

# 深層学習と fMRI を用いた聴覚野からの音階デコードにおける精度向上

## 1. 緒言

fMRI などの脳機能画像解析法の発達により、脳神経科学の研究は飛躍的進歩を遂げているが撮像時の動作音などにより聴覚野のメカニズムやデコード研究は、視覚野などと比べて遅れている。先行研究<sup>(1)</sup>において、15 度間隔 4 音 (B2(125Hz), B4(500Hz), B6(2kHz), B8(8kHz)) を聴覚刺激として呈示した際、高い周波数(B8)は聴覚野の後部、低い周波数(B2)は聴覚野の前部で脳賦活が確認され、周波数局在性が明らかになった。しかしながら、周波数解像度が低いという課題がある。

これまでの本研究グループでは、空間解像度 3mm<sup>3</sup> の fMRI で測定したデータに対し、深層学習を用いて 1 度差の脳賦活動による識別を行った結果、3 人中 1 人の被験者の学習が正常に行われず、また最大でも 67.5% の識別率であった。原因として、深層学習への入力データ処理の 1 つの正規化方法が適切でないと考えられる<sup>(2)</sup>。

本報では、これまでの研究結果を受け、全ての被験者において学習が可能となるよう、fMRI で測定したデータを処理する際の正規化方法の改善を行う。従来の方と、本報で提案する 2 つの方法を用いて深層学習で学習・評価を行い、学習過程や結果から音階デコードに適切な処理方法を見出す。その結果より、改善後の提案手法を用いて先行研究の 15 度差と比べ周波数解像度の向上を目指し、4 度差、2 度差、1 度差の聴覚刺激呈示時による識別を行い、これまでの 1 度差の識別結果より高い識別率を得られるか検討する。

## 2. 脳画像撮像・解析手法

本実験では、SIEMENS 社の fMRI 装置、MAGNETOM Verio3T を使用し脳賦活動の測定を行った。測定した画像データは DICOM 形式であるため、NIFTI-1 形式に変換した後、画像解析に SPM12 (Statistical Parametric Mapping) を用いて前処理、並びに統計解析を行った。統計解析では、ヒト聴覚野の活動のみを解析するために、先行研究と同様に聴覚野(ブロードマンマップ 41&42)を関心領域 (Region of Interest, ROI) に設定し、聴覚野の活動のみについて解析を行った<sup>(2)</sup>。統計解析において、各条件の聴覚刺激を受けているときを 1 としてコントラストを作成した。

## 3. DBN (Deep Belief Network)

本研究では、聴覚刺激による脳賦活動の識別を DBN で行う。DBN は深層学習の手法の一種であり、教師なし学習である制約付きボルツマンマシン (Restricted Boltzmann Machine, 以下 RBM) と、教師あり学習であるロジスティック回帰(Logistics Regression) から構成される(図 1)。DBN では多層に積み重ねられた RBM の各層で出力された特徴を次の層の RBM の入力とし、特徴抽出を行う。DBN の最後の層では、ロジスティック回帰を用い教師あり学習を行い効果的な学習を実現している。

DBN を含むニューラルネットワークでは、学習を行う中で過学習を抑制するために、学習誤差に任意の閾値を設定し、各出力細胞の教師値に出力値が近似しないようにする。学習

誤差  $\varepsilon$  の算出は、式(1)により行う。

$$\varepsilon = \frac{1}{N} \sum_{k=1}^N \sum_{n=1}^{n_0} |T_n^k - O_n^k| \quad (1)$$

ここで、 $N$  は総学習データ数、 $k$  は RBM の各層での学習回数、 $n_0$  は出力細胞数、 $T_n^k$  は  $n$  番目の出力細胞の教師値、 $O_n^k$  は  $n$  番目の出力細胞の出力値である。図 1 に示す  $k$  は RBM の学習回数、 $m$  はロジスティック回帰の学習回数である。

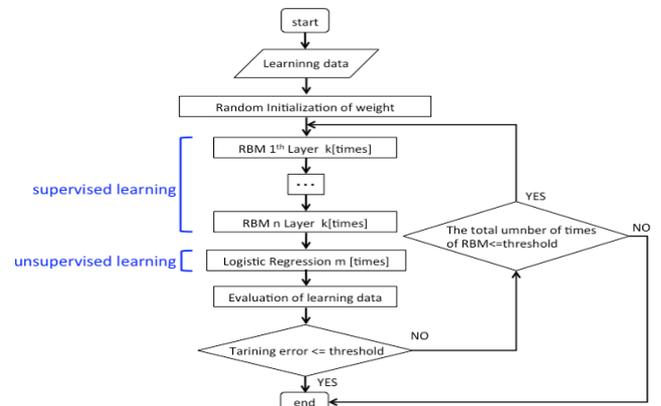


Fig. 1 Flow of DBN

## 4. 正規化方法別比較実験

### 4.1 実験目的

fMRI で測定したデータを DBN で学習するために正規化処理を行うが、従来の方が適切ではない可能性がある。本実験では、正規化の従来法と改善法を比較し、全ての被験者において学習可能な正規化方法を模索する。

### 4.2 実験方法

本実験では、任意の 30 秒間の fMRI の動作音に対して周波数解析を行ったピーク値の結果<sup>(2)</sup>より、人の可聴領域内で fMRI の動作音による影響が最も低いと考えられる C7 (2097.0Hz) の音を基準とする 1 度差 (C7 と C#7 (2217.5Hz)) を聴覚刺激とする。fMRI を用いて、C7, C#7 の聴覚刺激呈示時の脳活動をそれぞれ 60 枚ずつ、計 120 枚撮影する。実験デザインを図 2 に示す。撮像は、連続的に行い、DBN の識別で使用する測定データは、図 2 のように刺激提示直後とする。本実験では、聴覚野への動作音の影響を抑えるためにホワイトノイズを使用する<sup>(4)</sup>。実験で取得した脳画像を、SPM12 を用いて 2 章で述べた方法で解析を行う。正規化方法は以下の 3 通りである。(a) は従来の方法、(b)(c) は本研究で新たに提案する方法である。

- (a) 被験者毎に得られた聴覚野の座標上で、t 検定の統計量を [0, 1] の連続値に正規化
- (b) 聴覚刺激条件毎に得られた聴覚野の座標上で、t 検定の統計量を [0, 1] の連続値に正規化
- (c) 1 ボリューム毎に得られた聴覚野の座標上で、t 検定の統計量を [0, 1] の連続値に正規化

正規化したデータを DBN の一層目の入力データとする。本実験には、聴力検査(1000Hz, 4000Hz)で異常の無い 20 代健常男性 3 名の被験者に同意を得て実施した。

## 卒論要旨

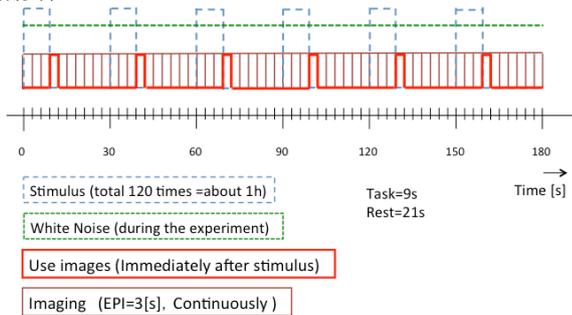


Fig. 2 Experiment design

作成したデータから各条件において、ランダムに 80 個(1音につき 40 個)を学習データに、残りの 40 個(1音につき 20 個)を評価データとする。DBN の RBM を 2 層とし、3 章で述べたパラメータの総学習データ数  $N$  を 80, 出力細胞数  $n_0$  を 2, RBM の総学習回数を 100000, ロジスティック回帰の学習回数を 300 と設定する。本実験では, RBM2 層の細胞数を 300~1000 の範囲内で「1 層目の細胞数 > 2 層目の細胞数」と設定し, 個人毎に識別を行う。学習誤差  $\varepsilon$  を 0.1 とし, 0.1 以下になった場合に学習を正常終了させる。

### 4.3 実験結果

DBN での 3 通りの正規化方法で作成したデータの学習過程を表 1 に示す。RBM2 層の細胞数の組み合わせを変更したが, 全て表 1 の様になった。

Table 1 Learning process

	a	b	c
Subject A	Forced termination	Forced termination	Normal termination
Subject B	Abnormal termination	Abnormal termination	Normal termination
Subject C	Forced termination	Forced termination	Normal termination

強制終了は, 学習過程で非数が発生したことにより学習が進まず, 人為的学習停止を表し, 異常終了は, 学習は RBM の総学習回数の上限により終了しているが, 学習誤差  $\varepsilon$  は閾値以下に達していないことを表す。正規化方法(c)が他の 2 つの方法に比べて, C7, C#7 の特徴の差が大きく現れたため, 全被験者において学習が可能となったと考えられる。今後, DBN への入力データの処理は正規化方法(c)を採用する。

## 5. 度数の差による識別実験

### 5.1 実験目的

本実験では, C7 を基準とした 4 度差, 2 度差, 1 度差の聴覚刺激呈示時の脳賦活動の測定を行い, 得た測定結果を 4 章の結果を受けた提案システムで識別し, 識別率, 及び解像度を向上させる。

### 5.2 実験方法

4.2 節同様の条件・手順で, fMRI で C7 の音を基準とし, C7 と F7(2793.8Hz) の 4 度差, C7 と D7(2349.3Hz) の 2 度差, C7 と C#7 (2217.5Hz) の 1 度差の脳賦活動を撮像し, 前処理, 及び統計解析を行う。4.3 節の結果を受け, t 統計量を 1 スキャン毎の最大値と最小値を使用して [0, 1] の連続値に正規化する。DBN への入力データも, 4.3 節同様の手順で作成する。RBM2 層の細胞数を 300~1000 の範囲内で「1 層目の細胞数 > 2 層目の細胞数」となり, 識別率  $\eta$  が高くなるように設定する。4.2 節同様に聴力検査で異常のない 20 代健常男性 10 名の被験者に同意を得て実施した。

### 5.3 実験結果

DBN での 4 度差, 2 度差, 1 度差の中間細胞数と識別率の結果を表 2~4 に示す。全ての被験者で学習誤差が 0.1 を下回り, 学習は, 正常に終了した。

Table2 C7, F7 Rate

	Number of Hidden Layer		C7[%]	F7[%]	Total [%]
	1th	2nd			
Subject A	700	650	65.0 (13/20)	90.0 (18/20)	77.5 (31/40)
Subject B	900	800	95.0 (19/20)	75.0 (15/20)	85.0 (34/40)
Subject C	900	800	70.0 (14/20)	90.0 (18/20)	80.0 (32/40)
Subject D	550	450	85.0 (17/20)	80.0 (16/20)	82.5 (33/40)
Subject E	400	300	65.0 (13/20)	75.0 (15/20)	70.0 (28/40)
Average			76.0	80.0	79.0

Table3 C7, D7 Rate

	Number of Hidden Layer		C7[%]	D7[%]	Total [%]
	1th	2nd			
Subject B	800	750	75.0 (15/20)	85.0 (17/20)	80.0 (32/40)
Subject D	600	550	85.0 (17/20)	80.0 (16/20)	82.5 (33/40)
Subject E	900	800	90.0 (18/20)	85.0 (17/20)	87.5 (35/40)
Subject F	900	850	75.0 (15/20)	65.0 (13/20)	70.0 (28/40)
Subject G	500	400	40.0 (8/20)	80.0 (16/20)	60.0 (24/40)
Average			73.0	79.0	76.0

Table4 C7, C#7 Rate

	Number of Hidden Layer		C7[%]	C#7[%]	Total [%]
	1th	2nd			
Subject A	850	750	70.0 (14/20)	50.0 (10/20)	60.0 (24/40)
Subject B	850	750	75.0 (15/20)	70.0 (14/20)	72.5 (29/40)
Subject C	500	400	85.0 (17/20)	40.0 (8/20)	62.5 (25/40)
Subject D	650	600	70.0 (14/20)	70.0 (14/20)	70.0 (28/40)
Subject E	900	850	65.0 (13/20)	60.0 (12/20)	62.5 (25/40)
Subject H	950	850	65.0 (13/20)	60.0 (12/20)	62.5 (25/40)
Subject I	950	850	75.0 (15/20)	55.0 (11/20)	65.0 (26/40)
Subject J	900	800	65.0 (13/20)	60.0 (12/20)	62.5 (25/40)
Average			71.25	58.13	64.7

## 5.4 考察

実験結果から, 提案システムを用いることで, 15 度差より細かい 4 度差, 2 度差において, 被験者によって, 度数の差による脳賦活動の識別が可能である見通しが得られた。1 度差の最大識別率は 72.5% となり, これまでの結果を上回った。しかしながら, 2 つの課題が残った。

- ① 各度数差の識別率が, 60% 程度にとどまっている被験者がいる(2 度差:被験者 D, 1 度差:被験者 A, C~H, J)。
- ② RBM の細胞数は, 経験からのみの選択である。

これらの課題の解決方法として, 被験者への負担を考慮しながら音圧を上げ, BOLD 効果<sup>(2)</sup>を踏まえて, 聴覚刺激に対して反応が最初に高くなる刺激呈示 6 秒後の脳賦活動の測定データを使用し, DBN で識別を行う。RBM の細胞数と識別率の規則性を見出し, 識別率のさらなる向上を目指す。

## 6. 結言

本研究では, 音階デコード技術開発のため fMRI で測定したヒト脳の聴覚野の活動から, 被験者がどのような純音を聴いているか DBN を用いて識別を行った。本論文では, データ処理における正規化方法を新たに提案し, 本実験の全被験者において学習が正常終了し識別精度が向上した。また, 先行研究より高い周波数解像度での識別が可能となり, かつ, これまでの研究より識別率が向上した。しかしながら, 最小単位である 1 度差の識別において高い識別率を得る必要や, 別の音階, 度数差の実験を行う必要がある。

## 文献

- (1) Dave R.M. Langers, et.al. “Representation of lateralization and tonotopy in primary versus secondary human auditory cortex”, NeuroImage34, (2007), pp.264–273.
- (2) 金子竜也 他, “ディープラーニングを用いたヒト脳聴覚野からの音階デコード技術”, 日本機械学会 中四国支部第 54 期総会・講演会, 講演番号 602, 2016