

TSUBAME 共同利用 令和7年度 産業利用 成果報告書

利用課題名 EEG を用いた大規模 Brain Computer Interface の開発
英文: Development of Brain Computer Interface from EEG

利用課題責任者 金井良太
First name Surname Ryota Kanai

所属 株式会社アラヤ
Affiliation Araya Inc.
URL www.araya.org

邦文抄録(300 字程度)

本研究は、耳周囲に装着する簡易な脳波計 (around-ear EEG) を用いたサイレントスピーチデコーディング (SSD) の実用性を検証した。24 名の健常者と 1 名のロックイン症候群患者から計 282 時間の EEG/EMG データを用い、大規模データや異種データの統合が精度向上に寄与することを示した。64 語分類で最大 47.3% (患者) の精度を達成した。結果は、軽量な装着型 EEG でも大規模学習により実用的な意思伝達手段となり得ることを示す。

英文抄録(100 words 程度)

This study investigates silent speech decoding (SSD) using a wearable around-ear EEG device. By collecting and integrating 282 hours of EEG/EMG data from 24 healthy participants and one individual with locked-in syndrome, the authors show that larger and more diverse datasets significantly improve decoding performance. The system achieved up to 47.3% for the patient in a 64-word task. These results demonstrate that lightweight, wearable EEG combined with large-scale training can enable practical, non-invasive communication systems.

Keywords: 5つ程度

背景と目的

ALS 等の神経変性症や、脳卒中後遺症により発声が困難な患者のコミュニケーションを補助する技術として、脳波 (EEG) や筋電 (EMG) を用いた発話解読のブレイン・コンピュータ・インターフェース (BCI) 技術が注目されている。従来の侵襲的手法 (ECoG, マイクロ電極) による発話解読 BCI は、高精度ながらも、外科的な手術を伴うため、感染症のリスクや心理的負担が大きい。そのため、非侵襲な代替手段として EEG/EMG による高精度な発話解読手法が求められる。

本プロジェクトでは、200 時間を超える大量の健常者 (言語によるコミュニケーションに問題がない方) の発話中の EEG/EMG データが、発話に困難を抱える方の EEG/EMG データから発話内容を解読する silent speech decoder の開発を行った。大量の発話中 EEG データを、対象となる患者の EEG データと混合して学習することにより、対象患者の silent speech の解読精度が上昇する結果が得られた。この結果は、対象患

者のデータ収集の負担を減らし、高精度なコミュニケーション補助を行う silent speech decoder システムの可能性を示唆する。

概要

1. EEG/EMG データ

学習に使用したデータは次の 3 種類である。一つ目は、4 人の健常者から g.Pangolin, g.SCARABEO, eego-sports で取得された、文章を自由なペースで有声発話中の 200 時間以上の EEG/EMG、text、audio のペアデータ。二つ目は、8 人の健常者と 1 人の imcompletely locked-in-syndrome (LIS) 患者から eegp-sports を用いて取得された、64 単語を決められたペースで一音ずつ有声・無声発話するタスクを実施時の、EEG/EMG、text、audio、発声タイミングのペアデータ (>16 時間)。最後に、装着が容易で患者への負担が少ない、around-ear EEG デバイスを用いて計測した、健常者 24 名、LIS 患者 1 名から

なる 64 単語を決められたペースで一音ずつ有声・無声発話するタスクを実施時の EEG, text, audio, 発声タイミングのペアデータ(>50 時間)。最後のデータセットでは、健常者からは、一部の被験者を除き各発話ごとに 10 runs ずつ、LIS 患者から 13 日に渡り計 136 runs を取得した。1 run はランダムに提示される 64 単語からなる。

2. 学習・評価手法

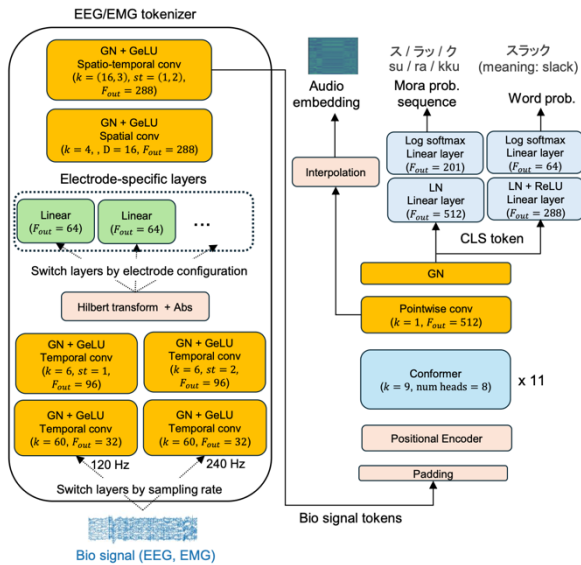


図 1 学習に使用したモデルのアーキテクチャ

評価時は一貫して、around-ear EEG で取得された、データセットから leave-out された run を test データとして学習後のモデルの評価に用いた。学習には、HTNet[1]と Conformer[2]ベースのモデルを用いた(図1)。様々な電極と発話タスクを処理可能なように、電極専用レイヤーの配置を行うとともに、EEG encoder に以下の3つの head を接続するように設計した。(1) text head : EEG latent を mora (日本語の音韻単位)の列を出力する,(2)word head: EEG latent の先頭に配置された [CLS token], から word を予測する, (3) audio head: EEG latent から Wav2Vec2.0[3]で計算された audio latent を予測する。この構成により、EEG デバイス、発話タスクについて異種混合のデータを有効活用しつつ、発話の種々の側面を学習することが可能になる。

結果および考察

1. データ量増加の silent speech decoding 精度への寄与

図2は、LIS 患者の test データにおける around-ear EEG から予想した単語分類精度を示す(語彙数 64)。左から順に(1)LIS 患者データから各 test run と同一日に取得された 8run を学習に用いた場合(2)LIS 患者データから test データを除く全ての run を学習に用いた場合、(3)cEEGrid-Extended: 健常者の around-ear EEG データを混合した場合、(4)Cross-Device: eego-sports のデータを追加した場合、(5)Cross-Task: 発話タスクとデバイスが異なるデータを追加で混合した場合の精度を示し、右進むにつれて、学習データの総量が大きくなる。

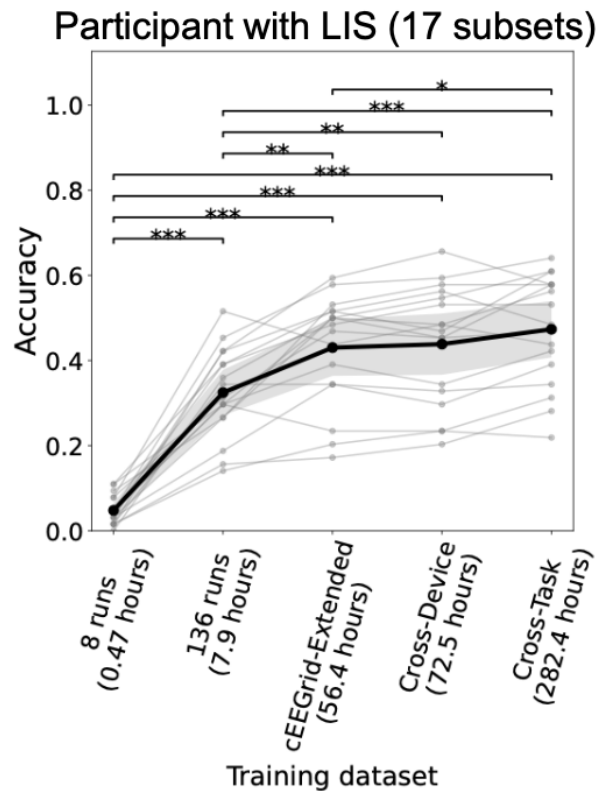


図 2 データ量増加と word classification accuracy

図 2 によると、Cross-Task 条件で、最も、LIS 患者における単語分類精度の平均が高く、cEEGrid-Extended 条件に対し、有意に精度の向上を示した(t-test, $\alpha=0.05$, Bonferroni 補正)。このことは、被験者間、デバイス間、さらにはタスク間の性能転移が可能であることを示し、ユーザーのキャリブレーションコストが

低い silent speech decoder の開発の可能性を示唆する。

まとめ、今後の課題

本研究では、多様な記録デバイス・タスクで大規模に収集した健常者の EEG/EMG データから、患者の EEG/EMG データへの転移学習が可能であることを示した。重要なことは、around-ear EEG デバイスという簡易で電極数も疎なデバイスに対しても、全脳を覆う、高密度なデバイスで収集されたデータが寄与したということであり、これは、研究室グレードではなく、実環境における silent speech decoder の可能性を提示する。

今後は、より多様で大規模な EEG データを収集し、EEG から単語ではなく文章レベルの解読を目指す。

[1] Peterson, S.M., Steine-Hanson, Z., Davis, N., Rao, R.P. and Brunton, B.W., 2021. Generalized neural decoders for transfer learning across participants and recording modalities. *Journal of Neural Engineering*, 18(2), p.026014.

[2] Gulati, A., Qin, J., Chiu, C.C., Parmar, N., Zhang, Y., Yu, J., Han, W., Wang, S., Zhang, Z., Wu, Y. and Pang, R., 2020. Conformer: Convolution-augmented transformer for speech recognition. *arXiv preprint arXiv:2005.08100*.

[3] Baeovski, A., Zhou, Y., Mohamed, A. and Auli, M., 2020. wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in neural information processing systems*, 33, pp.12449-12460.