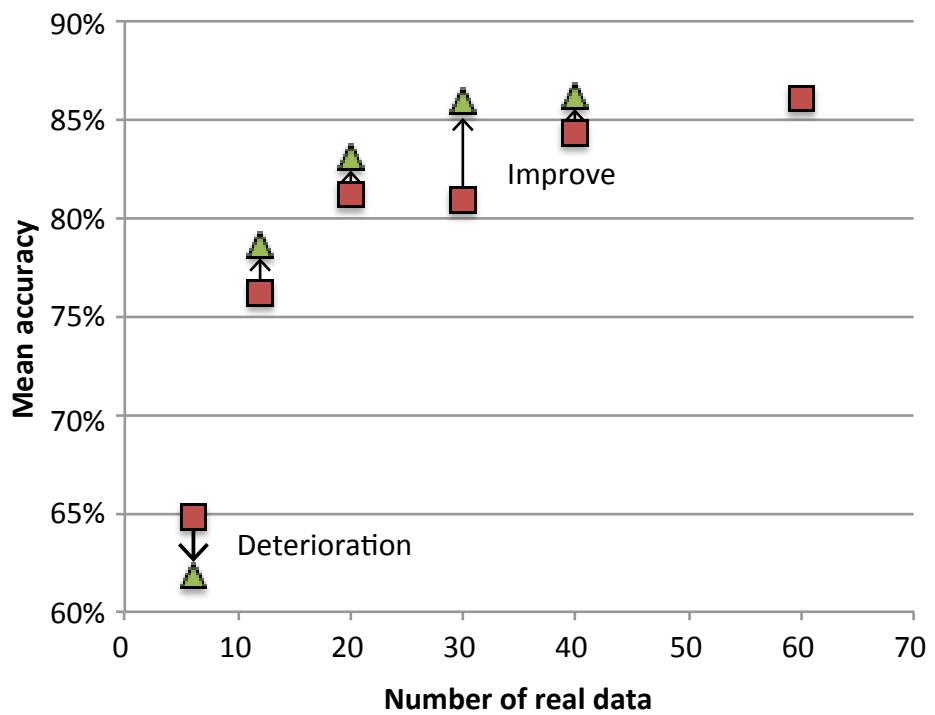


図 5-6 subject 1 における学習に用いた実測データ数に対する判別精度の変化. ▲は実測データのみで学習した場合を示し, ■は実測データに人工データを加え 60 個にした場合の判別精度を示す.

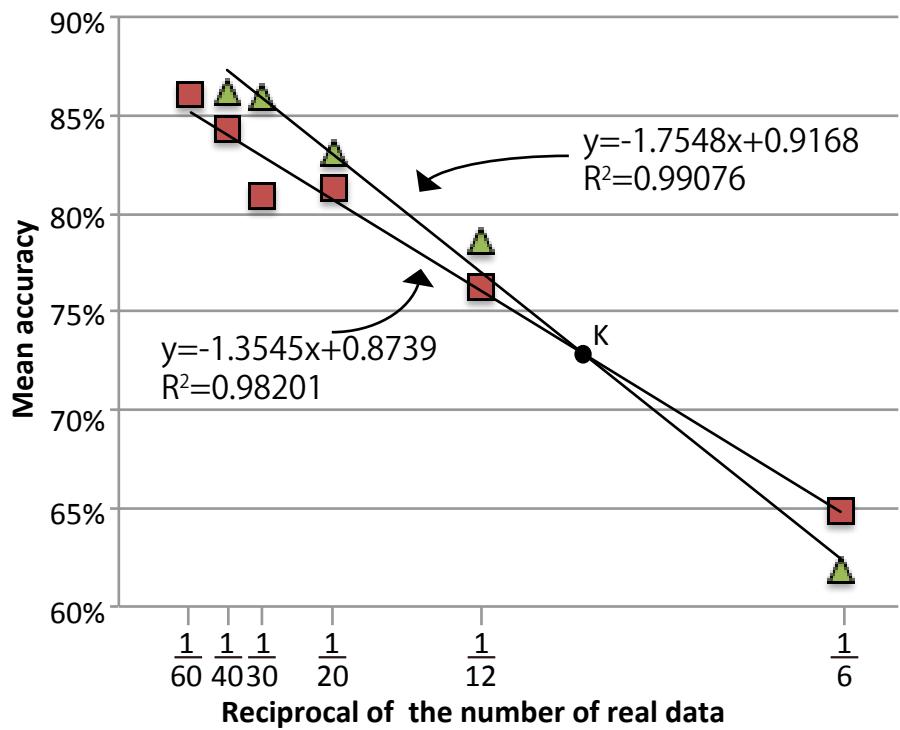


図 5-7 subject 1 における学習に用いた実測データ数に対する判別精度の変化. 横軸は実測データ数の逆数とする. ▲は実測データのみで学習した場合を示し, ■は実測データに人工データを加え 60 個にした場合の判別精度を示す. 黒線はプロットに対する近似直線を示す.

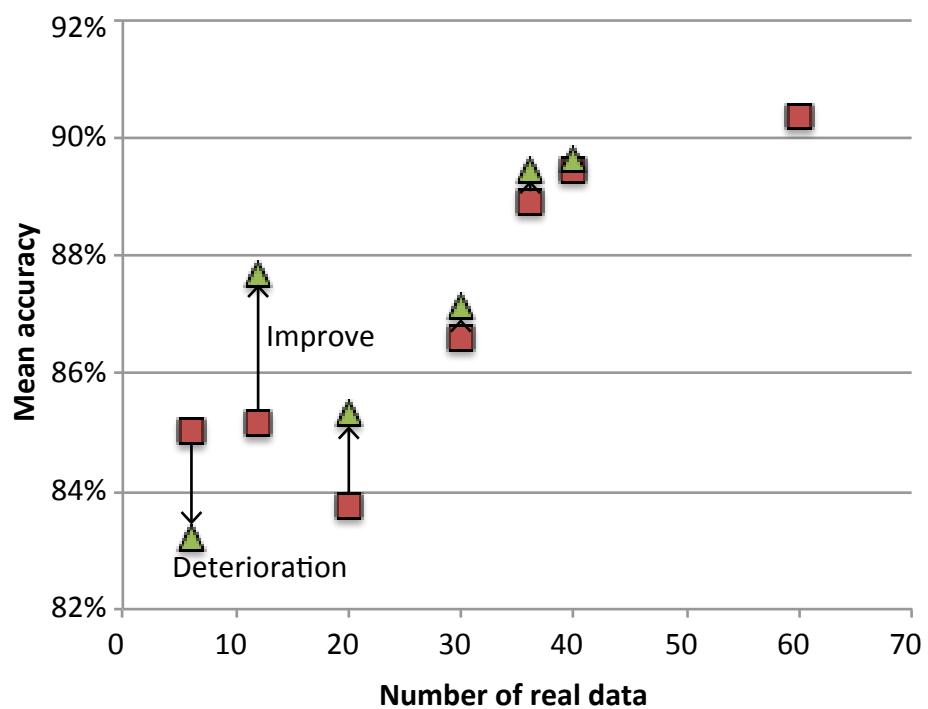


図 5-8 subject 2 における学習に用いた実測データ数に対する判別精度の変化. ▲は実測データのみで学習した場合を示し, ■は実測データに人工データを加え 60 個にした場合の判別精度を示す.

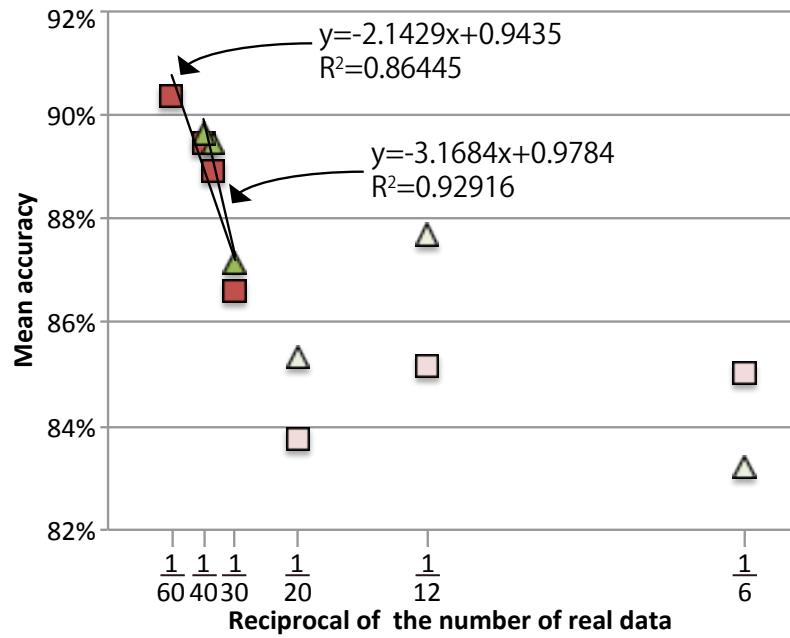


図 5-9 subject 2 における学習に用いた実測データ数に対する判別精度の変化. 横軸は実測データ数の逆数とする. ▲は実測データのみで学習した場合を示し, ■は実測データに人工データを加え 60 個にした場合の判別精度を示す. 黒線はプロットに対する近似直線を示す. 濃赤色及び濃緑色は近似直線を求めた際に用いたプロットを示す.

## 6 結論

本研究では、CNNとLSTMを組み合わせたANN構造を提案し、その判別精度を検証した。また、限られた実測脳波から人工脳波を作成し、学習データ数を増やすことでANNの判別精度向上を図り、人工脳波の有用性を検証した。本研究の主な結論について以下にまとめる。

- ・ 本研究が提案するCNNとCNN+LSTMによる運動想起判別手法は従来手法であるCSPと同程度、もしくはより高い識別精度が得られた。またその精度の向上率は被験者によって異なった。
- ・ 学習データ数が少ない場合、CNN+LSTMはCNNより判別能力が高かった。これはLSTMが脳波特性の時間変化をよく認識できたためだと考えられる。
- ・ CNN+LSTMを用いた運動想起判別において、学習に用いる実測データ数が増えるに従い判別精度は向上した。しかし向上率は徐々に低下して、頭打ちが発生することが示唆された。
- ・ ANNの学習に人工データを加えることで、判別精度向上が見られた。ただし実測データから人工データを作成する際の増加率には限界値があり、人工データ数に対して実測データ数が少なすぎると判別精度向上には寄与しない、もしくは逆効果であることが示唆された。

以上より、本研究が提案したCNNとLSTMを組み合わせた運動想起判別手法は、少ない実測脳波データによる学習でも高い判別能力が得られ、またその能力は人工脳波によって向上することがわかった。

## 7 文献

- Dinares-Ferran, J., Ortner, R., Guger, C., & Sole-Casals, J. (2018). A new method to generate artificial frames using the empirical mode decomposition for an EEG-based motor imagery BCI. *Front. Neurosci.*, 12:308. doi: 10.3389/fnins.2018.00308
- Erp, J., Lotte, F., & Tangermann, M. (2012). Brain-computer Interfaces: Beyond medical applications. *Computer*, 45(4), 26-34. doi: 10.1109/MC.2012.107
- Gers, A. F., Schmidhuber, J., & Cummins, A. F. (2000). Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, 12(10), 2451-2471. doi: 10.1162/089976600300015015
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. doi: 10.1162/neco.1997.9.8.1735
- Huang, E. N., Shen, Z., Long, R. S., Wu, C. M., Shih, H. H., Zheng, Q., Yen, N., Tung, C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences*, 454(1971), 903-995. doi: doi.org/10.1098/rspa.1998.0193
- 伊藤 伸一, 満倉 靖恵, 福見 稔, 赤松 則男 (2004). 脳波の個人特性を考慮した脳波分析法の提案. *電気学会論文誌 C*, 124(6), 1259-1266. doi: 10.1541/ieejeiss.124.1259
- 開, 一夫, 金山範明 (2016). 脳波解析入門 EEGLAB と SPM を使いこなす. 東京大学出版会.
- Lance, J. B., Kerick, E. S., Ries, J. A., Oie, S. K., & McDowell, K. (2012). Brain-computer interface technologies in the coming decades. *Proceedings of the IEEE*, 100, 1585-1599. doi: 10.1109/JPROC.2012.2184830
- Li, X., Song, D., Zhang, P., Hiu, Y., & Hu, B. (2017). Deep fusion of multi-channel neurophysiological signal for emotion recognition and monitoring. *International Journal of Data Mining and Bioinformatics*, 18(1), 1-27. doi: 10.1504/IJDMB.2017.086097
- Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., & Yger, F. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3). doi: 10.1088/1741-2552/aab2f2
- Muller, R. K., Anderson, W. C., & Birch, E. G. (2003). Linear and nonlinear methods for brain-computer interfaces. *IEEE Transactions on Neural Systems*

and Rehabilitation Engineering, 11(2), 165 - 169. doi: 10.1109/TNSRE.2003.814484

Nunez, L. P., & Srinivasan, R. (2007). Electroencephalogram. Scholarpedia, 2(2):1348, . doi: 10.4249/scholarpedia.1348

Pfurtscheller, G., & Lopes da Silva, F.H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol, 10(11), 1842-1857. doi:

Phan, A., & Cichocki, A. (2010). Tensor decompositions for feature extraction and classification of high dimensional datasets. IEICE, 1(1), 37-68. doi: 10.1587/nolta.1.37

Ramoser, H., Muller-Gerking, J., & Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Transactions on Rehabilitation Engineering, 8(4), 441 - 446. doi: 10.1109/86.895946

Reece, B. J., Urry, A. L., Cai, L. M., Wasserman, A. S., Jackson, B. R., & Minorsky, V. P. (2011). キャンベル生物学原書 9 版, 池内昌彦, 伊藤元己, 管本春樹(監訳). 丸善出版, , 1247-1248. doi:

Wolpaw, R. J., Birbaumer, N., McFarland, J. D., Pfurtscheller, G., & Vaughan, M. T. (2002). Brain-computer interfaces for communication and control. Clinical Neurophysiology, 113(6), 767-791. doi: 0.1016/S1388-2457(02)00057-3

山ノ井高洋 (2016). BMI/BCI の研究の現状. 計測と制御, 55(2), 107-113. doi: 0.11499/sicejl.55.107

Zhang, D., Yao, L., Zhang, X., Wang, S., Chen, W., Boots, R., & Benatallah, B. (2018). Cascade and parallel convolutional recurrent neural networks on EEG-based intention recognition for brain computer interface. AAAI Publications, Thirty-Second AAAI Conference on Artificial Intelligence, 1703-1710.

## 8 謝辞

本研究を実施するにあたり，ミケレット・ルジェロ教授には研究や発表についての熱心なご指導を賜りました。心より感謝いたします。

また脳波解析にあたり，技術，開発などあらゆる面においてご指導いただいた Cichocki Andrzej 先生，孫哲研究員，Phan Anh Huy 研究員，Huihai Zhao 研究員をはじめとする理化学研究所の方々に厚く御礼申し上げます。

そして MATLAB 解析プログラムについてご指導いただいた Sole-Casals Jordi さんに深謝いたします。

ニューラルネットワークを用いた時系列データを研究するにあたって，金亜伊准教授からは有意義かつ的確なご意見をいただいたほか，素晴らしい研究発表の機会をくださいました。大変感謝しております。

日々，研究室での学生生活において大変お世話になりましたミケレット研究室の皆様に感謝いたします。

最後に，執筆に携わってくださった全ての方々に，そしてこれまで温かく見守ってくれた家族に感謝の意を示し，謝辞といたします。

## 9 研究発表実績

### (1) 学会発表

1. 高橋馨子, 松本裕也, 上松大輝, 孫哲, ミケレットルジェロ, 金亜伊, ニューラルネットワークを用いた低 SN 比条件下における地震波検出手法の開発, 日本地震学会 2017 年度秋季大会, S02-P04 (ポスター発表), かごしま県民交流センター, 2017 年 10 月 26 日.
2. Kahoko TAKAHASHI, Hiroki UEMATSU, Zhe SUN, Ruggero MICHELETTO, Ahyi KIM, Detection of seismic signals under low SNR condition using an artificial neural network: toward the development of a dense low cost citizen seismic network in Japan, 2018 Seismological Society of America Annual Meeting, 11 (poster sessions), Miami, 17 May 2018.
3. Kahoko TAKAHASHI, Zhe SUN, Jordi SOL-CASALS, Andrzej CICHOCKI, Anh Huy PHAN, Ruggero MICHELETTO, Empirical Mode Decomposition for Improved EEG Signal Classification with Convolutional Neural Network in Brain-Computer Interface Experiments, The 28<sup>th</sup> annual conference of the Japanese neural network society, P1-15 (flash talk and poster sessions) Okinawa listitute of science and technology, 25 Oct 2018.

### (2) 学位論文

1. 卒業論文「人工ニューラルネットワークを利用した高ノイズ環境下における地震波検出手法の開発」(2017 年 2 月提出, 横浜市立大学).