

# 脳波解析における畳み込みニューラルネットワークの適用

曾我 洋史<sup>†</sup> 佐々木健太<sup>††</sup> 佐々木勇和<sup>††</sup> 平田 雅之<sup>†††</sup> 鬼塚 真<sup>††</sup>

<sup>†</sup> 大阪大学 〒 5650871 大阪府吹田市山田丘 2-1

<sup>††</sup> 大阪大学大学院情報科学研究科 〒 5650871 大阪府吹田市山田丘 1-5

<sup>†††</sup> 大阪大学国際医工情報センター 〒 5650871 大阪府吹田市山田丘 2-2

E-mail: †{soga.hirofumi,sasaki.kenta,sasaki,onizuka}@ist.osaka-u.ac.jp, ††mhirata@nsurg.med.osaka-u.ac.jp

あらし 現在、病気・けがなどによって運動や意思伝達に障害を抱える身体障害者のための生活支援技術として、脳波信号からロボットアームなどの外部機器を制御する技術の研究が行われている。ロボットアームの制御において、計測された皮質脳波信号から手の運動内容の推定を行うというタスクが存在する。先行研究ではガウス過程回帰により特徴量抽出を行い、SVM を用いて推定を行っているが、その推定精度は実用化に向けて十分でなく、精度向上が必要とされる。そこで本研究では、頭表脳波からの手の運動に関する推定において高い性能を発揮している畳み込みニューラルネットワーク (CNN) を皮質脳波からの運動内容推定に適用し、推定精度の向上を行う。

キーワード BMI, 脳波, CNN, 機械学習

## 1. はじめに

現在、ブレインマシンインタフェース (brain-machine-interface: BMI) という脳と機械の間で直接信号をやり取りする技術の研究が行われている [1]。BMI は出力型 BMI と入力型 BMI に分けられる。入力型 BMI はセンサで取得した外界の情報をコンピュータで変換し、脳に刺激を与えることにより、感覚情報を得る。出力型 BMI は脳信号を計測してコンピュータで解読を行い、信号が意味する脳機能の内容を推定し、ロボットアーム等の外部機器を制御することにより、神経機能の代替を行う。本研究では、身体障害者の運動・意思疎通の補助・代行のために研究開発が進められている出力型 BMI を対象とする。

BMI によるロボットアーム制御は主に身体障害者の生活支援を目的としている。対象疾患としては筋萎縮性側索硬化症、筋ジストロフィーなどの神経難病、脊髄損傷、頸椎損傷、切断肢、脳卒中後遺症などがあげられる。ALS 患者の数は国内で 8000 人、世界で 35 万人であり、規模は大きくないが、脊髄損傷患者は国内 10 万人、脳卒中後遺症患者は国内 150 万人と規模は非常に大きい。今後 BMI によるロボットアーム制御の性能が向上して思い通りに制御することが可能になれば、脊髄損傷や脳卒中後遺症による麻痺を持つ患者にも利用されていくと考えられる。

出力型 BMI における脳波の計測は電極を体の計測したい部位に配置し、電位を計測することによって行われる。頭蓋内電極を用いる BMI は頭蓋内の脳波を計測することができ、使用される電極により刺入針電極型と脳表電極型に分けられる。刺入針電極型では微小な針電極を刺入することにより神経細胞のスパイク活動やいくつかの神経細胞の集合電位が計測される。上肢の運動野の神経細胞は特定の運動方向で特異的に発火頻度が高まる特性があり、少数の神経細胞のスパイク活動を計測することで運動方向を推定することができる [2]。しかし、刺入針電極型は炎症反応により計測効率が低下するという問題があ

る [3]。脳表電極型では脳表面に皿状の電極を留置し脳波を計測する。脳表電極を用いて計測される脳波は皮質脳波と呼ばれ、頭皮脳波に対してノイズが少なく、高周波帯域まで計測が可能であるという特徴がある。脳表電極型は脳実質への侵襲が比較的少なく、長期安定性に優れていることが示されている [4]。皮質脳波を用いた研究により、運動内容の解読において high  $\gamma$  活動と呼ばれる 80–200Hz の高周波帯域の皮質脳波が脳機能所在を正確に反映することが明らかになっている [5] [6]。

ロボットアーム操作の実用化に向けて、計測機器、通信機器、脳波の解読、ロボットアームの制御など数多くの課題があり、現在研究が進められている。脳波の解読では脳波を分析し、どのような動作を行っているか推定を行う。柳沢ら [7] はガウス過程回帰と SVM を用いた手法により推定を行っているが、その推定精度は実用化に向けて十分ではなく、精度向上が必要とされる。脳波を利用した推定や分類は他の研究でも行われており、ニューラルネットワークを用いたモデルが高い性能を発揮している [8] [9]。本研究では脳表電極で計測された皮質脳波から体の動きを再現することを目標として、皮質脳波から運動内容の推定を行う。推定の方法として、頭表脳波からの手の運動に関する推定において高い性能を発揮している畳み込みニューラルネットワークを用いる。運動想起時刻前後の脳波信号を入力とし、運動内容の推定結果を出力するよう畳み込みニューラルネットワークを設計した。脳波信号に対して畳み込み処理等を行うことにより特徴量の抽出を行うが、フィルタの重みを機械的に学習することにより、前処理のコストを減らすとともに推定精度の向上を図る。

本稿の構成は以下の通りである。2. 章で関連研究について概説する。3. 章でデータセットについての説明を行い、4. 章で手法の詳細について説明する。5. 章において手法の評価と分析を行い、6. 章にて本稿をまとめ、今後の課題について論ずる。

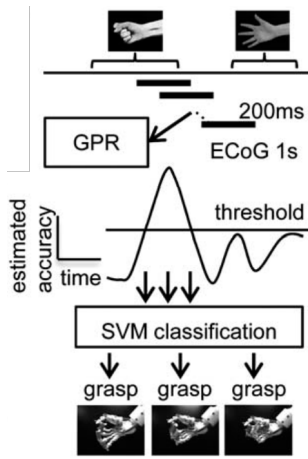


図 1 SVM decoder ([7] より改変)

## 2. 関連研究

### 2.1 ガウス過程回帰と SVM を用いた脳波解析

柳沢ら [7] は図 1 のようなアルゴリズムにより運動内容の推定を行っている。推定は 2 段階で行われる。まず、ガウス過程回帰 (GPR) で特徴量抽出を行う。学習用データでデータの分布モデルを作成し、200ms 分の脳波信号入力に対し、分布モデルとの差異を計算する。この差異を相互情報量と呼び、特徴量とする。その後、特徴量が域値を超えたかどうかを SVM により判定を行い、動作推定を行う。この手法では、脳波から運動内容を推定することは可能であると示されたが、その推定精度は実用化に向けて十分でなく、精度向上が必要とされる。

### 2.2 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 (CNN) は主に画像認識に利用されるニューラルネットワークの 1 つである。CNN の原型は、生物の脳の視覚野に関する神経生理学的な知見 [10] を元に、福島 [11] によって考案された Neocognitron に見ることが出来る。Neocognitron は、simple cells と complex cells を交互に階層的に配置したニューラルネットワークである。最初の CNN は LeCun ら [12] により提案された。Neocognitron は自己組織化による学習を行っていたのに対し、CNN では backpropagation アルゴリズムに基づく学習法が確立された。その後、文字認識において LeNet [13] が成功を収めた。

CNN は画像分類だけでなく、セグメンテーション [14] [15] や物体認識 [16] [17] などのタスクにおいても利用されている。また、画像以外にも、自然言語処理 [18] [19] や音声信号処理 [20] などの分野においても幅広く利用されている。

### 2.3 ニューラルネットワークを用いた脳波分析

kaggle において頭表脳波から手の運動に関するラベルを推定する問題が取り上げられている [21]。その中で精度の高い上位 3 つの手法では、いずれもニューラルネットワークが利用されている。第 1 位のチームは CNN と再帰的ニューラルネットワーク (RNN) を段階的に組み合わせて利用している。第 2 位のチームは CNN と RNN を融合させた RCNN [22] を使用して

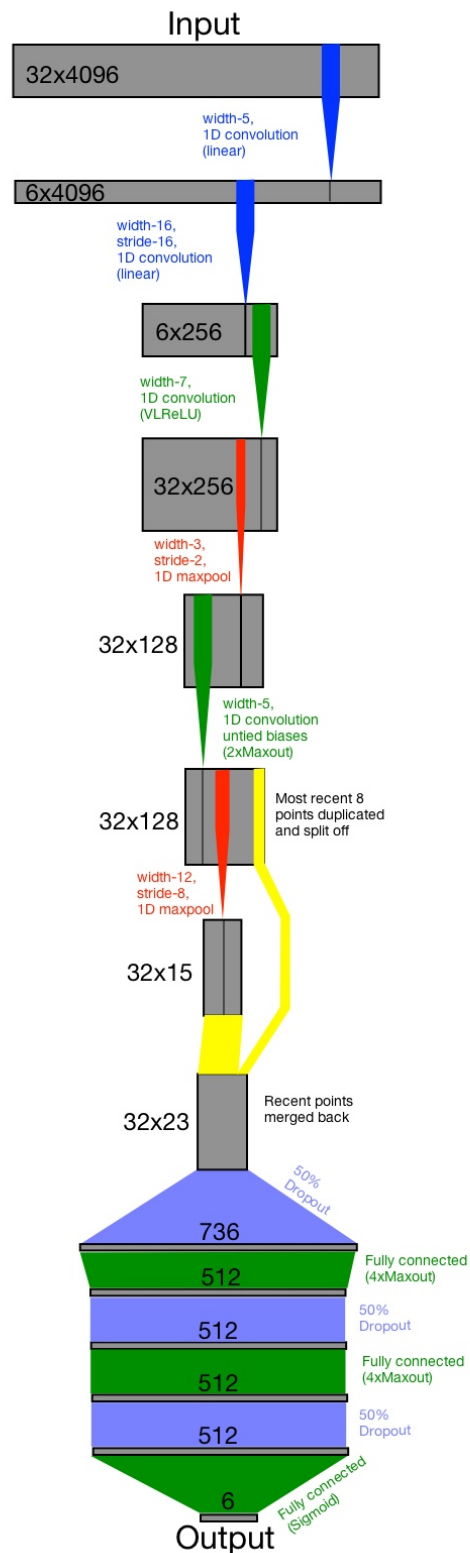


図 2 stf7 ネットワークの構造

(<http://blog.kaggle.com/2015/10/05/grasp-and-lift-ecg-detection-winners-interview-3rd-place-team-hedj/> より引用)

いる。第 3 位のチームは多数の CNN をアンサンブル学習させて使用している。これら 3 つの手法についてラベル推定の精度は僅差であり、脳波分析においてどの手法を使用すべきかは明

らかではない。3位のCNNを使用した手法[23]の概要を以下に示す。

3位の手法では特別な前処理は行っておらず、CNNを用いて脳波信号から運動ラベルの推定を行う。畳み込みニューラルネットワークの1つであるstf7ネットワークの構造を図2に示す。データセットを8秒ごとにスライスし、32チャンネル×4096タイムポイントを入力とする。複数の畳み込みニューラルネットワークの出力の加重平均を行い、最終的な予測結果としている。

- stf7 ネットワークの構成
  - 第1層：畳み込み層
  - 第2層：畳み込み層
  - 第3層：畳み込み層
  - 第4層：プーリング層
  - 第5層：畳み込み層
  - 第6層：プーリング層
  - 第7層：全結合層
  - 第8層：全結合層
  - 第9層：全結合層

stf ネットワークの多くは畳み込み層とプーリング層であり、画像認識において一般的に用いられるネットワークの構造に類似しているが、次の2つの領域が特徴的である。

#### 1. 初期2つの畳み込み層

第1の畳み込み層では、チャンネル数を32から6に減らす。これにより、ネットワークはデータのノイズを減らすための時空間フィルタを学習する。第2の畳み込み層では時間方向に対してストライドを16とし、ダウンサンプリングのためのフィルタが学習される。

#### 2. 第6層のプーリング層とそれに伴うバイパス

このプーリング層ではストライドを8に設定し過学習を防ぐが、イベントの開始位置を曖昧にするという副作用がある。そのため、直近8つのポイントを複製してプーリングをバイパスすることにより、過学習を防ぎつつイベントの開始位置を正確に推定する。

### 3. データセット

脳波データセットは計測手法の違いにより、計測部位、チャンネル（電極）数、サンプリング周波数、計測精度などにおいて異なる。2つのデータの差異を表1に示す。

#### 3.1 頭表脳波

kaggleの手法で使用されているデータは、図3の頭皮電極を用いて計測されたサンプリング周波数500Hz、32チャンネルの頭表脳波のデータセットである。12人の被験者が物体に手を伸ばし、持ち上げ、元に戻すという一連の動作を繰り返し行う。32極の頭皮型電極を用いて、この運動時の頭皮電圧を測定している。また、運動内容に対応して、HandStart, FirstDigitTouch, BothStartLoadPhase, LiftOff, Replace, BothReleasedの6つのイベントが設定されている。これらのイベントの有無を(0, 1)で表したものを正解ラベルとしている。なお、各イベントは1-hotではなく、複数イベントが共起する時刻やイベン

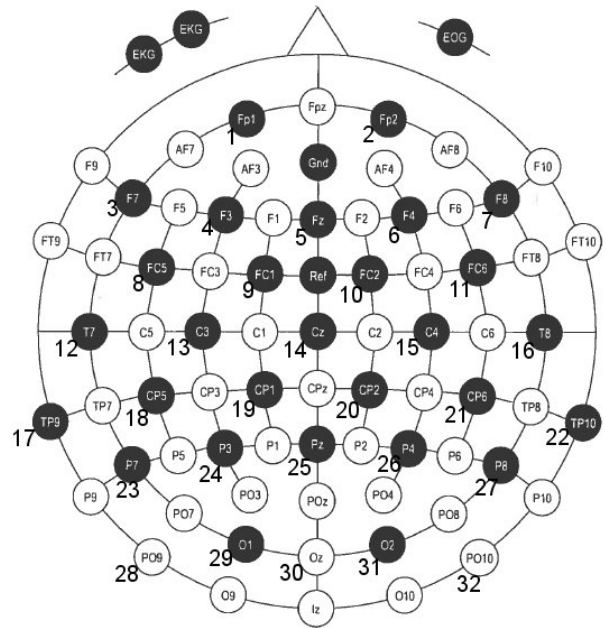


図3 頭表脳波の電極配置

(<https://www.kaggle.com/c/grasp-and-lift-eeeg-detection/data> より引用)

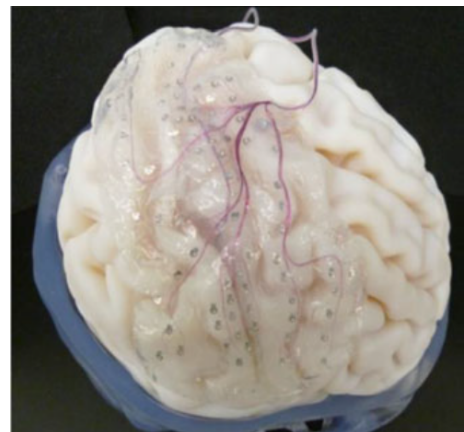


図4 皮質脳波の電極配置 ([24]より改変)

トのない時刻が存在する。

#### 3.2 皮質脳波

本稿で使用するデータは、図4の脳表電極を用いて計測されたサンプリング周波数1000Hz、94チャンネルの皮質脳波のデータセットである[24]。電極は手の運動を反映しているとされる脳の1次運動野付近に集中的に配置されている。脳表電極を用いて、ALS患者1名の運動想起時の脳表電圧を測定する。データは2週間のうち8日間計測されている。手と肘について運動想起課題が設定されており、画面上の指示に従い運動を想起する。この指示により運動を想起したとされる時刻を同時に記録する。手について、握る/開くという運動想起課題が1日当たり40回ずつ計80回行われる。肘についても、曲げる/伸ばすという運動想起課題が40回ずつ計80回行われる。運動想起の種類”握る・開く”, ”曲げる・伸ばす”を正解ラベルとした。

表 1 脳波データの特徴

	頭表脳波	皮質脳波
タスク	動作検出	動作方向の推定
電極	頭表面	脳表面
位置	全体	1次運動野
サンプリング周波数 [Hz]	500	1000
電極 (チャンネル) 数	32	94
ノイズ	多い	少ない

## 4. 手 法

皮質脳波から運動内容を推定する手法について説明する。本研究では、stf7 ネットワークを参考に、畳み込みニューラルネットワーク (CNN) を構築し、皮質脳波から運動内容の推定を行う。前処理として、運動想起時刻毎にその前後の脳波信号を切り取り、切り取った時刻毎に標準化を行い、チャンネル数×時間の2次元ベクトルとして扱ひ入力とする。今回はデータ数が少なく、過学習を防ぐためにモデルを小さくする必要があり、5層のネットワークとした。構築したネットワークは畳み込み層、プーリング層、そして全結合層から構成される。

- 提案手法のネットワークの構成
  - 第1層：畳み込み層
  - 第2層：畳み込み層
  - 第3層：プーリング層
  - 第4層：全結合層
  - 第5層：全結合層

### 畳み込み層

第1, 第2の畳み込み層では線形畳み込みを行う。線形畳み込みではチャンネル数×フィルタサイズのフィルタを用いて畳み込みを行う。1層目の線形畳み込みでは入力のチャンネル数の削減を行う。今回は6種類のフィルタを使用し、電極数94チャンネルの入力から6チャンネルの特徴量に変換する。2層目の畳み込み層では畳み込みフィルタの適用間隔をフィルタサイズと同じ幅に変更し、フィルタサイズ分の特徴量を1つの特徴量にマッピングすることによりダウンサンプリングを行う。

### プーリング層

プーリング層では線形マックスプーリングを行う。時間方向に3タイムポイントの値をマックスプーリングにより集約する。ストライド幅を2とすることで集約される領域にオーバーラップが生じる。AlexNet [25] ではこのオーバーラッピングプーリングにより、過学習を低減し、わずかに精度が向上すると主張されている。

### ドロップアウト

ドロップアウト [26] は学習時のネットワークについて、隠れ層のニューロンを一定確率で無効化する手法である。これにより疑似的に毎回異なるアーキテクチャで学習を行うことにより、アンサンブル学習と同様の効果をもたらす、汎化性能を高める。今回は kaggle の手法と同様に2つの全結合層の入力前でそれぞれ50%ドロップアウトを行う。

### 全結合層

表 2 提案手法のネットワークのパラメータ

Type	Filters	Size/Stride	Output
Convolutional	6	7	3000 × 6
Convolutional	6	10/10	300 × 6
Maxpool		3/2	149 × 6
Fully Connected			10
Fully Connected			1

input = 3000 timepoint × 94 チャンネル のとき

2つの全結合層で特徴マップを集約し、最終的な出力を得る。

## 5. 実 験

### 5.1 前 処 理

本実験では、皮質脳波データセットについて簡単な前処理を行った。まず、運動想起時刻を基に、脳波信号の切り出しを行った。切り出し幅等については5.2節で説明する。次に、切り出した信号ごとに式1で表される標準化を行った。

$$x_{std}^{(i)} = \frac{x^{(i)} - \mu}{\sigma} \quad (\mu : \text{平均}, \sigma : \text{標準偏差}) \quad (1)$$

### 5.2 推定精度の評価

畳み込みニューラルネットワーク (CNN) を用いて皮質脳波データセットから運動内容を推定した際の推定精度について評価を行った。今回、提案手法の CNN の実装は Keras (2.1.3) と Tensorflow-gpu (0.12.1) で行った。パラメータを表2に示す。

8日間のデータのうち最終日のデータをテストデータとし、その他のデータでネットワークの学習を行い、手の運動 (握る/開く) と肘の運動 (曲げる/伸ばす) についてそれぞれ2値分類を行った。分類精度の評価指標として式2で表される Accuracy を利用した。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

切り出し幅を変更しながら5回ずつ実験を行った結果を図5,6に示す。CNNにおける推定精度は95%信頼区間で、手の運動について  $93.00 \pm 1.18\%$ 、肘の運動について  $79.00 \pm 6.16\%$  という結果が得られた。共同研究者が行ったガウス過程回帰 (GPR) と SVM による実験結果<sup>(注1)</sup>は手の運動について81.7%、肘の運動について73.8%の推定精度となっている。CNNを用いた手法により、GPR・SVMを用いた手法に対して、手の運動において11.3%、肘の運動において5.2%推定精度の向上を行うことができた。

実験結果では、肘の運動の推定精度は手の運動の推定精度に対し、全体的に低くなっている。医学的知見 [27] によると、肘の運動野は手の運動野に比べて脳の領域が小さいことが分かっており、肘の運動の推定は手の運動の推定と比較して難しいと考えられる。また、手の運動は複数のチャンネルに表れており、肘の運動は特定のチャンネルに表れていると仮定すると、手と肘を異なるモデルで推定する方法等が考えられる。また、切り出し幅については、GPR・SVMの実験において運動想起時刻

(注1) : GPR・SVMの実験は提案手法の実験とテストデータの設定、データの切り出し方法等、実験条件が一部異なる

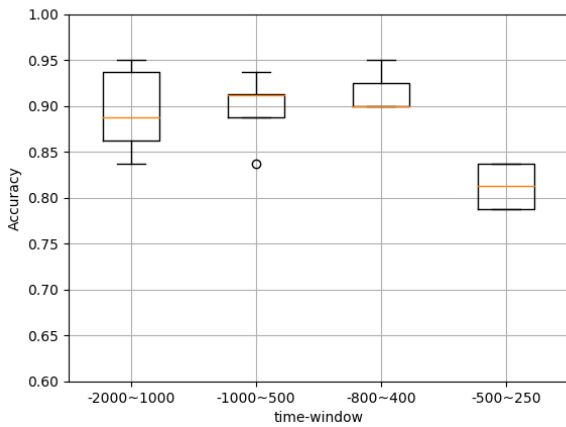


図 5 手の動作推定の精度

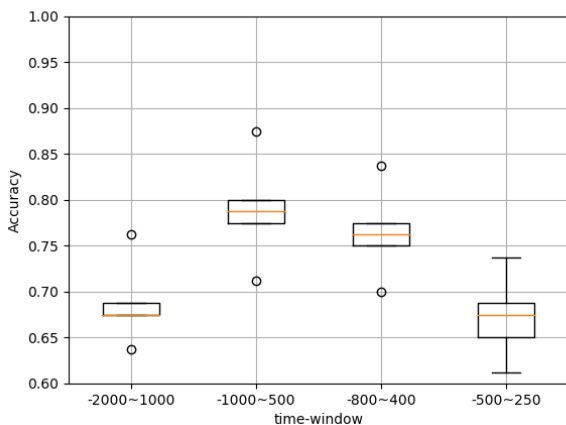


図 6 肘の動作推定の精度

の少し前において推定精度が最も高くなっていたため、今回の設定とした。手と肘共に運動想起時刻の前 1000ms・後 500ms のときが最も高い結果となった。これは他の幅が、運動の特徴量が現れている幅よりも短いことや、長過ぎて運動と関係ない情報を踏んでいることが考えられる。切り出しの幅とタイミングについては、検討する必要がある。

## 6. ま と め

本稿では皮質脳波からの運動内容の推定において畳み込みニューラルネットワークを構築し、推定精度についての実験を行った。その結果、推定精度において、先行研究の結果を上回る推定精度を達成した。今後の課題としては、推定精度のさらなる向上のため、データオーグメンテーションや他の手法の検討が考えられる、また、より自在なロボットアームの制御のため、手と肘の運動の同時分類（多クラス分類・マルチラベル推定）も行っていくたい。

## 7. 謝 辞

本研究開発は総務省「次世代人工知能技術の研究開発」の委託を受けたものです。

- [1] M. Hirata, "Implantable brain-machine interfaces," 神経治療, vol.33, p.399404, 2016.
- [2] M. Velliste, S. Perel, M.C. Spalding, A.S. Whitford, and A.B. Schwartz, "Cortical control of a prosthetic arm for self-feeding," Nature, vol.453, no.7198, pp.1098–1101, 2008.
- [3] E. Fernández, B. Greger, P.A. House, I. Aranda, C. Botella, J. Albiusua, C. Soto-Sánchez, A. Alfaro, and R.A. Normann, "Acute human brain responses to intracortical microelectrode arrays: challenges and future prospects," Frontiers in Neuroengineering, vol.7, no. July, pp.1–6, 2014.
- [4] Chao, "Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkey," Frontiers in Neuroengineering, vol.3, no. March, 2010.
- [5] N.E. Crone, A. Sinai, and A. Korzeniewska, "High-frequency gamma oscillations and human brain mapping with electrocorticography," Progress in Brain Research, vol.159, pp.275–295, 2006.
- [6] T. Yanagisawa, M. Hirata, Y. Saitoh, A. Kato, D. Shibuya, Y. Kamitani, and T. Yoshimine, "Neural decoding using gyral and intrasulcal electrocorticograms," NeuroImage, vol.45, no.4, pp.1099–1106, 2009.
- [7] T. Yanagisawa, M. Hirata, Y. Saitoh, H. Kishima, K. Matsushita, T. Goto, R. Fukuma, H. Yokoi, Y. Kamitani, and T. Yoshimine, "Electrocorticographic control of a prosthetic arm in paralyzed patients," Annals of Neurology, 2012.
- [8] K. Sivasankari, and K. Thanushkodi, "An Improved EEG Signal Classification Using Neural Network with the Consequence of ICA and STFT," J Electr Eng Technol, vol.9, no.3, pp.1060–1071, 2014.
- [9] R.T. Schirrmester, L. Gemein, K. Eggensperger, F. Hutter, and T. Ball, "Deep learning with convolutional neural networks for decoding and visualization of EEG pathology," Human Brain Mapping, vol.38, pp.5391–5420, 2017.
- [10] T. Inoue, H. Miyakawa, K. Ito, K. Mikoshiba, and H. Kato, "RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX," Neuroscience research, vol.15, no.4, pp.265–271, 1992.
- [11] K. Fukushima, and S. Miyake, "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position," Pattern Recognition, vol.15, no.6, pp.455–469, 1982.
- [12] Y. Le Cun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel, "Handwritten Digit Recognition with a Back-Propagation Network," Advances in Neural Information Processing Systems, pp.396–404, 1990.
- [13] 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.2278–2324, 1998.
- [14] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.07-12-June, pp.3431–3440, 2015.
- [15] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," arXiv preprint arXiv:1606.00915, 2016.
- [16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," Advances in neural information processing systems, pp.91–99, 2015.
- [17] J. Redmon, and A. Farhadi, "YOLO9000: Better, Faster, Stronger," arXiv preprint, vol.1612, pp.6517–6525, 2016.
- [18] Y. Kim, "Convolutional Neural Networks for Sentence Classification," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP),



pp.1746–1751, 2014.

- [19] X. Zhang, J. Zhao, and Y. LeCun, “Character-level convolutional networks for text classification,” *Advances in neural information processing systems*, pp.649–657, 2015.
- [20] A.v.d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, “WaveNet: A Generative Model for Raw Audio,” <http://arxiv.org/abs/1609.03499>, pp.1–15, 2016.
- [21] “<https://www.kaggle.com/c/grasp-and-lift-eeg-detection>,” .
- [22] M. Liang, and X. Hu, “Recurrent Convolutional Neural Network for Object Recognition,” *IEEE*, no.Figure 1, pp.3367–3375, 2015.
- [23] “[https://github.com/bitsofbits/kaggle\\_grasp\\_and\\_lift\\_eeg\\_detection](https://github.com/bitsofbits/kaggle_grasp_and_lift_eeg_detection),” .
- [24] S. Morris, M. Hirata, H. Sugata, T. Goto, K. Matsushita, T. Yanagisawa, Y. Saitoh, H. Kishima, and T. Yoshimine, “Patient-Specific Cortical Electrodes for Sulcal and Gyral Implantation,” *IEEE Transactions on Biomedical Engineering*, vol.62, no.4, pp.1034–1041, 2015.
- [25] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances In Neural Information Processing Systems*, pp.1–9, 2012.
- [26] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research*, vol.15, pp.1929–1958, 2014.
- [27] W. Penfield, and E. Boldrey, “Somatic Motor and Sensory Representation in Man,” *Brain*, pp.389–443, 1937.