

時間・空間・周波数を統合した深層状態空間モデルによる 脳波からの行動予測

A Deep State-Space Model Integrating Time, Space, and Frequency
for Predicting Actions from EEG/ECoG

鈴木 駿太郎 *1
Shuntaro Suzuki

長嶋 隼矢 *1
Shunya Nagashima

平田 雅之 *2
Masayuki Hirata

杉浦 孔明 *1
Komei Sugiura

*1 慶應義塾大学
Keio University

*2 大阪大学
Osaka University

Classification of electroencephalogram (EEG) / electrocorticogram (ECoG) signals during motor-imagery (MI) exhibits significant application potential, including communication assistance and rehabilitation support for patients with motor impairments. However, these signals remain inherently susceptible to physiological artifacts (e.g., eye blinking, swallowing), posing persistent challenges in achieving high classification accuracy with existing approaches. To overcome these limitations, we propose a novel architecture that extends deep state-space models to capture integrated dependencies of EEG/ECoG signals across temporal, spatial, and frequency domains. For evaluation, we conduct comprehensive evaluation across two benchmarks: 1) A large-scale public MI EEG datasets, containing >50 subjects, 2) A clinical MI ECoG dataset recorded from an amyotrophic lateral sclerosis patient. Across these two benchmarks, our method outperformed baseline methods on standard evaluation metrics. Furthermore, visual explanations derived from our model indicate that it effectively captures neurophysiologically significant regions in both EEG/ECoG signals.

1. はじめに

ブレイン・コンピュータ・インターフェース (BCI) は、脳活動の計測を通じて、コンピュータやロボット等の外部機器との直接的な通信を可能とするシステムの総称である [Hramov 21]. その応用範囲は広く、神経変性疾患の診断や脳機能マッピングの深化等、多岐にわたる。BCIは、対象とする脳波応答に基づき、Active, Reactive, および Passive BCI に分類される [Hramov 21]. 特に、Active BCI は被験者の自発的な脳活動を解析対象とすることから、身体麻痺患者に対するコミュニケーション支援やリハビリテーション支援への応用が期待されている。そこで、本研究では、Active BCI において重要な運動想起時の脳活動に着目し、頭皮脳波および頭蓋内脳波を用いて、想起された行動の多クラス分類タスクを扱う。

ここで、図 1 に、本タスクの代表例を示す。この例では、被験者が「肘の伸展」を想起した際の頭皮脳波または頭蓋内脳波が入力として与えられ、モデルは各行動に対応する確率の予測値を出力する。ここでは、「肘の伸展」クラスに対する確率の予測値が最も大きいことが望ましい。本タスクにおいて、運動想起に関連する脳波応答は、活性領域や変動する周波数帯域が被験者ごとに異なることが知られており [Pfurtscheller 01], 高精度な分類が困難である。実際に、運動想起時頭皮脳波の 2 クラス分類タスクを扱う OpenBMI [Lee 19] データセットにおいて、代表的な脳波分類モデルである EEGNet [Lawhern 18] は約 22.1% の分類を誤る。

そこで、本研究では、時間・空間・周波数領域における頭皮脳波および頭蓋内脳波の統合的な依存関係を捕捉可能な新しい手法を提案する。さらに、時間・空間・周波数の各領域に対する視覚的説明生成を可能とする。

2. 問題設定

本研究では、頭皮脳波および頭蓋内脳波を用いた、運動想起時における多クラス分類タスクを扱う。本タスクでは、被験者

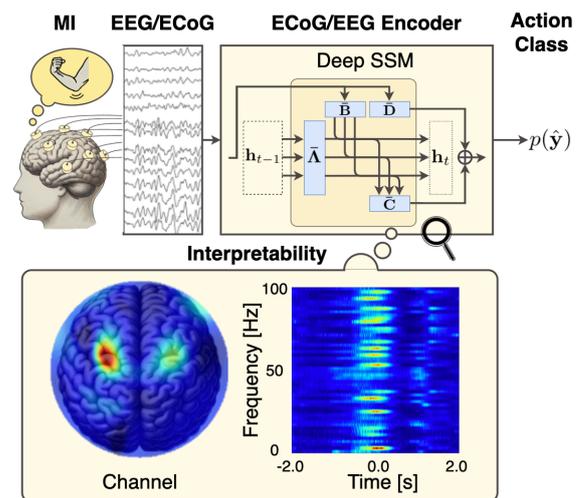


図 1: 運動想起時の頭皮脳波・頭蓋内脳波を用いた行動分類

が任意の行動を想起した際に、対応する頭皮脳波および頭蓋内脳波を正しく分類できることが望ましい。ここで、本タスクの入力は、被験者が運動を想起する際に計測された頭皮脳波および頭蓋内脳波であり、出力は想起された行動に対応する確率の予測値である。なお、本研究で扱う頭皮脳波および頭蓋内脳波は運動想起時に計測された信号を前提とし、その他の脳波応答 (例: 視覚応答, 事象関連応答) は対象としない。また、評価指標として、分類精度および Macro-F1 を使用する。

3. 提案手法

提案手法では、深層状態空間モデルを拡張し、脳波分類における時間・空間・周波数特徴の統合的な依存関係を捉える。さらに、各領域に対する視覚的説明生成を可能とする。なお、提案手法における、時間・空間・周波数特徴の依存関係を統合的に捉えるアプローチは、脳波を用いた他の外部および内部刺激のデコーディングに対しても適用可能であると考えられる。加えて、多変量時系列予測の手法にも広く適用可能である

連絡先: 鈴木駿太郎, 慶應義塾大学, 神奈川県横浜市港北区日吉 3-14-1, shuntaro20021227@keio.jp

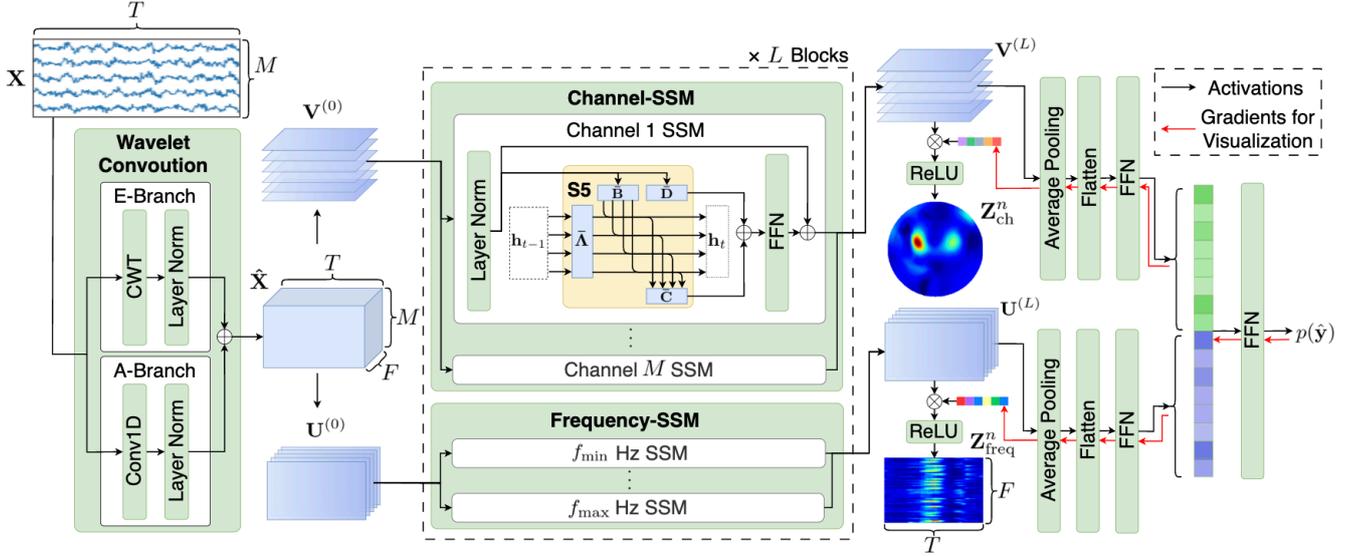


図 2: 提案手法のモデル構造

と考えられる。提案手法は3つの主要モジュールから構成され、それぞれ Wavelet-Convolution, Frequency-SSM, および Channel-SSM である。図2にモデルの構造を示す。

3.1 入力

本モデルへの入力は、頭皮脳波および頭蓋内脳波の計測信号

$$\mathbf{X} = \left\{ \mathbf{x}_m \mid \mathbf{x}_m \in \mathbb{R}^T, m = 1 \dots M \right\} \in \mathbb{R}^{M \times T}.$$

である。ここで、 M および T は電極数および系列長を表す。

3.2 Wavelet-Convolution

Wavelet-Convolution では、決定的に得られた周波数成分および適応的に得られた周波数特徴を統合する。先行研究より、深層学習に基づく脳波分類では、1次元畳み込み層におけるカーネルを適応的に更新することで周波数領域の特徴を抽出するアプローチが広く用いられている [Lawhern 18, Song 22]。しかし、これらの手法により得られた周波数特徴は解析が困難であり、直接的に周波数領域の説明性が得られない。一方で、短時間フーリエ変換やウェーブレット変換に代表される伝統的な時間・周波数領域の特徴抽出手法を用いた解析では、決定的に周波数成分が抽出されるため、直接的に説明性が取得可能である。しかし、これらの手法では高精度な分類が困難である。

そこで、本モジュールでは Explanation Branch (E-Branch) により決定的に得られた周波数成分および Adaptation Branch (A-Branch) により適応的に得られた周波数特徴を統合する。これにより、周波数解析可能な特徴量を抽出しつつ、頭皮脳波および頭蓋内脳波のより高精度な分類が可能となる。ここで、 \mathbf{x}_m に対する出力 $\tilde{\mathbf{x}}_m \in \mathbb{R}^{F \times T}$ は以下で定式化され、 $\tilde{\mathbf{X}} = \{\tilde{\mathbf{x}}_m \mid m = 1 \dots M\} \in \mathbb{R}^{M \times F \times T}$ が得られる。

$$\tilde{\mathbf{x}}_m = \underbrace{\frac{1}{2} \text{LayerNorm}(\text{CWT}(\mathbf{x}_m))}_{\text{E-Branch}} + \underbrace{\frac{1}{2} \text{LayerNorm}(\text{Conv1D}(\mathbf{x}_m))}_{\text{A-Branch}}. \quad (1)$$

なお、 $\text{Conv1D}(\cdot)$, $\text{CWT}(\cdot)$, および $\text{LayerNorm}(\cdot)$ は1次元畳み込み層、連続ウェーブレット変換およびレイヤー正規化をそれぞれ表す。また、 F は周波数領域の次元数を表す。

3.3 Frequency-SSM

Frequency-SSM では、周波数成分ごとに独立して時間・空間特徴を捉える。頭皮脳波および頭蓋内脳波では、運動想起の前後で、特定の脳領域における一部の周波数成分において顕著な信号強度の変動が知られている [Pfurtscheller 01]。そこで、本モジュールでは、周波数成分ごとに独立して時間・空間特徴の依存関係を捉える。これにより、運動想起に関連する周波数成分ごとの強度変動を捕捉することが期待される。本モジュールは、 L 個のブロックからなる階層構造を持ち、各ブロックはレイヤー正規化、Feed-Forward Network, および深層状態空間モデルから構成される。以降では l 番目のブロック ($l = 0, \dots, L$) における入力 $\mathbf{U}^{(l)}$ に対する処理の詳細を述べる。ただし、 $\mathbf{U}^{(0)} = \tilde{\mathbf{X}}$ である。

はじめに、次式より $f \in \{1, \dots, F\}$ に関して $\mathbf{U}^{(l)}$ を時系列方向に正規化する。

$$\tilde{\mathbf{u}}_f^{(l)} = \text{LayerNorm}(\mathbf{u}_f^{(l)}). \quad (2)$$

次に、 $\tilde{\mathbf{u}}_f^{(l)}$ に対する時系列方向の依存関係を捕捉する。脳波における時系列依存関係の捕捉においては、EEG Conformer [Song 22] をはじめ、Transformer ベースのアプローチが広く用いられている。一方で、長系列信号に対する時系列依存関係の捕捉では、深層状態空間モデルにおいて Transformer を上回る結果が報告されている [Smith 23]。そのため、本モジュールでは深層状態空間モデルの一種である S5 [Smith 23] を組み込み、頭皮脳波および頭蓋内脳波における時系列方向の依存関係を捉える。ここで、離散時間システムにおける S5 の操作を $\text{SSM}(\cdot)$ と定義する。これにより、 $\tilde{\mathbf{u}}_f^{(l)}$ における時系列方向の依存関係を捉えた特徴量 $\mathbf{u}_f^{(l+1)}$ は次式として得られる。

$$\mathbf{u}_f^{(l+1)} = \text{FFN} \left(\text{SSM} \left(\tilde{\mathbf{u}}_f^{(l)} \right) \right) + \tilde{\mathbf{u}}_f^{(l)}. \quad (3)$$

ここに、 $\text{FFN}(\cdot)$ は Feed-Forward Network を表す。そして、 l 番目のブロックにおける本モジュールの出力 $\mathbf{U}^{(l+1)} = [\mathbf{u}_1^{(l+1)}, \mathbf{u}_2^{(l+1)}, \dots, \mathbf{u}_F^{(l+1)}] \in \mathbb{R}^{M \times F \times T}$ が得られ、周波数成分ごとに独立して時系列依存関係が捉えられる。

3.4 Channel-SSM

Channel-SSM では、電極ごとに独立して時間・周波数特徴を捉える。ここでは、Frequency-SSM では捉えられない電極ごとの時間・周波数特徴の依存関係を捉えることで、運動想起に関連

表 1: OpenBMI [Lee 19] および ECoG-ALS における定量的結果.

手法	OpenBMI		ECoG-ALS	
	Accuracy[%]↑	Macro-F1[%]↑	Accuracy[%]↑	Macro-F1[%]↑
Chance Level	50.00	50.00	25.00	25.00
EEGNet [Lawhern 18]	74.93±4.48	74.80±4.52	65.47±6.64	63.33±10.07
Informer [Zhou 21]	73.92±5.19	73.71±5.25	59.45±7.70	58.32±8.69
EEG Conformer [Song 22]	77.90±5.27	77.78±5.33	64.06±9.70	62.41±12.14
PathcTST [Nie 23]	76.80±4.73	76.73±4.77	58.91±5.88	57.10±7.23
iTransformer [Liu 24]	69.83±3.86	69.75±3.88	56.09±8.74	53.84±11.43
[長嶋 24]	72.40±6.17	72.31±6.22	60.78±5.37	58.73±9.23
提案手法	81.62±5.20	81.56±5.24	75.08±9.69	73.87±12.84
	(+3.72)	(+3.78)	(+9.61)	(+10.54)

する局在的な脳領域の信号強度変動を検出する [Pfurtscheller 01]. 本モジュールは, Frequency-SSM と同様, L 個のブロックからなる階層構造を持ち, 各ブロックはレイヤー正規化, Feed-Forward Network, および深層状態空間モデルから構成される. ここで, l 番目のブロック ($l = 0, \dots, L$) における入力 $\mathbf{V}^{(l)}$ は次式にて変換され, $\mathbf{V}^{(l+1)} = [\mathbf{v}_1^{(l+1)}, \mathbf{v}_2^{(l+1)}, \dots, \mathbf{v}_M^{(l+1)}] \in \mathbb{R}^{M \times F \times T}$ が得られる. なお, $\mathbf{V}^{(0)} = \tilde{\mathbf{X}}$ である.

$$\mathbf{v}_m^{(l+1)} = \text{FFN} \left(\text{SSM} \left(\text{LayerNorm} \left(\mathbf{v}_m^{(l)} \right) \right) \right) + \text{LayerNorm} \left(\mathbf{v}_m^{(l)} \right). \quad (4)$$

3.5 出力

Frequency-SSM および Channel-SSM の出力 $\mathbf{U}^{(l)}$ および $\mathbf{V}^{(l)}$ は, 次式にて統合され \mathbf{X} に対応する行動の確率予測値 $p(\hat{\mathbf{y}})$ が得られる. なお, N および AvgPool(\cdot) は想起する行動の種類および平均プーリング層をそれぞれ表す.

$$p(\hat{\mathbf{y}}) = \text{FFN} \left(\left[\text{AvgPool} \left(\mathbf{U}^{(l)} \right); \text{AvgPool} \left(\mathbf{V}^{(l)} \right) \right] \right). \quad (5)$$

そして, 損失関数は交差エントロピー誤差関数を使用した.

4. 実験設定

本研究では, 運動想起時の頭皮脳波および頭蓋内脳波データセットを用いた. 頭皮脳波については, 公開データセットである OpenBMI [Lee 19] を使用し, 頭蓋内脳波については, ALS を患った被験者から運動想起時に計測された臨床データセット (以降 ECoG-ALS とする) を使用した.

4.1 OpenBMI データセット

OpenBMI は健常者 54 名から計測された運動想起時頭皮脳波の公開データセットである [Lee 19]. 本データセットに含まれるタスクは, 右手および左手の掌握の 2 種類の運動想起から構成される. 被験者は 2 回のセッションに参加し, 1 名あたり 400 回の運動想起を行った. そして, 54 名の被験者から計 21600 サンプルが計測された. 各サンプルにおけるサンプリング周波数は 1000Hz であり, 注視点としての十字が 3 秒間提示された後に, 4 秒間の運動想起および 6 秒間 (± 1.5 秒) の安静が記録された. 計測においては, 国際 10-20 システムに則って 62 個の銀-塩化銀電極が配置された. なお, 本研究では運動想起時の 4 秒間を使用した. そして, データセットの分割では, k -fold 交差検証 ($k = 8$) を採用した.

4.2 ECoG-ALS データセット

ECoG-ALS は, ALS を患った被験者 1 名から計測された運動想起時の頭蓋内脳波データセットである. 本データセットに含まれるタスクは, 肘の伸展および屈曲, 手の伸展および掌

握の 4 種類の運動想起から構成される. 被験者は 8 日間にわたって, 4 種類の運動想起を実験日ごとに各 40 回行った. そして, 1 名の被験者から合計 1280 サンプルが計測された. 各サンプルにおけるサンプリング周波数は 1000Hz であり, 4 秒間の安静の後に, 2 秒間の運動想起が記録された. なお, 本研究では運動を想起する 2 秒前から 2 秒後に渡る 4 秒間を使用した. また, 頭蓋内脳波は左脳の中心溝近辺に配置された 94 個の硬膜下電極から計測された. そして, データセットの分割では, OpenBMI と同様の交差検証を用いた.

4.3 学習設定

OpenBMI [Lee 19] および ECoG-ALS を用いた評価の詳細について述べる. 本研究における頭皮脳波および頭蓋内脳波の前処理では, 計算量を削減するべく 250Hz へのダウンサンプリングのみを実施し, モデルへの入力とした. そして, 各データセットにおける訓練集合, 検証集合およびテスト集合はそれぞれモデルの訓練, ハイパーパラメータの探索およびモデル性能の評価に使用した.

5. 実験結果

5.1 定量的結果

表 1 に OpenBMI [Lee 19] および ECoG-ALS における提案手法およびベースライン手法の定量的結果を示す. 表中の値は k -fold 交差検証 ($k = 8$) における 8 試行の平均および標準偏差を表す. 評価指標として, 分類精度および Macro-F1 を用いた. また, 以下の観点からベースライン手法を選択した.

- (i) 脳波分類において, 良好な結果が得られているモデルとして EEGNet [Lawhern 18], EEG Conformer [Song 22], および長嶋らの手法 [長嶋 24] を用いた.
- (ii) 頭皮脳波および頭蓋内脳波が多変量時系列信号として計測されることを考慮し, 多変量時系列予測において良好な結果が報告されている Informer [Zhou 21], PatchTST [Nie 23], および iTransformer [Liu 24] を用いた.

表 1 より, 提案手法は OpenBMI における分類精度および Macro-F1 において, それぞれ 81.62% および 81.56% であり, 最良の性能であった. また, 各評価指標において次点のモデルと比較して, 提案手法はそれぞれ 3.72, および 3.78 ポイント上回った. さらに, 提案手法は ECoG-ALS においても各評価指標で最良の性能が得られ, 分類精度および Macro-F1 において, それぞれ 75.08% および 73.87% であった. そして, 各評価指標において次点のモデルと比較して, 各評価指標それぞれ 9.61 および 10.54 ポイント上回った. また, OpenBMI および ECoG-ALS の各評価指標において, Wilcoxon 符号付き順位検定を実施し, 提案手法がベースライン手法に対して統計的に有意な性能向上を示した ($p < 0.05$).

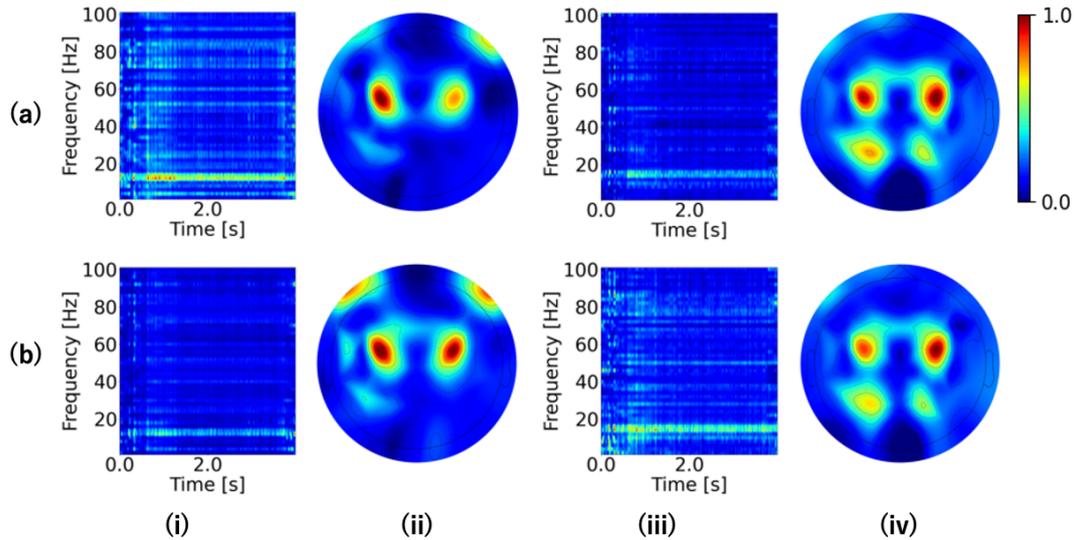


図 3: OpenBMI [Lee 19] における提案手法の視覚的説明。

5.2 視覚的説明

図 3 に OpenBMI [Lee 19] における提案手法の視覚的説明を示す。なお、視覚的説明の生成においては、Grad-CAM [Selvaraju 17] を拡張した手法を使用した。図において、(a) 行および (b) 行は、出力クラス（左手および右手の掌握）ごとの予測成功例に対する視覚的説明をそれぞれ示す。また、(i) 列および (iii) 列では、時間・周波数領域における視覚的説明を示し、(ii) 列および (iv) 列では空間領域における視覚的説明を国際 10-20 システムに則りトポグラフィックマップとして投影した。さらに、(i)(ii) 列および (iii)(iv) 列はそれぞれ被験者 02 および 11 に対する視覚的説明を示す。図 3(i) および図 3(iii) より、提案手法の説明では 10Hz 付近のミュール波が時間に依らず注目されていたことが分かる。ここで、10Hz 付近のミュール波は、頭皮脳波において運動想起との強い関連が報告されている周波数帯域である [Pfurtscheller 06]。また、図 3(ii) および図 3(iv) より、C3 および C4 電極の近傍が注目されていたことが分かる。ここで、C3 および C4 電極は運動皮質に位置する電極であり、左手および右手の運動制御との強い関連が報告されている [Pfurtscheller 06]。したがって、これらの結果から、提案手法が被験者に依らず、頭皮脳波において神経生理学的に重要な領域に注目していたことが示唆される。

6. おわりに

本研究では、頭皮脳波・頭蓋内脳波に基づく想起された行動の多クラス分類タスクを扱った。本研究の貢献は以下である。

- 深層状態空間モデルを拡張し、時間・空間・周波数領域の統合的な依存関係を捕捉する Frequency-SSM および Channel-SSM を導入した。
- 決定的および適応的に得られた周波数特徴を統合する Wavelet-Convolution を導入した。
- 運動想起時脳波を扱う 2 つのベンチマークにおいて、提案手法がベースライン手法を上回る結果を得た。
- 提案手法の視覚的説明より、頭皮脳波において神経生理学的に重要な特徴を捉えられたことが示唆された。

将来研究では、頭皮脳波および頭蓋内脳波におけるセッション間および被験者間のドメインシフトへの対処が挙げられる。
謝辞

本研究の一部は、JSPS 科研費 23K28168, JST ムーンショットの助成を受けて実施されたものである。

参考文献

- [Hramov 21] Hramov, A., et al.: Physical principles of brain-computer interfaces and their applications for rehabilitation, robotics and control of human brain states, *Physics Reports*, Vol. 918, pp. 1–133 (2021)
- [Lawhern 18] Lawhern, V., et al.: EEGNet: A Compact Convolutional Neural Network for EEG-based Brain-Computer Interfaces, *Journal of Neural Engineering*, Vol. 15, No. 5, p. 056013 (2018)
- [Lee 19] Lee, M.-H., Kwon, O.-Y., et al.: EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy, *GigaScience*, Vol. 8, No. 5, p. giz002 (2019)
- [Liu 24] Liu, Y., Hu, T., Zhang, H., Wu, H., Wang, S., Ma, L., and Long, M.: iTransformer: Inverted Transformers Are Effective for Time Series Forecasting, in *ICLR* (2024)
- [Nie 23] Nie, Y., Nguyen, N., Sinthong, P., and Kalagnanam, J.: A Time Series is Worth 64 Words: Long-term Forecasting with Transformers, in *ICLR* (2023)
- [Pfurtscheller 01] Pfurtscheller, G. and Neuper, C.: Motor imagery and direct brain-computer communication, *Proceedings of the IEEE*, Vol. 89, No. 7, pp. 1123–1134 (2001)
- [Pfurtscheller 06] Pfurtscheller, G., et al.: Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks, *NeuroImage*, Vol. 31, No. 1, pp. 153–159 (2006)
- [Selvaraju 17] Selvaraju, R. R., Cogswell, M., Das, A., et al.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, in *ICCV*, pp. 618–626 (2017)
- [Smith 23] Smith, J., Warrington, A., et al.: Simplified State Space Layers for Sequence Modeling, in *ICLR* (2023)
- [Song 22] Song, Y., Zheng, Q., Liu, B., and Gao, X.: EEG Conformer: Convolutional Transformer for EEG Decoding and Visualization, *TNSRE*, Vol. 31, pp. 710–719 (2022)
- [Zhou 21] Zhou, H., Zhang, S., Peng, J., et al.: Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting, in *AAAI*, Vol. 35, pp. 11106–11115 (2021)
- [長嶋 24] 長嶋 隼矢, 兼田 寛大, 飯田 紘, 田口 美紗 他: Retention 機構に基づく頭蓋内脳波の分類と BMI の構築, 第 38 回人工知能学会全国大会資料, pp. 4N1GS101–4N1GS101 (2024)