

〔経常研究B〕

深層学習による筋負荷解析の汎用化およびリアルタイム化

福田 純, 福井 航, 平田 一郎

1 目的

近年、ユーザーを第一に据えて製品の設計を行う「人間中心設計¹⁾」の考え方が広まっている。人間中心設計においては、ユーザーの使いやすさを定量的に評価する必要があるが、その評価指標の一つとして、筋負荷量が挙げられる。筋負荷量の算出の際には、筋電位を取得するが、この手法では製品が使用される現場環境での検証が難しく、人間中心設計の思想に適合しない。そこで、本研究では、カメラ動画のみ用いてデータを取得し、機械学習を用いて筋電位を予測する。これにより、任意環境で汎用的に使用可能かつリアルタイムな筋負荷の定量化が可能になり、人間中心設計を行う際に非常に有用なツールを開発することができる。

2 実験方法

2.1 システムの全体像

本研究では、動作動画から筋電位の予測を行う。このため、動作動画を学習データとし、動画の取得と同時に測定した筋電位を正解データとする。この際、動画のデータをそのまま学習データとして用いるのではなく、深層学習を利用した姿勢推定技術により、関節の座標（キーポイント）を取得する。これらキーポイントが記録された時系列データを学習データとする。システムの全体像としては図1のようになる。

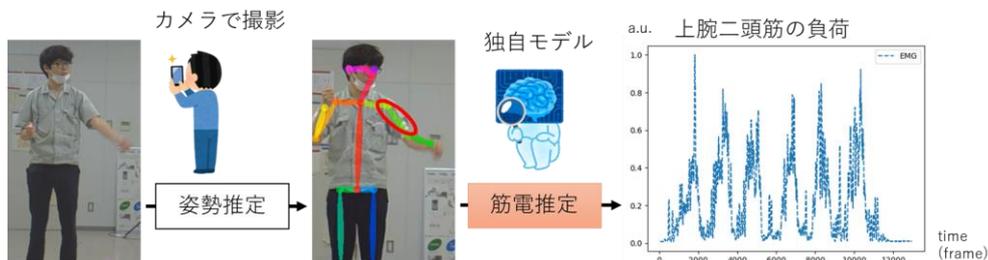


図1 システムの全体像

2.2 モデルおよび筋電位計の選定

姿勢推定を行うAIは多くの種類のもので存在するが、今回我々は、精度が高く、複数人物の検出にも対応可能な、OpenPose²⁾を選定した。筋電位の測定には、クレアクト社 biosignalsplux を使用した。サンプリングレートは1000Hzとした。筋電位計から得られる生データには、ノイズが多く含まれる上、正負の符号を持つ波形が測定される。今回我々が最も興味があるのは、波形の大きさであるから、符号は不要である。そこで、生データの絶対値を取る整流処理および、移動平均の算出（ウィンドウサイズ50）による、ノイズの除去を行った（ノイズ除去誤の信号サンプルは図1の通り）。

2.3 動作取得実験

動作を取得する実験においては、被験者は1名とし、20回の動作測定を行った。筋電位の測定部位はSN比が良好と見られる、上腕二頭筋とした。上腕に筋電計を装着したまま、RGBカメラにより動画を撮影し、後処理としてOpenPoseで姿勢推定を行った。なお、本研究においては、特定の動作の指定は行わず、自由に上腕を運動させた。これは、将来的な設計の現場に対応できるよう、汎用性を

高めるためである。動作測定の開始・終了について、動画の撮影と筋電位の測定とで同期させなければならないが、これは動作開始時および終了時にライトを点灯させることで実現した。また、動作動画のフレームレートと、筋電位の測定フレームレートは一致しない。そこで、スプライン補完によるアップサンプリングを OpenPose で解析した時系列データに施し、フレームレートを 1000Hz に統一した。これにより、学習データ（OpenPose による画像中の間接座標のデータ）と、筋電位の測定結果が、各フレームにおいて1対1で対応するようになった。

機械学習においては、シャローラーニングと呼ばれる伝統的な手法とディープラーニングと呼ばれる表現力の高い手法が存在する。今回我々は、以下のモデルを選定し、筋電位の予測を試みた。① Linear Regression（線形回帰） ② Random Forest ③ Gradient Boosting ④ Dense 層のみ ⑤ LSTM（長・短期記憶） ⑥ Transformer Encoder。1~3 はシャローラーニング・4~6 はディープラーニングに属する。1~3 のシャローラーニングにおいては、ハイパーパラメータ探索などは行わず、フレームワークで設定されたデフォルト値とした。4~6 のディープラーニングモデルの構造は図 2 の通りである。学習時のエポック数はおおよそ 20 程度とした。5・6 の時系列データを扱うモデルにおいては、入力を 10 フレームのシーケンスに分割し、次の 1 フレームを予測するようにデータの加工を行った。その他のモデルにおいては、時系列を考慮せず、フレームごとに予測を行う。

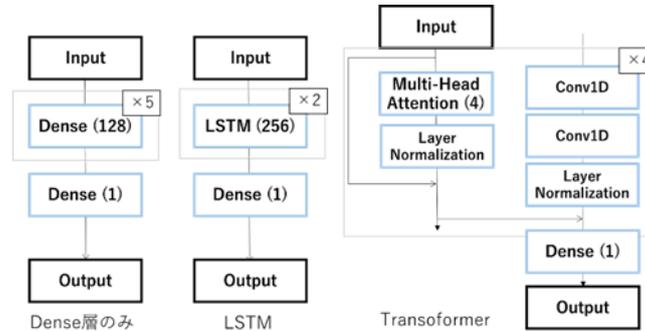


図 2 ディープラーニングモデルの構造

3 結果と考察

テストデータに対し、各モデルで予測を行った結果と、正解データ（実測の筋電位）を重ねたものを図 3 に示す。

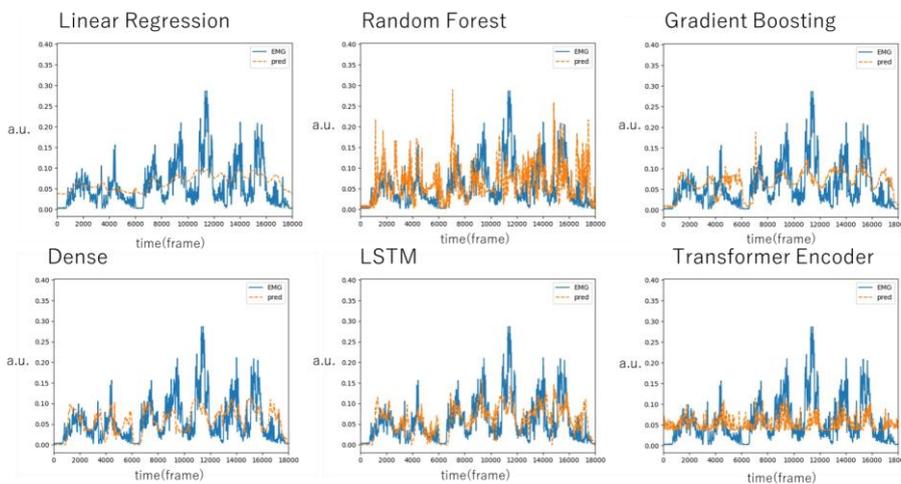


図 3 正解データ（青）とモデルごとの予測結果（橙）

図 3 に示されるように、定性的な評価（筋電位の時間的変化を追えているかどうか）においては時

系列上の前後関係を扱うことができる LSTM が最も高く、定量的な評価においても、MAE（平均絶対誤差）が 0.0344 と、誤差が最小となっている。

線形モデルにおいては筋電位の複雑なパターンが掴めておらず、モデルの表現力が足りていない。Random Forest においては、複雑なパターンは追えているものの、過学習しているような傾向が見られる。Gradient Boosting はシャローラーニングに分類されるモデル中では最も良い性能を示した。推論速度の要求によっては、選択肢の一つとなりうる。

Dense 層のみのモデルは高い表現力を見せているが、LSTM には及ばない。Transformer Encoder はそれらしい傾向を掴めてはいるが、余計なオフセットが出力されている。これは Transformer が全体的なパターンを捉えやすい一方で、個々のデータポイントの特性を見落としがちであるためと思われる。なお、どのモデルにおいても予測値が実測値よりも小さくなる傾向があるが、これは損失関数を調整することで、改善できる可能性がある。

4 結論

本研究では、任意環境下で測定が可能な、人間中心設計支援ツールの開発を目指し、姿勢推定 AI および独自モデルによる、筋電位のリアルタイム推定を試みた。複数のモデルを試行した結果、LSTM モデルが最も優れていることが分かり、その MAE は 0.0344 であった。筋電位の時間推移の大まかな傾向はつかめているものの、予測結果にはまだずれが大きい。特にピークを捉えきれない傾向が全てのモデルで見られるが、これは今後の工夫により改善される可能性がある。

参考文献

- 1) ISO 9241-210:2010
- 2) Shih-En Wei and Varun Ramakrishna and Takeo Kanade and Yaser Sheikh, CVPR2017 pp. 1302-1310 (2017),

(問合せ先 福田純)