

スケートボーダー重心位置予測チャレンジ

ユーザー名 kimpar

スコア 0.6727162

順位 1 位

要旨

筋電図 (EMG) 信号を用いて重心速度を推定するモデルを構築した。EMG 信号の時間・周波数特性を抽出するために、連続ウェーブレット変換を用いてスカログラムを作成し、これを入力データとした。

LED 側と LED 逆側とで左右反転した筋負荷から速度を予測するため、左右チャンネルの振幅を相乗平均し、左右反転に頑健な特徴量を作成した。モデルには画像分類モデル ResNet18 利用し、これを回帰タスクに応用することで限られたデータセットでも高い精度を実現した。損失関数には評価関数と同じ 3 次元 RMSE を使い、評価指標を考慮したモデルの最適化を行った。さらに、推論結果の異常検知を行い、精度を向上させた。

データ処理の工夫点と考察

EMG 信号には振幅と周波数の時間変化が含まれているため、これらの情報を効率よく抽出するために連続ウェーブレット変換を施した。

連続ウェーブレット変換の周波数帯は、ノイズが少なくなる 10Hz から、EMG 信号の有効範囲と言われる 500Hz までを選択した。特に低周波成分に個々のデータの特徴が多く見られたため、周波数を対数スケールにすることで低周波成分を強調した。1,000 タイムステップのデータに対して 112 個の周波数で変換を行うことで、解像度 1000×112 のスカログラムが得られる。この画像を時間方向に縮小し、 112×112 の画像を生成した (図 1)。

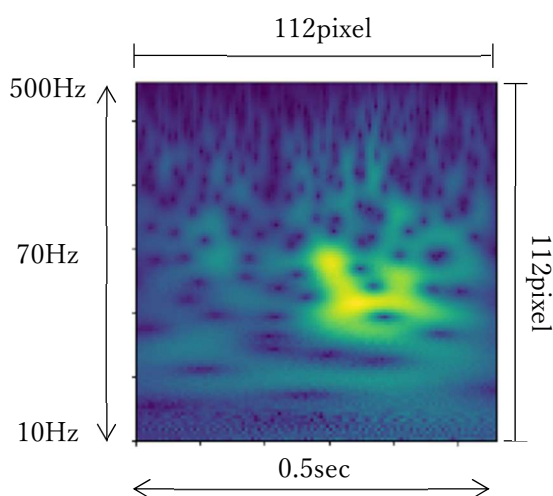


図 1 EMG 信号からのスカログラム

今回のコンペでは LED 側の学習で LED 逆側の重心速度を予測する必要がある。被験者が行う運動である左右ターンとポンピング運動のうち、左右ターンは LED 側・LED 逆側でボードスタンスが同一だが、ポンピング運動は LED 側と LED 逆側でボードスタンスが逆になり、体の動作が左右反対となる。このため学習時と予測時とでデータ分布が異なる可能性が高い。

そこで、左右の筋に対応する二つのチャンネルを平均して一つのチャンネルにすることで、左右反転の動作に対して同一の特徴量が得られるようにした。

EMG 信号の振幅は筋負荷におおよそ正比例すると言われている⁽¹⁾。しかし、その比例係数は電極と測定する筋との位置ずれにより、チャンネルによって異なることが予想される。

LED 側と LED 逆側とで EMG 信号と筋負荷の関係を以下のように仮定する。

$$\begin{aligned} \text{LED 側} \quad E_r &= k_r P_r, \quad E_l = k_l P_l \\ \text{LED 逆側} \quad E'_r &= k_r P'_r, \quad E'_l = k_l P'_l \end{aligned}$$

ここで、 E は EMG 信号の振幅、 k は電極の比例係数、 P は筋負荷、 r, l の添え字は体の左右を表す。

左右反転の動きで、 $P_r = P'_l, P_l = P'_r$ であったとすると、

$$E_r E_l = k_r P_r k_l P_l = k_r P'_l k_l P'_r = E'_r E'_l$$

となり、左右の EMG 信号の振幅の積をとれば左右反転しても等しい特徴量が得られる。これにより、LED 側と LED 逆側とを区別することなく、学習と予測を行うことが可能となった。

モデリングの工夫点と考察

モデルアーキテクチャには ResNet18 を用いた。ResNet は、深層学習における残差学習の手法を用いたモデルであり、画像分類タスクで高い性能を示している。分類タスクの場合は、モデルの出力層を分類クラス数と等しいノード数の全結合層とするが、これを X,Y,Z 方向速度の 30 タイムステップ分を表現する 3×30 ノードの層に付け替えた。

初期重みとして、PyTorch で公開されている学習済みモデルを利用し、スカログラム画像で転移学習

を行う。ResNet は $3\text{ch} \times 224 \times 224$ の画像に最適化されている。それを今回のデータ形状にあわせて、 $8\text{ch} \times 112 \times 112$ の画像で学習できるよう変更した。具体的には第 1 層の 2 次元畳み込みのチャンネル数を 3 から 8 へ、ストライドを 2 から 1 に変更することで、第 1 層の出力形状を元の ResNet と等しくなるようにし、それ以降の層の学習済み重みをそのまま利用できるようにした。第 1 層の 2 次元畳み込みのカーネルサイズを 3, 5, 7...13 に変化させて、6 種類のモデルを作成した。

学習および推論後の処理

予測モデルは被験者毎に作成した。対象被験者のデータを 3 分割して交差検証を行った。最初にリファレンスデータを合わせた 5 人分のデータで学習を行い、その後、対象被験者のデータのみを使って追加学習を行った。

データ拡張のため、時間方向に最大 7 ピクセル (1/32 秒相当) のランダムオフセットを加えた。

損失関数は今回のコンペの評価指標である 3 次元 RMSE を使用した。回帰タスクで多く利用される損失関数である MSE を使用すると、誤差の大きい試行に過度に影響されてしまう恐れがあるためである。最適化手法には Adam オプティマイザーを使用した。

交差検証による分割、複数のカーネルサイズ、複数のシードで被験者一人当たり 66 回の学習を行い、1 つのテストデータに対して 66 の推論結果を得た。

異常検知アルゴリズムである Local outlier factor を用いて学習データの速度分布とかけ離れた推論結果を排除し、残った結果を平均した。

最後に、LED 側と LED 逆側では、測定座標系から見た滑走方向が 180° 反転するため、X, Y 速度成分の符号を反転させた。

得られた知見

EMG の 1D 信号を処理するにあたり、1D-CNN, LSTM などを試したが、時間・周波数領域の特徴を画像に変換し、2D-CNN で処理する手法が最も高精度であった。短時間フーリエ変換でスペクトログラムを作成し、2D-CNN に入力することも可能だが、連続ウェーブレット変換を使用することで、時間・周波数領域共に高い解像度で特徴量を抽出できる。画

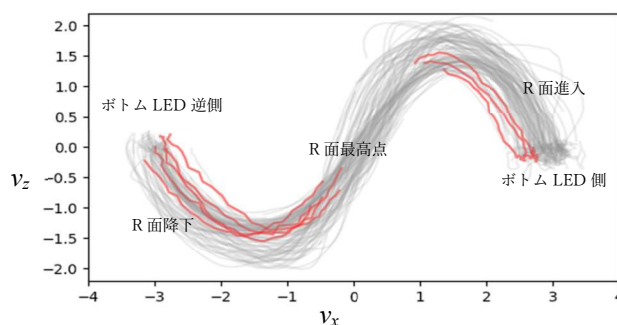


図2 複数回の推論によるばらつき

灰色は学習データ、赤はある 1 データの推論結果 (180° 反転) 学習時のシードの違いでばらつきが生じている。
 V_x の値からおおよそのランプの滑走位置が分かる。

像の解像度が上がったことで、既存の優れた画像分類モデルや学習済みモデルを利用できるようになり、限られたデータセットでも高い性能を発揮することができる。

学習データの X, Z 方向速度を重ねてプロットすると、Sin カーブ状の速度の分布が確認できる。テストデータで複数回の推論を行うと、推論結果に大きなばらつきが生じることがある。特に、R 面への進入と R 面から降り切った時とで、ばらつきが生じていた (図 2)。どちらも +Z 方向の加速度が発生するタイミングであり、下肢の筋負荷から Z 方向の速度を予測した結果、予測結果が二分してしまったと考えられる。

今回のコンペのような繰り返し動作の予測では異常検知アルゴリズムで異常な予測結果を取り除くことが可能だが、この手法は多種多様な動作を予測する場合には適用できない可能性が高い。

対象速度の分布範囲があらかじめ分かっている場合には、異常検知モデルによる後処理の代わりに、ニューラルネットワークの最終層に速度分布の相関関係や傾向をとらえた活性化関数を加えることが考えられる。今回の例であれば Z 速度を X 速度で多項式フィッティングするなどして、近似関数 $z = f(x)$ を定義する。モデル内部の変数を x, y, w にしたうえで、モデルの出力を $z = f(x) + \tanh(w)$ のような活性化関数に通す。このようにして、推論結果が目的の関数に沿うように制限することで、後処理が不要なエンドツーエンドの予測が実現できる可能性がある。

参考文献

- (1) 中村英夫. いまさら聞けない生体計測 1 (筋電). 生体医工学, 2019, 57.4-5: 149-157.