

筋力トレーニング支援のための多種目トレーニング動画からの 複数部位の筋電位推定

藤本 彬仁[†] 牛尼 剛聡[‡]

[†]九州大学大学院芸術工学府 〒815-0032 福岡県福岡市南区塩原 4-9-1

[‡]九州大学大学院芸術工学研究院 〒815-0032 福岡県福岡市南区塩原 4-9-1

E-mail: [†]2ds21064n@s.kyushu-u.ac.jp, [‡]ushiana@design.kyushu-u.ac.jp

あらまし 本研究では、スマートフォン等に搭載されたカメラで撮影した筋力向上トレーニングの動画から、トレーニングの効果を正確に評価し、効果的なトレーニングを支援するシステムの開発を目的とする。トレーニングの効果を正確に評価する手法の一つとして、トレーニング中に筋電位を取得することが考えられる。しかし、筋電位センサの導入はコストが高く、また、ユーザがセンサを装着することを煩雑に感じることもある。そのため、一般のユーザが日常的に筋電位センサを装着してトレーニングすることは現実的でない。この問題を解決するために、我々は、トレーニング中の様子をスマートフォン等で撮影するだけで、運動効果を推定するシステムを開発中である。本論文では、複数種目のトレーニング動画から、それぞれの種目に対応する筋肉の筋電位を予測する手法を提案する。提案手法では、トレーニング中の動画に対して姿勢推定することで得た関節座標とその時の筋電位を教師データとして、深層学習を利用したトレーニング種目の検知、及び、筋電位予測モデルの学習を行う。そして、学習済みのモデルを利用して、トレーニング動画のみから運動効果を予測する。

キーワード トレーニング支援, 筋電位, 姿勢推定, 深層学習, 筋トレ

1. はじめに

近年、健康増進等を目的として、スポーツに関する注目が高まっている。図1はスポーツ庁のスポーツ実施状況に関する世論調査のグラフであり、スポーツを継続的に実施する人が増加していることがわかる。スポーツには様々な種類が存在するが、すべてのスポーツの基礎となる運動として、筋力トレーニング(筋トレ)がある。

筋トレを効果的に行うには、ジムに行き、トレーナーによる指導を受けることが重要である。しかし、ジムに行くことは金銭的なコストが高いことや、最近では新型コロナウイルス感染症に対する感染リスクのため、ジムに行くことを控えたいと考える利用者も多い。

そのような状況の中、ゲーム機やスマホ等のデジタルデバイスを利用して、自宅で一人で行う筋トレをサポートするシステムが注目されている。

例えば、NintendoSwitch上で動作するリングフィットアドベンチャー[2]では、バネの様な特殊なリング状のコントローラーなどを用いて身体の動作を検出し、その動作を基にしてゲームを遊ぶことで運動をサポートする。一方、スマートフォンで動作するトレーニング支援アプリケーションの一つであるNikeTrainingClub[3]では、ユーザは正しいトレーニングフォームを記録した動画を再生しながらトレーニングを行ったり、運動の記録を蓄積することが可能である。

これらのシステムは、正しい運動の手順やフォームを示し、利用者が行った運動に基づいて、予め決められた消費カロリーの推定などにより運動を評価する。しかし、運動の手順やフォームをユーザに提示したとしても、利用者のそれぞれ運動が適切に行われたかを評価をしているわけではないため、ユーザが効果的に運動をしたかどうかはわからない。

一方、トレーニング中の動画を深層学習によって自動的に解析してスクワットの回数を数えるアプリが開発されている[4]。しかし、回数を数えるだけでは、利用者が行った運動によってどの筋肉にどの程度の負荷

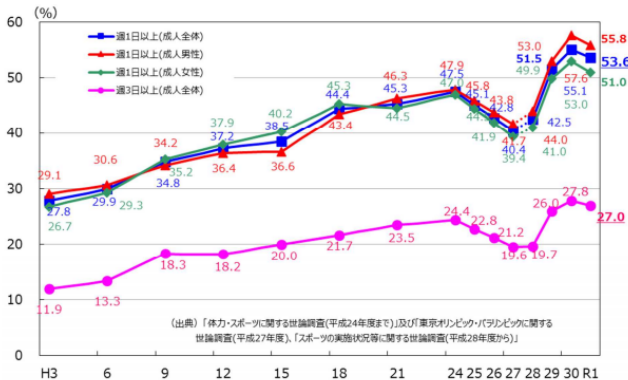


図1: スポーツの実施状況等に関する世論調査[1]

がかかっているかはわからないため、適切にどの程度運動が行われたかを評価することができない。

運動の効果を正確に評価するための一つの代表的な手法として、筋電位を計測することが挙げられる。文献[5]では、センサを利用してトレーニングをするユーザ（トレーニー）の筋電位などの生体情報を取得することで、適切な運動評価をフィードバックすることによって筋トレやリハビリをサポートするサポートシステムが提案されている。しかし、筋電位センサを利用するアプローチは、筋電位センサ等の高価な専用のデバイスを利用しなければならないため導入コストが高く、一般の人々が家庭で利用することは現実的でない。

上記の背景の下、我々は、スマートフォンに搭載されたカメラ等でトレーニング中の動画を撮影するだけで、個人での筋力向上トレーニングの効果を正確かつ容易に評価可能なトレーニング支援システムを開発すること目的として研究を行っている。本研究で開発するシステムの概要を図2に示す。

本システムでは、スマートフォン等のカメラでユーザのトレーニング中の様子を撮影し、撮影された動画をサーバに送り、サーバは送られた動画の解析と評価を行う。その評価をスマートフォン等に送り返し、ユーザは自分が行ったトレーニングに関する効果をフィードバックされることでサポートを受ける。

上記のシステムを実現するためには、サーバでどのようにしてトレーニング動画からそのトレーニングの効果を評価するかが重要である。本論文では、上記のシステムを実現するために必要な、トレーニング動画から、トレーニングの対象となる筋肉にかかる負荷を予測することにより、運動の効果を推定する手法を提案する。具体的には、トレーニング中に人体のいくつかの部位の筋電位を計測し、それを正解ラベルとして学習を行うことで、動画のみから運動の効果を予測する。

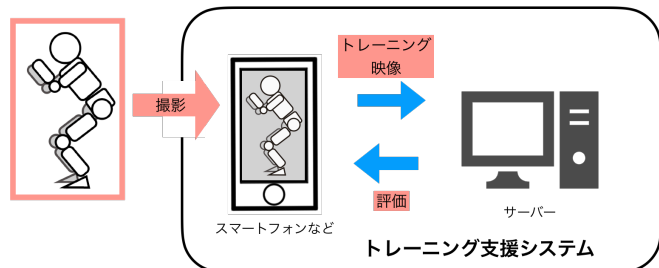


図2：提案システムの概要

2. 関連研究

本論文では、トレーニング中の動画から姿勢を抽出し、運動の効果を予測する手法を提案する。提案手法

では、効果的に学習を行うためにトレーニング動画そのものを直接学習させるのではなく、トレーニング動画からトレーニーの姿勢を推定し、姿勢データを学習させる。

動画からの人物の姿勢推定を目的として、これまでにいくつかの研究が行われている。OpenPose[6]やPoseNet[7]などの姿勢推定ライブラリでは、深層学習を利用して単一カメラ画像からの姿勢推定が可能である。また、複数カメラからより高精度に姿勢推定を行う手法も提案されている[8]。

こうした背景の下、近年、姿勢推定によって得られた姿勢情報を利用したシステムが提案されている。PoseNet を利用したスポーツ反復練習のためのスマートミラーシステム[9]や、接触型のモーションキャプチャシステム OptiTrack を用いた静止状態からの動作の予備動作に関する研究[10]や、非接触型のモーションキャプチャシステムである Kinect を用いた筋力トレーニング支援システムの研究[11]も行われている。しかし、これらの研究では、カメラ画像のみから筋肉の部位ごとの運動効果を推定しトレーニングを支援することは行っていない。

一方、筋電位センサを利用したトレーニング支援システムに関する研究も行われている。高久ら[12]は、VR ゴーグルにリアルタイムで筋電位をフィードバックすることによりサポートを行うシステムを提案している。大塚ら[13]は6週間のウェイトトレーニング中の筋電位を取得し、その筋電位から個人毎に対応した、適切な重量、及び筋力増加に伴って重量を増加させるタイミングの判断が行えるかの研究も行われており、筋電位は筋力トレーニングの評価に適していると考えられる。しかし、これらの研究では、トレーニーの支援のために筋電位センサを利用する必要があるが、本研究ではトレーニーは筋電位センサを用いずにカメラだけでトレーニングを支援するシステムを開発する。

馬場ら[14]は、ユーザーから取得した姿勢情報を基にしたCGアバターに、筋電位センサーから得られた筋電位情報に基づいた計算した力みの大きさを重畳し可視化することで、サポートを行うシステムを提案している。これは、トレーニング中の筋電位を姿勢情報に基づいて対応する筋肉に対して筋電位の可視化を行うことでサポートを行う。このサポートシステムは、本研究のアプローチと類似しているが、本研究では、トレーニングを行うユーザはセンサを付けずにサポートを受けることを目指す。

我々のこれまでの研究[15]では、単一のトレーニング種目かつ単一部位の筋電位予測について述べており、本研究では複数トレーニングかつ複数部位での筋電位予測を目指す。

3. 動画からの運動効果の推定

3.1. 手法の概要

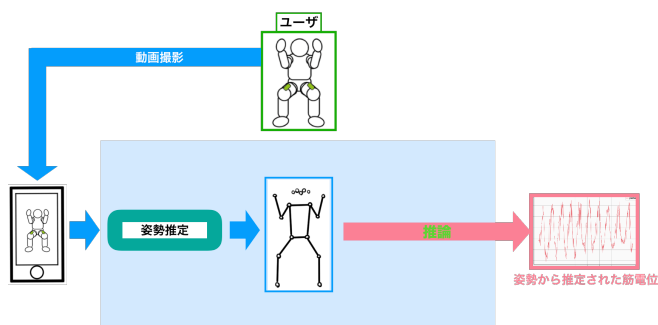


図 3：提案手法の全体像

本論文では、トレーニングを撮影した動画から運動効果を推定する手法を提案する。提案手法では、運動効果の推定のためにトレーニング中の筋電位に注目する。筋電位とは、筋肉にどの程度の力が入っているかを示す指標であり、筋電位を測定することで、筋肉に対する負荷を計測可能である。そのため、本研究で目的とする「適切な運動効果」を推定するために利用可能であると考えられる。

提案手法の具体的な流れとしては、トレーニング中の動画データを学習データ、トレーニング中の筋電位データを正解ラベルとして取得しておき、それらのデータをデータセットとして学習を行うことで、動画のみから筋電位を予測する深層学習モデルを作成する。

トレーニング動画からトレーニング中の任意の時刻における筋電位を予測するために、本研究では、トレーニング動画を直接入力とするのではなく、撮影された人物に対して姿勢推定を行い、そこで得られた姿勢情報を入力として機械学習を用いて筋電位を予測する。予測した筋電位と、実測の生体情報との誤差が小さければ、生体センサを使用せずに、生体情報に基づいて筋トレの効果を評価する。

動画から推定された骨格情報を入力とし、入力されたデータから予測された仮定の筋電位情報を推定するため、提案システム使用者は、撮影された筋トレ動画のみから、その筋トレに対応した予測した筋電位情報に基づいたフィードバックを受けることにより効果的なトレーニングを実現することを目指す。

3.2. 筋電位情報の取得

訓練用のデータを取得するために、運動中の筋電位筋電位センサを利用して取得する。本研究では、筋電位センサとして bitalino 社の Bitalino Muscle BIT[16]を用いる。図 4 に Muscle BIT の外見を示し、図 5 に腕の

筋肉の筋電位を測定している様子を示す。

Muscle BIT では、筋電位のサンプリングレートを指定できる。本研究では筋電位を測定する際のサンプリングレートは 1000[Hz]とする。図 6 に Muscle BIT を用いて取得した生データ（電圧に変換する前のデータ）を示す。Muscle BIT では測定値は 0 から 1023 の整数値として返される。時刻 t における測定値 x_t を筋電位 v_t に変換するには以下の式を利用する。

$$v_t = \alpha |x_t - EMG_ave|$$

ここで、 α は変換係数であり、 EMG_ave は筋電位の平均値である。この式を利用して、筋電位に変換したデータの例を図 7 に示す。

測定された筋電位データにはノイズが含まれているため、移動平均を利用してデータの平滑化を行う。本研究では窓幅を 50 とした移動平均を取ることで平滑化を行う。図 7 のデータに対して、平滑化を行った例を図 8 に示す。

それぞれの動画のフレームにおける筋電位を取得するために、筋電位データ系列をフレーム数で分割して 1 フレームあたりのデータ数を求め、1 フレームに対応付けることができるデータの平均値をフレームに対する筋電位として扱う。図 9 は、図 8 のデータそれぞれのフレームに対応付けたグラフを表す。

動画撮影開始時と筋電位センサでデータは、開始時刻を合わせることで二つのデータの同期させる。



図 4：Bitalino Muscle BIT

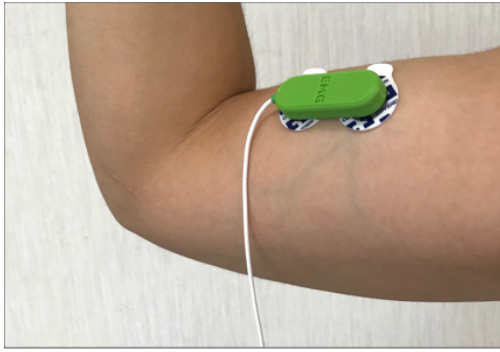


図 5 : Bitalino Muscle BIT の装着例

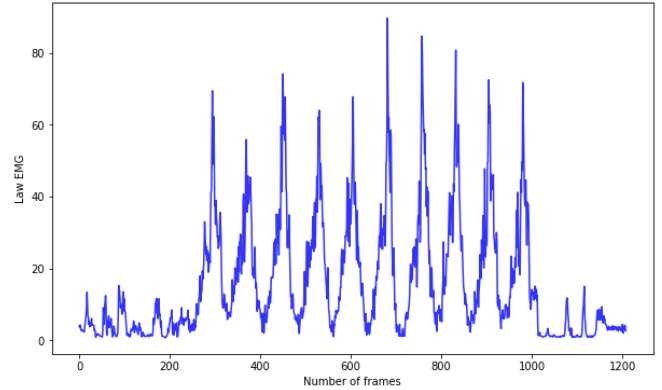


図 9 : 筋電位をフレーム数で割ったグラフ

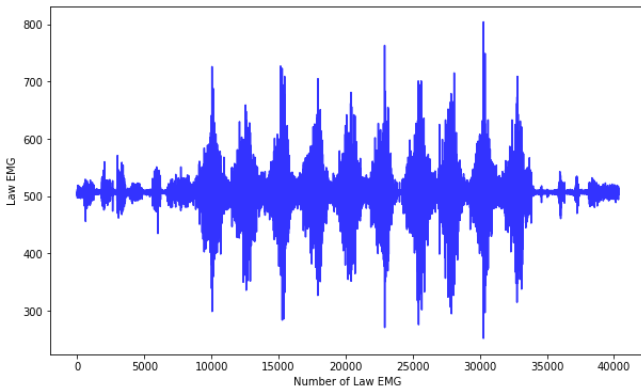


図 6 : 実測筋電位の例

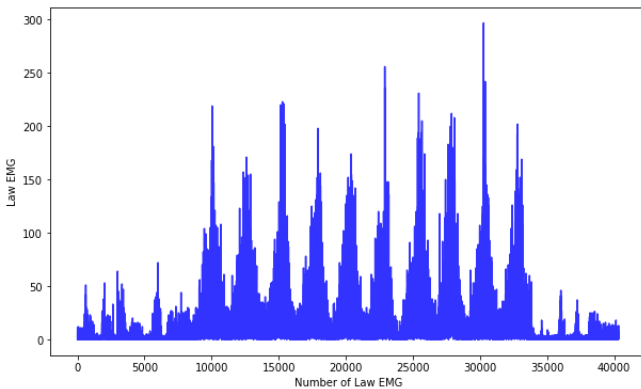


図 7 : 筋電位を反転処理したグラフ

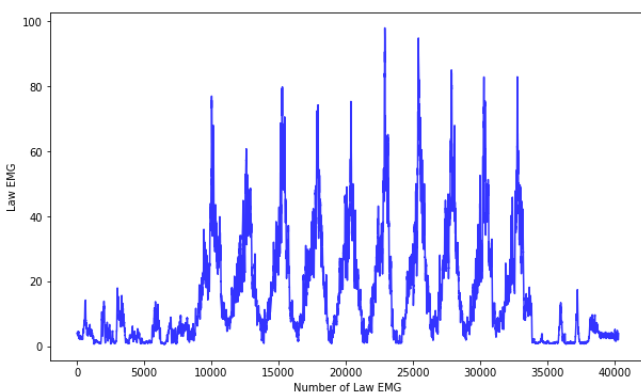


図 8 : 筋電位の移動平均を取ったグラフ

3.3. 動画からの骨格推定

ユーザのトレーニング動画から深層学習で筋電位を推定する際には、トレーニング動画からユーザの運動の状態を認識する必要がある。一般的に、動画像には様々な対象が撮影されたり、撮影する画角や大きさなどが変化する。そのため、動画像をそのまま入力として機械学習をしたとしても高精度に予測可能な汎用的なモデルを得ることは難しいと考えられる。そこで、本研究では、動画像に対して一般的な姿勢推定アルゴリズムを利用して姿勢を検出し、姿勢情報を入力として筋電位を推定するアプローチを採用する。

姿勢情報を推定するための方法として、10年ほど前までは Kinect などの深度センサを用いた手法[17]やマーカを利用して推定する方法が一般的であった。しかし、近年、深層学習を利用してマーカーレスで動画のみから姿勢情報を推定する研究が活発に行われている。

マーカーレスで動画から姿勢検出を行う方法としては、人の画像に関節座標をアノテーションしている MS COCO Keypoint Detection Task dataset などのデータセットを畳み込みニューラルネットワークを利用したモデルによって学習を行い、関節に対するヒートマップを推定することで画像から関節座標の推定を行う手法が代表的である。

動画からの姿勢推定を行う代表的なライブラリとして OpenPose[6]と PoseNet[7]がある。これらのライブラリは、両者とも高い精度で検出を行えるが、OpenPoseの方が姿勢推定処理に必要な時間的なコストが大きいため、リアルタイム性を必要とする本研究では、処理コストが小さい姿勢推定ライブラリである PoseNet を用いる。

PoseNet では、画像から手や足などのキーポイントの 17 関節の予測確率とポーズ全体としての予測確率を検出する。PoseNet で検出可能な関節としては、鼻、目、耳、肩、肘、手首、臀部、膝、足首であり、PoseNet から出力される ID と対応する各部位をまとめたものを表 1 に示す。PoseNet は、一般的な PC でもリアルタ

イムで推定可能であり，本研究で開発するサポートアプリに適していると考えられる．PoseNet を利用して姿勢検出を行った例を図 11 に示す．



図 11：骨格推定前と後の人物画像

表 1：キーポイント 17 点の内訳

ID	部位
0	鼻
1	左目
2	右目
3	左耳
4	右耳
5	左肩
6	右肩
7	左肘
8	右肘
9	左手首
10	右手首
11	左臀部
12	右臀部
13	左膝
14	右膝
15	左足首
16	右足首

4. 複数トレーニングに対する筋電位予測

我々はこれまでに，単一のトレーニング種目の単一部位の筋電位に関する予測を行う手法を開発した．本論文では複数のデータに対しての応用を目的とする．

4.1. 予備実験

被験者は 1 名とし，トレーニング種目はスクワットとバーベルプッシュアップ，筋電位取得部位は両腿と両肩とした．

データ数は，各種目ごとに 10 回*10 セット取得した．

4.2. トレーニング種目の分類

学習を行う際に，どのトレーニングを行っているかの分類を行い，その分類結果を学習データセットに加えることで，精度の向上を目指した．トレーニングの分類にはトレーニング動画を姿勢推定することで得られたデータの 0~100 フレーム分を入力とし，CNN を用いて分類を行った．分類モデルは，2 次元畳み込み層とプーリング層を 4 ループ、それに加えて、平坦化層、全結合層を連結させることで構成した．学習結果のグラフを図 12 に示す．このグラフより，分類は高精度に実現できたことがわかる．

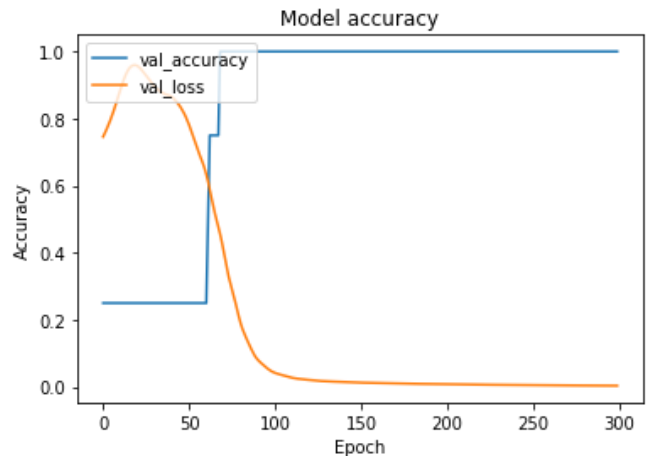


図 12：種目分類の学習結果

4.3. 予備実験結果

前述の分類ラベルを学習データに加えて，LSTM で回帰学習を行った際の学習結果を図 13 に示した．

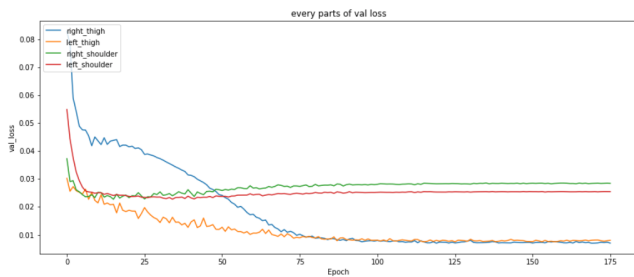


図 13：複数種目及び複数部位の回帰の学習結果

この学習したモデルを用いて学習データを予測した結果を図 14~17 に示す。色つきの線が実測筋電位であり、黒線が予測筋電位を示す。

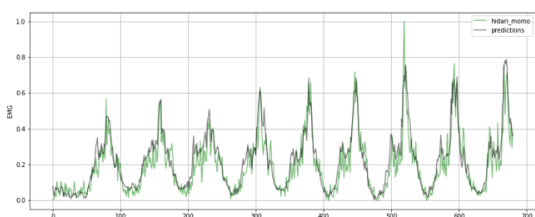


図 14：スクワットにおける太腿の筋電位予測結果

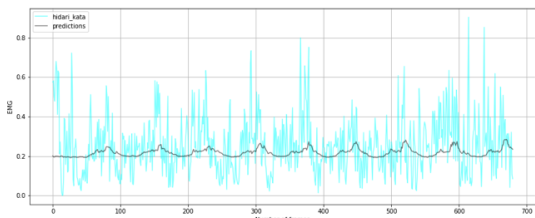


図 15：スクワットにおける肩の筋電位予測結果

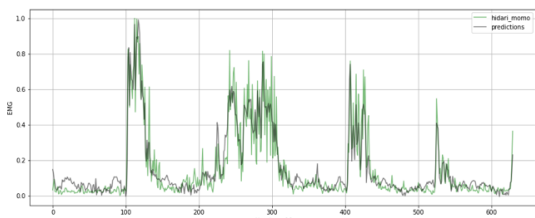


図 16：バーベルプッシュアップにおける太腿の筋電位予測結果

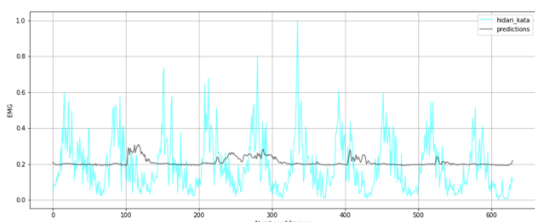


図 17：バーベルプッシュアップにおける肩の筋電位予測結果

次に目的変数を両肩の筋電位のみにした場合の予測筋電位を示す。

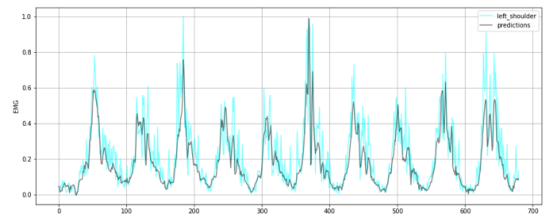


図 18：肩の筋電位のみを目的変数とした場合のバーベルプッシュアップ際の肩の筋電位予測結果

4.4. 考察

予備実験結果より、太腿の筋電位は正確に予測できているが、肩の筋電位が正確に予測出来ていない。

これは、肩の筋電位のピークが発生する時の骨格の動きが太腿の時の動作に比べて小さい事が原因と考えられる。

肩の筋電位が発生するために必要な動作は腕の動きのみだが、太腿の筋電位が発生するためには全身が沈み込むことによる全ての部位の骨格の動きが必要となる。そのため、太腿の方が全部の骨格データと相関するため学習しやすく、反対に肩の方は腕まわりの骨格データ以外はノイズとなるため相関が低く、これらの相関が異なるトレーニングデータを同時に学習すると、相関の高い方に学習が寄ってしまったと考えられる。

また、図 18 に示すように、肩の筋電位データのみを目的変数として学習を行い、筋電位予測を行ったところ正確に予測出来たため、学習の方法を再考する事が必要と考えられる。）

5. トレーニング支援アプリ

上記で述べた学習済みモデルを利用して、トレーニング支援を行う Web アプリケーションを開発する。

本システムでは、まず、ユーザーのトレーニング中の映像をブラウザ経由で取得し、その映像を PoseNet に入力することで解析を行う。次に、tensorflow.js 形式に変換した学習済みモデルに PoseNet の解析結果を入力することで、予測される筋肉の筋電位を可視化を行う。

図 18 は開発中のアプリの実行画面の例であり、以下にその詳細について述べる。

本アプリでは画面全体に、ブラウザから取得した映像が流れ、その映像に PoseNet で解析された関節と骨格が重ね合わせて表示される。左上の“映像の読み込み”部分では、ブラウザから取得している映像の状態を示しており、ロード中、ロード完了、エラーの 3 種

類のラベルが映像の状態に応じて切り替わる。

「トレーニング種目」部分では、取得した映像を基に、トレーニング分類モデルで予測したトレーニング種類を示しており、スクワット、バーベルプッシュアップ、データ収集中の3種類のラベルが、予測結果に応じて切り替わる。このトレーニング種目によって、どの種類の筋電位予測モデルを利用するかを決定する。

また、PoseNetによって解析されたポーズ全体の予測確率が0.7未満の場合、全身がカメラに適切に写っていない可能性が高いため、画面中央に警告文を表示する。

ポーズ全体の予測確率が0.7以上の場合、対象トレーニングに対応した筋電位予測モデルに、PoseNetの解析結果を入力することで予測筋電位を取得する。この予測筋電位を、種目に対応する筋肉が存在する骨格の大きさと色に対応させることで、ユーザーへの可視化を行う。

トレーニングを終了する際は、右下の終了ボタンを押すことで、トレーニングが終了し、最終的にどの程度筋肉が使用されたかをユーザーに提示する。

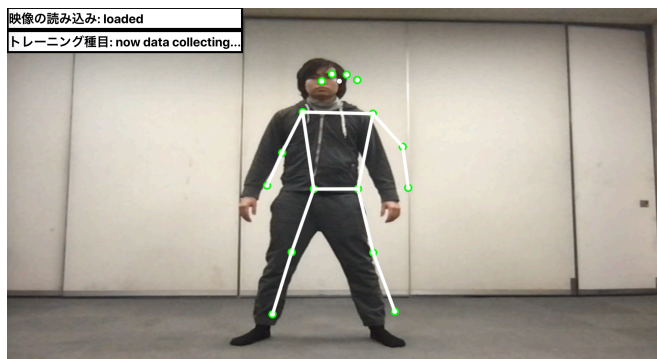


図 18 : 開発中のアプリの画面例

本アプリでは、まず最初に、トレーニングの分類を行い、その分類結果に応じて、単一種目のトレーニングに対応した筋電位予測モデルを切り替えて使用することで、複数の筋力トレーニングに対応する。

しかし、実際に使用してみると、リアルタイムでPoseNet、及び2つの学習済みモデルを使用するため、処理コストが大きくなった。そのため、映像を正確にい取得できず、適切な予測が行われなかった。

また、カメラに全身が映る程度の距離でサポートを受けると、画面が遠すぎて見辛く、サポートが受けづらいなどの問題があった。

5. まとめ

本論文では、カメラで撮影された筋トレ動画をPoseNetによって姿勢推定し、その推定結果とその筋トレに対応する筋肉の筋電位をデータセットとして、

深層学習を行うことで、筋トレ動画のみから筋電位を予測するモデルの作成を行い、予備実験を行った。

予備実験の結果より、姿勢と筋電位との相関の高い部位の筋電位予測は正確に行えたが、相関の低いトレーニングデータに関しては正確に予測を行うことが出来なかったため、transformerなどのLSTMとは別の深層学習方法を検証する必要があると考えられる。

また、サポートアプリの開発を行い、Webアプリケーション上で学習済みモデルが動作することを確認したが、計算量の多さによる処理落ち問題や、サポート方法の問題などを解決する必要がある。

謝辞

本研究はJSPS科研費19H04219の助成を受けたものです。

参考文献

- [1] スポーツ庁, “スポーツの実施状況等に関する世論調査”, 令和元年度「スポーツの実施状況等に関する世論調査」について, 2021年1月閲覧, <https://www.nintendo.co.jp/ring/>
- [2] Nintendo, “リングフィットアドベンチャー | Nintendo Switch |任天堂”, リングフィットアドベンチャー公式サイト, 2020年12月閲覧, <https://www.nintendo.co.jp/ring/>
- [3] Nike, “Nike Training Club”, Nike Training Club公式サイト, 2020年12月閲覧, <https://www.nike.com/jp/ntc-app>
- [4] NEXT-SYSTEM, “AI筋トレカウンター 家トレ”, 2020年12月閲覧, <https://www.next-system.com/ietore>
- [5] 株式会社エヌ・シー・ピー, “Measee ~筋(近)未来サポートシステム メーシー~”, 2020年12月閲覧, <https://www.ncp.co.jp/measee/>
- [6] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, Yaser Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2019
- [7] Dan Oved and Irene Alvarado, “Real-time Human Pose Estimation in the Browser with TensorFlow.js” (online), <https://medium.com/tensorflow/real-time-human-pose-estimation-in-the-browser-with-tensorflow-js-7dd0bc881cd5>, 2018
- [8] Takuya Ohashi and Yosuke Ikegami and Kazuki Yamamoto and Wataru Takano and Yoshihiko Nakamura, “Video Motion Capture from the Part Confidence Maps of Multi-Camera Images by Spatiotemporal Filtering Using the Human Skeletal Model”, IROS 2018
- [9] 新野大輔, 井尻敬, “姿勢検出法を活用したスポーツ反復練習のためのスマートミラーシステム”, エンタテインメントコンピューティングシンポジウム2019論文集 (p.138-144), 2019
- [10] 許超舜, 藤原正浩, 牧野泰才, 篠田裕之, “静止状態からの動作における予測可能性の検討及び予備動作の推定”, 第25回バーチャルリアリティ学会大会, 2020

- [11] 高久大輔, 中島克人, ” Kinect を用いた筋力トレーニング支援システム”, 情報処理学会第 77 回全国大会講演論文集 (p.437-438), 2015
- [12] 大塚勇人, 西垣一馬, 酒井元気, 岩井将行, ” MR と筋電位センサを組み合わせた筋力トレーニング効果のリアルタイムセルフモニタリングシステム”, インタラクシオン 2019 論文集 (p.798-802), 2019
- [13] 馬場裕子, 木竜徹, 山崎健, ” 表面筋電図を用いた短期筋力トレーニング時における個人対応プログラムに向けた運動機能評価”, 生体医工学 45 巻 1 号 (p.48-54), 2007
- [14] 浦脇浩二, 増田泰, 眞鍋佳嗣, 千原國宏, ” スポーツフォーム練習支援のための生体情報可視化”, 情報科学技術フォーラム一般講演論文集 (p.471-472), 2004
- [15] 藤本彬仁, 牛尼剛聡, トレーニング支援のための動画からの運動効果の推定, 第 13 回データ工学と情報マネジメントに関するフォーラム, 2021
- [16] creact, “BITalino MuscleBIT Bundle Data Sheet”, 2021 年 1 月閲覧, <https://www.creact.co.jp/wp-content/uploads/2017/07/BITalino-MuscleBIT-Bundle-Datasheet.pdf>
- [17] Microsoft, “Azure Kinect DK”, 2020 年 12 月閲覧, <https://azure.microsoft.com/ja-jp/services/kinect-dk/>
- [18] 内田祐介, 山下隆義, ” [サーベイ論文] 畳み込みニューラルネットワークの研究動向”, 電子情報通信学会技術研究報告 信学技報 (vol. 117, no. 362, p. 25-38), 2017
- [19] Aurélien Géron, ” scikit-learn と TensorFlow による実践機械学習”, p.381, p.403~405, 長尾高弘訳, 株式会社オーム社, 2018