

スケートボードトリック分類チャレンジ

ユーザー名 kimpar

データ①スコア 0.9101877 4位

データ②スコア 0.7037534 8位

要旨

脳波 (EEG) を用いた動作種別の推定モデルの開発と評価を行った。前処理のない EEG データでは、エポックの標準偏差の中央値を基準にした正規化を提案する。モデリングでは、1次元、2次元、3次元の畳み込みニューラルネットワーク (CNN) を用いて、時間的および空間的情報を考慮したモデルを構築してモデル間の比較を行った。

解析対象となる脳波成分の考察

解析対象となる脳波成分は、運動に先行して現れる運動関連脳電位と考えられる。

スケートボードのトリックは全身を使った一連の運動であるため、運動開始タイミングを一意に定義することはできないが、被験者の上方向加速度が増加し下肢の筋負荷が高まる、R 進入タイミングを運動開始とする。ボトムの滑走時間とテストデータの取得タイミングから、解析対象は運動開始前の 0.2 ～0.7 秒前となる。

対応する運動関連脳電位の成分は運動開始の約 1 ～2 秒前に出現する Bereitschafts potentials (BP) と運動開始の 0.4～0.5 秒前に急峻に陰性方向に増大する negative slope (NS') とが考えられる。利き手・非利き手、手か足かによって、同側か対側かが異なるが、BP,NS' とともに運動肢と活動する半球の関連が報告されている⁽¹⁾。

スケートボードトリックで主に活動する足はボードスタンスによって決まり、左右ターンで活動肢が変わるわけでは無いが、回転動作を行うための体幹の筋活動は左右で変わる。これらのことから、トリック判別においては、脳電位の陰性変動と両半球の活動差が主なポイントになると考えた。

データ処理の工夫点と考察

データ 1 の処理

セッションの連続データからトリガー後 0.5 秒間を切り出してエポックとする。連続データには低周

波の基線の揺動が見られたため、エポック分割後に直流成分を除去する。

エポックの特徴を捉えるために、セッション間やチャンネル間のスケールを揃える正規化が必要となる。同一セッションかつ同一チャンネルのデータであればスケールが同等であると考えられるため、セッションを通してのチャンネル標準偏差でスケールリングすることが望ましい。

一方、基線の揺動により、同一セッションであってもエポック間の値が大きく異なることがあるため、各チャンネルのエポック毎の標準偏差でスケールリングすることが適切な場合もある。

そこで、各チャンネルにおいて、同一セッション内のエポックの標準偏差の中央値 Std_{Median} を算出して、それぞれのエポックの標準偏差が Std_{Median} より大きい場合はそのエポックの標準偏差で、変動幅が小さくエポックの標準偏差が Std_{Median} より小さい場合は Std_{Median} で正規化する。これにより、同一セッションのエポックの大小関係を概ね維持しつつ、大きなドリフトがある場合でもスケールを揃えることができ、予測精度の向上が確認できた (図 1, 表 1)。

全被験者で無効チャンネルとして 0 埋めされたチャンネルが存在したが、これらはデータ形状を統一するために残した。

データ 2 の処理

データ 2 は低周波成分が除去されていたため、基

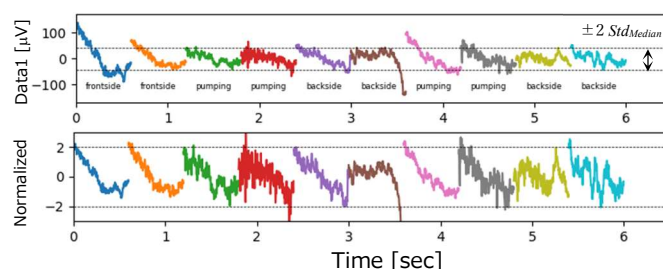


図 1 正規化前後の比較(Cz)

表 1 正規化の違いによる CV 精度

	Accuracy
エポック標準偏差で正規化	0.866
セッション標準偏差で正規化	0.869
提案手法	0.878

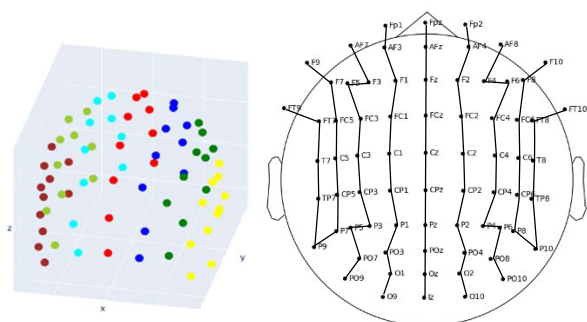


図2 2D-CNN 入力データのグループ分け
色分け・線で結んだチャンネルでグループ化した

線の揺動は見られなかった。そのため、直流成分の除去は行わず、標準化もセッション全体のチャンネル標準偏差で行った。

モデリングの工夫点と考察

72 チャンネルの時系列データを入力するにあたり、1次元畳み込みモデル、2次元畳み込みモデル、3次元畳み込みモデルの3つのモデルを検討した。

1次元畳み込みモデル (1D-CNN)

このモデルは、72 チャンネル×250 タイムステップのデータを時間方向に畳み込む。このモデルは、すべてのチャンネルを等しく見渡して時間的特徴をとらえることができる利点がある一方、電極配置の空間情報を考慮できない弱点がある。

そこで、正中線を除く左右のチャンネルの替わりに、左右チャンネルの差と和を入力する。これにより、左右半球活動の違いを強調し、モデルの判別能力を向上させた。

2次元畳み込みモデル (2D-CNN)

1D-CNNでは空間情報が利用できないため、左右チャンネルの差と和で疑似的に表現したが、時間と空間の両方を考慮するモデルとして、2D-CNNを検討した。

電極を3次元空間上の左右方向の位置によってグループ分けし、順に並べることで空間次元とする(図2)。グループ内にまとめられた前後方向の電極はチャンネル次元とする。

左右方向に7グループに分け、グループ内の電極

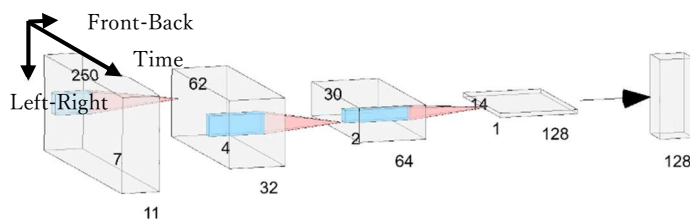


図3 2D-CNN モデル

数は 10 または 11 とした。電極数の少ないグループは 0 埋めした 1 チャンネルを追加してチャンネル数を合わせた。これを時間方向に並べて、 $11\text{ch} \times 7 \times 250$ のデータとすることで、左右方向軸と時間軸の 2 次元畳み込みを行った。(図 3)

3次元畳み込みモデル (3D-CNN)

測定点の EEG のパワースペクトルから脳波トポグラフィーを作成することができる。各時間ステップの脳波トポグラフィーを時間方向に並べることで動画データとして扱うことができ、3D-CNN による分類を行うことができる。

3D-CNNとして、PyTorchの事前学習済み動画分類モデルを利用した。これらの動画分類モデルはカラー3チャンネルの入力をとるため、第1層のレイヤ数を1チャンネルに変更し、 $1\text{ch} \times 32 \times 32$ の脳波トポグラフィーを入力できるようにした。使用した学習済みモデルはr3d_18, mc3_18及びr2plus1d_18である。学習済みモデルを初期重みとして、脳波トポグラフィーを使って転移学習する。

学習および推論

1D-CNN, 2D-CNN, 3D-CNN の 3 つのモデル, さらに LSTM を加えた 6 個のモデルで被験者毎に学習を行った.

損失関数としてクロスエントロピーロスを採用し、最適化手法には Adam オプティマイザーを用いた。3つのセッションに対して leave-one-out 交差検証を行い、各学習での最小 loss のエポックの重みを保存してテストデータの推論に使用した。

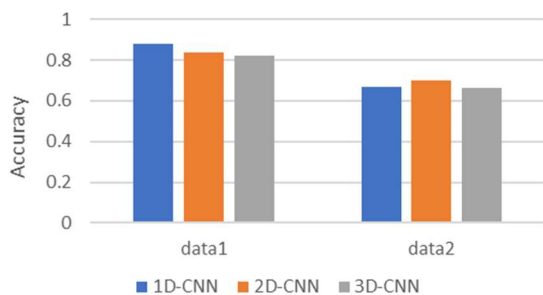


図4 CNNモデルの精度比較

学習された6モデルの推論結果を機械学習モデルでスタッキングした。2ndモデルの選択および学習はAutoMLであるPycaretを利用した。交差検証で得られた3つの2ndモデルの推論結果を多数決し、最終サブミットを作成した。多数決が同数であった場合は学習データ数の最も多かったポンピング運動を選択した。

考察

図4にそれぞれのデータでの各モデルの交差検証の精度を示す。データ1を使用した1D-CNNモデルが最も高い性能を示した。データ1にはBP,NS'の陰性変動が含まれることから、1D-CNNの時間的特徴に対する高い検出能力により、高精度となったと考えられる。2D-CNN,3D-CNNでは、畳み込みフィルタが時間軸と空間軸の両方に適用される。そのため、時間的特徴の検出に対するフォーカスが弱まり、重要な特徴が捉えにくくなった可能性がある。

データ2はバンドパスフィルタにより陰性変動は除去されている。時間的特徴が減少した結果、空間情報も考慮できる2D-CNNの性能が相対的に向上したと考えられる。3D-CNNは時空間を最も正確に捉えられる構造ではあるが、今回与えられたデータに対して複雑すぎるモデルであったため、精度が上がらなかったと考えられる。

生体信号データの取得における課題に対するAIによる解決策に関する考察

EEGは、脳の活動を示す多くのシグナルが含まれているが、測定ではさまざまなノイズが混入することが避けられない。主なノイズとして、電極の接

触不良や頭皮と電極間の抵抗変化、EMGの混入、環境からの電磁ノイズ、眼球運動に伴うアーチファクトや呼吸や心拍によるノイズが挙げられる。現状、EEGデータを分析するためにはバンドパスフィルタやICAなどでノイズを除去することが必要となる。

そこで、これらのノイズを自動的に正確に除去するためのアプローチとして、AIによる自動除去が考えられる。ノイズ除去した少数のEEGデータに対して、様々なノイズ元を模した人工的なノイズを加えた「ノイズ付き信号」を生成する。このノイズ付き信号を入力としてノイズ除去された元のEEGデータを復元するAIを学習する。これにはオートエンコーダーやVAEなどの深層学習モデルが利用できる。このようなモデルがあれば、人の手を介さないリアルタイムでのノイズ除去が実現でき、EEGデータの利用のハードルが下がるのではないかと考える。

社会課題への応用・展開に関する考察

EEGのBMIやBCIへの活用が期待されているが、広い分野での活用を実現していくために、EEGデータから得られる情報の分解能を上げていくことが求められる。そのためには、信号のもつ情報量を最大限生かして処理する必要がある。これまでノイズと一緒に捨てられていたデータからも情報も抽出しなければならない。

そうすると、人がデータを見て、信号かノイズかを判断することに限界が来ると考えられる。このような人では情報を抽出することが困難なデータこそ、AIによる自動的な特徴検出の仕組みが有効なのではないかと考える。今回のデータ1のように最低限の前処理でAIに入力し、大量のデータからAIで特徴を捉える手法が有望ではないかと考える。

参考文献

- (1) 西平賀昭. 運動と脳電位. 体力科学, 1995, 44.1: 173-179.