# CNN を用いた音高想起時の脳波デコーディングの試み

清水菜々子\*1, 曽我真人\*2 \*1 和歌山大学大学院システム工学研究科 \*2 和歌山大学システム工学部

# Attempt of EEG Decoding When Recalling Pitch Using CNN

Nanako Shimizu\*1, Masato Soga\*2
\*1 Graduate School of Systems Engineering, Wakayama University
\*2 Faculty of Systems Engineering, Wakayama University

In recent years, research on a technique called brain decoding has been actively conducted. The data obtained by measuring brain activity is thought to be the conversion of human mental states and thoughts into codes. Therefore, attempts have been made to clarify a human mental state and thoughts by decoding the data obtained by measuring brain activity. This attempt is called brain decoding. Brain decoding attempts have been made for various tasks, but much has not been clarified for the task of recalling sound. Therefore, in this study, we measured the electroencephalogram (EEG) when recalling sounds with different pitches. In addition, we applied the obtained data to a convolutional neural network to verify whether it is possible to identify the recalled pitches.

キーワード: ブレイン・デコーディング, 脳波, 音高, 想起, CNN

#### 1. 研究背景

#### 1.1 ブレイン・デコーディング

近年,ブレイン・デコーディングと呼ばれる技術の研究が盛んに行われている. 脳活動を計測して得られたデータは,人の精神状態や思考がコード化されたものであると考えられている. そのため, 脳活動を計測して得られたデータをデコーディングすることで人の精神状態や思考を明らかにしようする試みがなされている. この試みをブレイン・デコーディングと呼ぶ(図1). ブレイン・デコーディングに関する研究事例として,睡眠中の脳活動データを解読することで,見ている夢の内容を明らかにする研究や(1),映像を見ているときの脳活動データを解読することで,映像を見て感じた内容を言語化する研究などが挙げられる(2).



図 1 ブレイン・デコーディング

### 1.2 脳と音楽

聴取した音楽を脳内で処理するときには、聴覚野が深く関係している. 聴覚野は一次聴覚野, 二次聴覚野, 三次聴覚野に分けられる. 一次聴覚野は音圧や音高, 音色などの音を構成する要素の認識に関係していると考えられており、二次聴覚野はリズム, メロディ, ハーモニーなどの音楽を構成する要素の認識に関係していると考えられている. そして, 三次聴覚野は音楽の全てを統合する処理を行なっていると考えられている。 音楽に関する処理については一般的に脳の右半球

が優位となることが多い<sup>(4)</sup>. また,楽曲の想起時において運動前野が活性化した事例が報告されている<sup>(5)</sup>. ただし,実施する課題の内容や被験者の違いによって異なる結果が得られることもあり,音楽に関する脳内の働きは明確にはなっていない.

聴取した音の音高の違いが脳波に与える影響について検証を行なっている研究が存在する(6). この研究では、110、220、440、880、1760[Hz]の周波数の純音を聴取しているときの脳波に含まれる $\alpha$ 波(8~13[Hz])、 $\beta$ 波(14~30[Hz])、 $\theta$  波(4~7[Hz])について分析を行なっている. 結果として、 $\alpha$ 波の総和と $\beta$ 波の総和の比について有意差があること、440[Hz]において $\alpha$ 波、 $\beta$ 波、 $\theta$ 波のパワーが低下する傾向があること、脳の右半球と左半球を比較すると $\beta$ 波において有意差があることが明らかになった.

#### 1.3 脳波と CNN

脳波を識別する際に深層学習を活用する試みが近年 では活発に行われている. 通常, 脳波を解析し識別を 行うためには専門的知識や経験が必要とされる. その ため,深層学習を用いることで,専門的知識や経験が なくとも脳波を識別できるようになることが期待され ている. ここでは、CNN (畳み込みニューラルネット ワーク)を用いて脳波の識別を行なっている研究につ いて説明する. なお、CNN は主に画像認識の分野で用 いられる深層学習の一種である. 先行研究として, 4つ の異なる思考を行なっているときの脳波を画像化し, CNN を用いて識別を行なっている研究が存在する(7). この研究では、「利き腕を動かす」、「計算する」、「回転 する3次元物体を想像する」、「食べ物を食べる」とい う4つのイメージを被験者に行ってもらい、そのとき の脳波を計測している、そして、脳波に対してウェー ブレット変換を施し得られた画像を CNN の入力デー タとして使用している. 結果として, 40.4%の平均識別 率が得られた.

# 2. 研究目的

ブレイン・デコーディングは様々な課題遂行時の脳 活動データに対して行われているが、音を想起したと きのブレイン・デコーディングについては明らかにな っていない点が多く存在している。そこで、本研究で は音高の異なる音を想起したときの脳波を計測し、得られたデータを CNN に入力することで、被験者がどの音高を想起していたのか識別することが可能かどうかの検証を行なった(図 2).

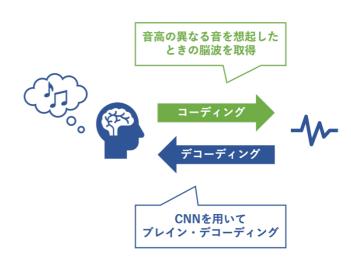


図 2 音高想起時の脳波デコーディング

## 3. 研究の意義

本研究の意義は、ほとんど知見のない異なる音高を 想起したときの脳波を識別するという課題に対して新 しい知見を得ることができることである。また、本研 究が発展することで、医学領域およびエンターテイン メント領域における活用が期待できる。脳からの命令 が筋肉に伝わらなくなり、コミュニケーション手段を 失ってしまう ALS 患者の音楽的表出に寄与できる可 能性が考えられる。また、例えば頭の中でメロディを 思い浮かべるだけで作曲ができるようになるなど、作 曲の知識がない人の制作活動に役立つ可能性も考えら れる。

# 4. 関連研究

本研究と関連性のある脳と音と深層学習に関する研究として2つの先行研究が存在する.1つは及川らが2019年に発表した「畳み込みニューラルネットワークを用いた脳波による音響信号再構成」(®)、もう1つは2019年に繁本らが発表した「fMRIを用いたヒト聴覚野からの音階デコードアルゴリズムの検討」である(®).1つ目の先行研究では、聴覚刺激として262[Hz]の純音を呈示したときの脳波から聴覚刺激を再構成することを試みている.結果として、脳波に短時間フーリエ

変換を施し得られたスペクトログラムを CNN の入力 データとして用いることで、262[Hz]の周波数の波形 を再構成することに成功している. 2 つ目の先行研究 では、聴覚刺激として C7(2097[Hz])と C#7(2217[Hz]) の 2 種類の純音を呈示したときの fMRI データを深層 学習の一種である Deep Belief Network に入力し識別 することを試みている. 結果として、1 人の被験者に おいて 62.50%、もう 1 人の被験者において 70.83%の 識別率を得ることができている.

これら2つの先行研究と比較した上での本研究の位置付けについて説明する.これら2つの先行研究は,深層学習を活用することで脳活動の内容を明らかにしようと試みている点、および音に関係する脳活動を対象としている点で本研究と類似している.一方で,これら2つの先行研究は,音を聴取しているときの脳活動データを深層学習に適用しているのに対し,本研究では音を想起しているときの脳活動データを深層学習に適用するという点で異なっている.

# 5. 実験手法

#### 5.1 実験課題

本研究では、被験者に「低いラ」、「ラ」、「高いラ」の3つの異なる音を想起してもらう実験を実施し、そのときの脳波を計測した。そして、計測した脳波をCNNの入力データとし、識別率を算出した。また、想起課題を実施してもらう前に聴覚刺激を呈示した。聴覚刺激の呈示は、想起する音高を全ての被験者でなるべく統一すること、想起課題は繰り返し行なってもらうため、試行の度に想起する音高がずれていくのをなるべく防ぐことを目的としている。「低いラ」を想起する実験ではA3(220[Hz])、「ラ」を想起する実験ではA4(440[Hz])、「高いラ」を想起する実験ではA5(880[Hz])の聴覚刺激を呈示した。なお、聴覚刺激の種類として、純音とピアノの音の2種類を用意した。

#### 5.2 実験システム

実験システムの流れを図3に示す.実験システムは, スタート画面を呈示した後,「注視点の呈示」,「聴覚刺激の呈示」,「注視点の呈示」,「想起課題の実施」を一連の流れとして繰り返し実行される.

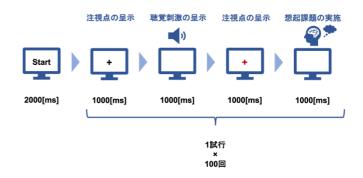


図 3 実験システム

### 5.3 脳波計

脳波計測には、BIOSEMI Active Two を使用した. 64 電極を頭部全体に設置し、512[Hz]のサンプリングレートで脳波の計測を行なった. 電極の配置は図 4 に示す通りである. なお、両耳の乳様突起に設置した電極の平均電位を基準電位としている.

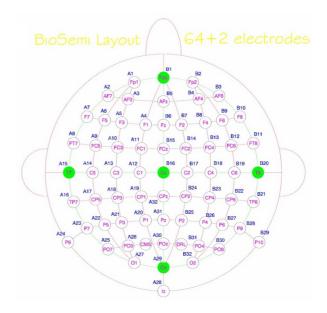


図 4 電極の配置

(引用元: BIOSEMI, "Headcaps"(10))

#### 6. CNN による識別

### 6.1 脳波解析の方針

先行研究より、音楽に関係する脳の働きは右半球が優位であること<sup>(4)</sup>、楽曲を想起したときに運動前野が活性化することが分かっている<sup>(5)</sup>. したがって、本研究では右半球の運動前野に相当する位置に設置した電極 (FC6) のデータを使用することとした。また、異なる音高の音を想起したときに $\alpha$ 波、 $\beta$ 波、 $\theta$ 波のパワーに影響がある可能性が示唆されている<sup>(6)</sup>. したが

って、本研究は想起したときの脳波を対象としているが同様に $\alpha$ 波、 $\beta$ 波、 $\theta$ 波のパワーに影響があると仮定する。また、脳波は時間分解能が高いため時間変化も重要な要素となる。そのため、 $\alpha$ 波、 $\beta$ 波、 $\theta$ 波を含む周波数帯域(0.5~30[Hz])のパワーの時間変化に着目することとした。

#### 6.2 脳波の前処理・解析

脳波の前処理では、ノイズ除去を目的として0.5[Hz]のハイパスフィルタと 30[Hz]のローパスフィルタを適用した。その後、解析対象となる電極のデータのみを抽出し、「想起課題の実施」中の脳波のみを全体から切り出した。そして、切り出した脳波のデータ群から $\pm 100[\,\mu\,V]$ を超える脳波をまばたきや身体を動かしたことなどによるノイズを含むものとして除外した。これらの前処理を施した後にウェーブレット変換を行った。

#### 6.3 CNN への適用

CNN に適用するにあたり、ウェーブレット変換の結果得られるスカログラムを画像化した.このスカログラム画像に被験者が想起していた音高を正解ラベルとして紐付けたものをデータセットとした.なお、ノイズを含む脳波を除外したことでデータセットの各クラスに属するサンプル数に偏りができたため、アンダーサンプリングを実施し、各クラスに属するサンプル数を揃えた.そして、データセットの3/4を学習データとして、1/4をテストデータとして使用した.本研究で使用した CNN の構造を図5に示す.なお、ネットワークの学習にはモーメンタム項付き確率的勾配降下法を使用し、初期学習率を0.01、ミニバッチサイズを16、エポックの最大数を3とした.

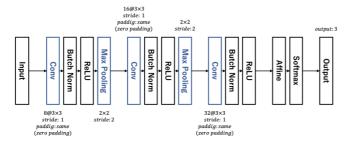


図 5 CNN の構造

## 6.4 結果·考察

10名の被験者の脳波を計測した. 各被験者における 識別率を表 1 に示す. どの被験者においても聴覚刺激 の種類によらず 33.3%の識別率となった. これはチャ ンスレベルと変わらない値である. また, CNN による 学習の過程を確認したところ, 識別率が向上する傾向 は見られなかった.

「低いラ」、「ラ」、「高いラ」をそれぞれ想起したと きの脳波に対して CNN を適用することで脳波を識別 することが可能かどうか検証したが、 結果として高い 識別率を得ることはできなかった. 高い識別率を得る ことができなかった原因として3つの可能性が考えら れる. 1つ目は、CNN に入力したデータが不適切であ った可能性である. 本研究では、「低いラ」、「ラ」、「高 いラ」をそれぞれ想起したときには、 $\alpha$ 波、 $\beta$ 波、 $\theta$ 波のパワーに影響があると仮定した. そして, 0.5~30[Hz]の周波数帯域を解析対象とした. しかし, 本研究での仮定に誤りがあった場合,解析対象とした 脳波に CNN が識別するに足る特徴が含まれていなか った可能性が考えられる. 今後の課題として, 脳波に 含まれる要素のうちどの要素を深層学習の入力データ として使用することが適切かを検証する必要があると 考えられる. 2 つ目は、脳波の前処理・解析の手法が 不適切であった可能性である. 本研究での仮定に誤り がなかった場合において、脳波の前処理・解析の手法 が適切でなかったために脳波に含まれていた特徴をう まく表現できなかった可能性が考えられる. 例えば, 計測した後にどのような解析手法も適用していない生 波形の状態で特徴が現れている可能性も考えられる. 今後の課題として、脳波に対してどのような前処理・ 解析手法を施すことが本研究での課題において適切で あるかを検証する必要があると考えられる.3つ目は、 学習データの数が不足していた可能性である. 深層学 習においてはデータセットの数が多ければ多いほど良 いとされる. そのため、本研究で用意したデータセッ トの数が不十分であったために高い識別率を得ること が出来なかった可能性が考えられる. 脳波のような生 体情報は被験者負担の観点から多量にデータを収集す ることが困難であるという課題を抱えている. そのた め、人工的に脳波のデータを増やす試みをしている先 行研究も存在する(11). 今後の課題として, 実験内容の 工夫により取得できるデータ数を増やすことやデータ 拡張を行うことが考えられる.

表 1 識別率

被験者	A	В	С	D	E
純音 (%)	33.3	33.3	33.3	33.3	33.3
ピアノの音(%)	33.3	33.3	33.3	33.3	33.3

F	G	н	- 1	J	平均	標準偏差
33.3	33.3	33.3	33.3	33.3	33.3	7.49 ×10 <sup>-15</sup>
33.3	33.3	33.3	33.3	33.3	33.3	7.49 ×10 <sup>-15</sup>

## 7. 脳波の特徴量の検証

#### 7.1 検証手法

6.4節の考察を受け、CNNに入力したデータが不適切であった、すなわち本研究での仮定に誤りがあったため CNNに入力したデータに脳波を識別するに足る特徴が含まれていなかった可能性についての検証を行うこととした。そこで、 $\alpha$ 波、 $\beta$ 波、 $\theta$ 波のパワーに対して、想起する音高の違いが影響を与えているかどうかについて検定を実施した。この検定において、「低いラ」、「ラ」、「高いラ」をそれぞれ想起したときの脳波に含まれる $\alpha$ 波、 $\beta$ 波、 $\theta$ 波それぞれについて音高による優位な差があれば、脳波の前処理・解析手法を変更する、もしくはデータセットの数を増やすことで、脳波を識別することができるようになる可能性が考えられる。

データに対して正規性と等分散性の検定を行い,正 規性と等分散性を仮定できるパターンに対してはパラメトリック検定である対応のない二元配置分散分析を 適用し,仮定できないパターンに対してはノンパラメ トリック検定であるフリードマン検定を適用した.検 定では、帰無仮説を「想起する音高の効果はない」、対 立仮説を「想起する音高の効果はある」とした.また, 有意水準を 5%とした.

#### 7.2 結果·考察

対応のない二元配置分散分析の結果を表 2~5,7に,フリードマン検定の結果を表 6 に示す.全てのパターンにおいて p 値が有意水準を上回ったため、帰無仮説は乗却されない.すなわち,聴覚刺激の種類によらず,

 $\alpha$ 波、 $\beta$ 波、 $\theta$  波それぞれのパワーは、想起する音高 の影響を受けているとは言えないことが明らかになった

以上より、脳波に含まれる $\alpha$ 波、 $\beta$ 波、 $\theta$ 波のパワーに着目する方法では深層学習を用いて脳波を識別することは困難である可能性が示唆された。深層学習を用いて脳波から想起した音高を識別する際に高い識別率を得るためには、想起する音高の違いが脳波にどのような影響を与えるかについての検証をさらに行う必要があると考えられる。今後の課題として、本研究において解析対象としなかった $\gamma$ 波( $26\sim70[Hz]$ )、HFO(80[Hz]以上)などの高周波数成分について検証することが挙げられる。また、音楽の想起に関する脳内の働きは明確にはなっていないため、脳波の検証と並行して、解剖学・生理学的な観点から音楽想起時の脳内の働きを明らかにしていくことも重要であると考えられる。

表 2 α波のパワー (聴覚刺激:純音)

因子	平方和	自由度	平均平方	F値	p値
音高	46.04	2	23.022	1.62	0.2251
標本	1606.25	9	178.472	12.57	0
誤差	255.48	18	14.193		
全体	1907.77	29			

\* p < 0.05

表 3 β波のパワー (聴覚刺激:純音)

因子	平方和	自由度	平均平方	F値	p値
音高	11.784	2	5.8921	0.77	0.4793
標本	417.977	9	46.4419	6.04	0.0006
誤差	138.401	18	7.6890		
全体	568.162	29			

\* p < 0.05

表 4 θ波のパワー (聴覚刺激:純音)

因子	平方和	自由度	平均平方	F値	p値
音高	7.135	2	3.5677	1.49	0.2527
標本	840.736	9	93.4151	38.92	0
誤差	43.204	18	2.4002		
全体	891.076	29			

\* p < 0.05

表 5  $\alpha$ 波のパワー (聴覚刺激:ピアノの音)

因子	平方和	自由度	平均平方	F値	p値
音高	10.18	2	5.091	0.52	0.6018
標本	1678.34	9	186.483	19.14	0
誤差	175.41	18	9.745		
全体	1863.94	29			

\* p < 0.05

表 6  $\beta$  波のパワー (聴覚刺激:ピアノの音)

因子	平方和	自由度	平均平方	カイ二乗値	p値
音高	3.8	2	1.9	3.8	0.1496
誤差	16.2	18	0.9		
全体	20	29			

\* p < 0.05

表 7  $\theta$ 波のパワー (聴覚刺激:ピアノの音)

因子	平方和	自由度	平均平方	F値	p値
音高	4.542	2	2.271	0.26	0.7711
標本	467.906	9	51.9896	6.04	0.0006
誤差	154.971	18	8.6095		
全体	627.419	29			

\* p < 0.05

### 8. まとめ

人が異なる音高で音を想起したときの脳波を計測し、 深層学習に適用することで、脳波からどの音高の音を 想起していたか識別することが可能かどうかの検証を 行なった. 本研究では、 $\alpha$ 波、 $\beta$ 波、 $\theta$ 波を含む周波 数帯域のパワーの時間変化に着目し、計測データにウ ェーブレット変換を行い得られたスカログラム画像を CNN の入力データとして使用したが、高い識別率を 得ることはできなかった. 高い識別率とならなかった 原因として、CNN に入力したデータが不適切であっ た可能性, 脳波の前処理・解析の手法が不適切であっ た可能性, 学習データの数が不足していた可能性が考 えられる. そこで、CNN に入力したデータが不適切で あった可能性について焦点を当てて,  $\alpha$  波.  $\beta$  波,  $\theta$ 波のそれぞれのパワーが想起する音高による影響を受 けているかについて検証を行なった. 検定の結果, α 波、 $\beta$ 波、 $\theta$ 波それぞれのパワーは、想起する音高の 影響を受けているとは言えないことが明らかになった. 深層学習を用いて異なる音高を想起したときの脳波を 識別する課題において高い識別率を得るためには、想

起する音高の違いが脳波に与える影響についてさらに 検証を重ねることが必要であると考えられる。今後の 課題として、本研究で解析対象としたかった周波数帯 域や、脳波の生波形について検証することが挙げられ る。また、脳波の検証と並行して解剖学・生理学的な 観点から音楽の想起に関する脳内の働きを明らかにし ていくことも重要であると考えられる。

#### 参考文献

- T. Horikawa, et al., "Neural Decoding of Visual Imagery During Sleep", Science 340(6132), pp.639-642 (2013)
- (2) NICT 国立研究開発法人 情報通信機構,"映像を見て感じた内容を脳から言葉で読み解く脳情報デコーディング技術を開発 ~名詞・動詞だけでなく"印象"を形容詞の形で解読に成功~",
  - <a href="https://www.nict.go.jp/press/2017/11/01-1.html">https://www.nict.go.jp/press/2017/11/01-1.html</a> (2020 年 10 月 22 日確認)
- (3) Alison Abbott, "Music, maestro, please!", Nature 416, pp.12-14 (2002)
- (4) 安井拓也ら、"言語と音楽と脳科学", JOHNS Vol.27 No.8 (2011)
- (5) 澤隆史ら、"音楽を鑑賞する脳",情報処理 Vol.50 No.8 (2009)
- (6) 中村貴展ら, "音高が脳波に及ぼす影響",音楽情報科学 45-6 (2002)
- (7) 藤中雄大ら、"畳み込みニューラルネットワークを用いた多思考脳波判別手法の開発"、第33回ファジィシステムシンポジウム 講演論文集 (2017)
- (8) 及川大樹ら、"畳み込みニューラルネットワークを用いた脳波による音響信号再構成", IPSJ SIG Technical Report (2019)
- (9) 繁本成美ら、"fMRI を用いたヒト脳聴覚野からの音階デコードアルゴリズムの検討"、日本機械学会 中国四国支部 第57期総会・講演会(2019)
- - <a href="https://www.biosemi.com/headcap.htm">(2020年10月19日アクセス)</a>
- (11) Kahoko Takahashi, et al., "Empirical mode decomposition for improved EEG signal classification with Convolutional Neural Network in Brain-Computer interface experiments", The 28th Annual Conference of the Japanese Neural Network Society (2018)