

トレーニング支援のための動画からの運動効果の推定

藤本 彬仁[†] 牛尼 剛聡[‡]

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

[‡]九州大学芸術工学部 〒815-0032 福岡県福岡市南区塩原 4-9-1

E-mail: [†] 1ds17444y@s.kyushu-u.ac.jp, [‡] ushiama@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、トレーニング動画から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右肩上がりであることがわかる。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]というゲームでは、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。

しかし、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に、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。

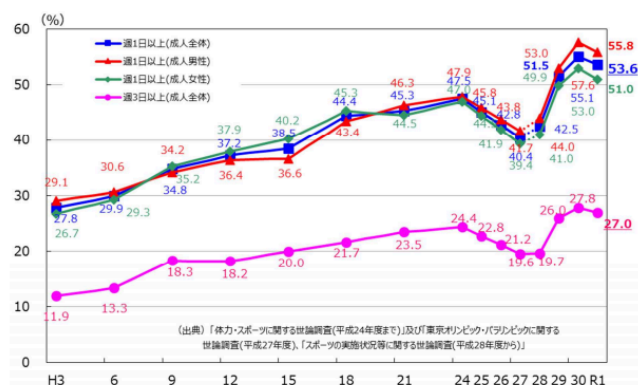


図1: スポーツの実施状況等に関する世論調査[1]

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ー（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ーのトレーニング中の様子を動画として撮影し、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、ユーザー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。

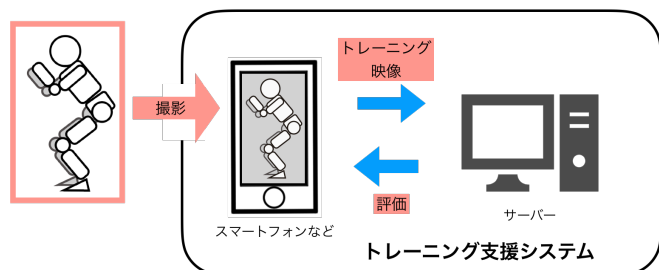


図2：提案システムの概要

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できていない。

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ーはセンサを付けずにサポートを受けることが可能である。

ドップラーセンサを利用して非接触で心電を深層学習により予測する手法[15]も研究されているが、本研究ではトレーニング動画のみから筋電位の推定を行うことを目標とする。

3. 動画からの運動効果の推定

3.1. 手法の概要

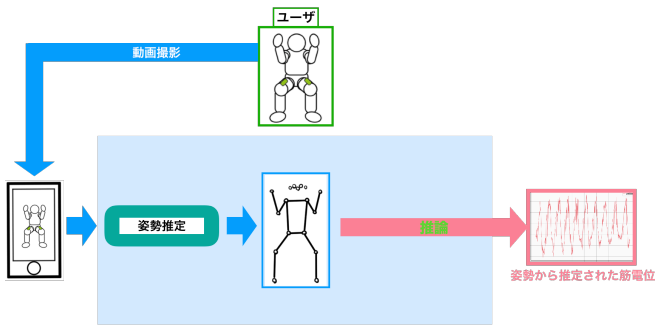


図 3：提案手法の全体像

本論文では、トレーニング中の動画のみから適切な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を利用する。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それぞれ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、トレーニングする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深層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の評価できる。

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できるため、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

学習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。

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

ここで、 α は変換係数であり、 ave は平均値である。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より平滑化を行う。図 7 のデータに対して、平滑化を行った例を図 8 に示す。

次、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

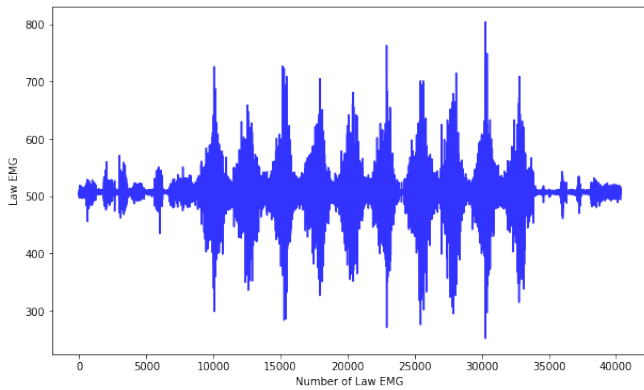


図 6：実測筋電位の例

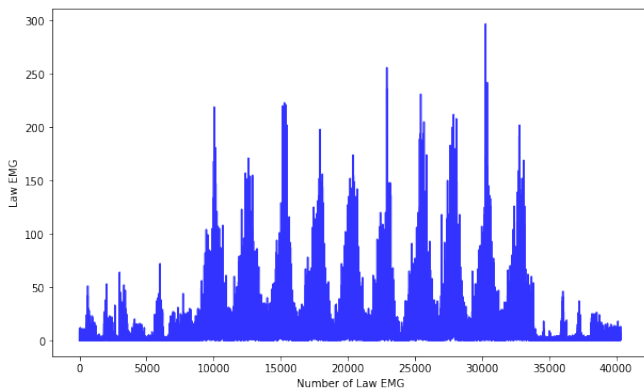


図 7：筋電位を反転処理したグラフ

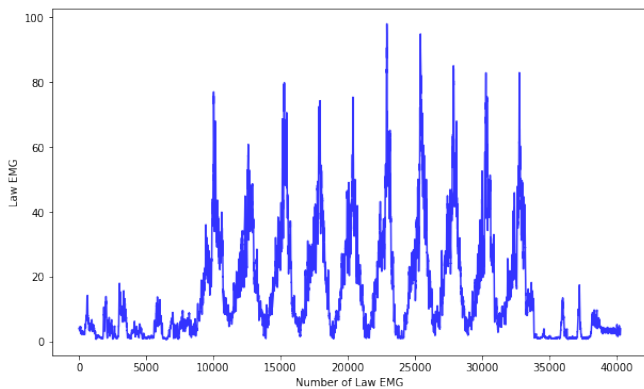


図 8：筋電位の移動平均を取ったグラフ

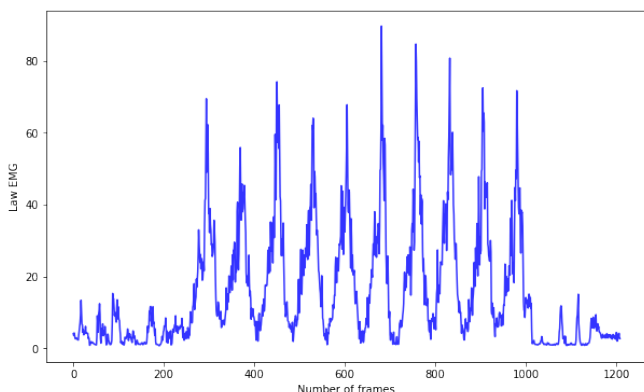


図 9：筋電位をフレーム数で割ったグラフ

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 (図 10) [18] 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。



図 10：MS COCO Keypoint Detection Task dataset の例

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時間を必要としたため、リアルタイム性を必要とする本研究では、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

3.4. 深層学習モデル

本研究では、トレーニング動画から PoseNet を利用して抽出された姿勢情報から機械学習を利用して筋電位を予測する。機械学習のモデルとしては様々な手法が提案されているが、本研究では近年様々な領域で利用され、成功的な成果を収めている深層学習を利用する。

深層学習とは、多層ニューラルネットワークを利用した機械学習であり、対象とするデータの特徴に応じていくつかの手法が提案されている。本論文では以下に示す 3 種類のモデルを提案し、評価実験により、それらの性能を比較する。

1. 全結合層モデル
2. CNN モデル
3. LSTM モデル

これらのモデルの概念図を図 12 に示す。

1 つ目のモデルは全結合層モデルである。このモデルでは、全結合層 (Dense) を 5 層連結した全結合型ニューラルネットワークモデルである。

2 つ目のモデルは CNN モデルである。CNN とは畳み込みニューラルネットワーク (Convolutional Neural Network) のことである。CNN には、特徴抽出を行う畳み込み層と、位置ずれを許容する働きを持つプーリング層を交互に階層的に配置したニューラルネットワークである [19]。CNN は画像認識等で一般的に利用される。このモデルでは、次元畳み込み層 (Conv-1d) とプーリング層 (MaxPooling) を組み合わせたブロックを 5 つ連結し、最後に全結合層で筋電位を出力する。

3 つ目のモデルは LSTM モデルである。LSTM は再帰型ニューラルネットワーク (Recurrent Neural Network: RNN) の一つである。RNN は時系列データの分析に用いられるニューラルネットワークであるが、長期間の時系列データの解析を行う際に非常に深いネットワークとなってしまう、そのために勾配消失や爆発問題が発生してしまうという問題が存在する。その問題を解決したのが、LSTM 層 (long short-term memory) である。このセルはネットワークが長期的な状態に格納すべきもの、長期的な状態から捨てるもの、読み取るべきものを学習することで、長期的なパターンを取り出すことができるようになる [20]。このモデルでは LSTM 層と全結合層から構成される。ここでは、姿勢情報と筋電位データは時間経過にしたがって変化するため、時系列データとする学習を行うことで時系列的な特徴を推定できるようにする。

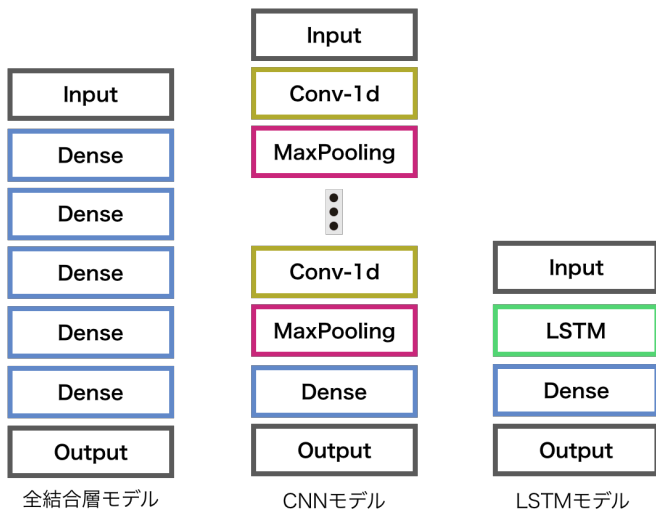


図 12 : 3 種類のモデルの概要

4. 実験

本論文で提案する 3 つのモデルの予測性能を評価するために、それぞれのモデルの学習を行った後に、未知のデータに対する予測実験を行いそれぞれのモデルの性能を評価する。

4.1. 実験手法

学習用のデータセットを作成し、それを訓練データとテストデータに分け、性能の評価を行った。

学習用のデータセットを作成するにあたって、トレーニング種目はスクワットを採用した。筋電位センサは両太腿に装着しトレーニング時の筋電位を計測した。トレーニングは被験者 1 名に対して 10 回*30 セットのトレーニングを行ってもらい動画データを撮影すると同時に筋電位を取得した。なお、30 セットの内訳としては、10 セットごとにセット a~c を用意した。詳細は表 2 に示す。

表 2 : セット a~c の詳細

	カメラに対する角度	スクワットの沈み込み度合い
セット a	正対	一定
セット b	正対	不定
セット c	1 セット毎に異なる角度	不定

取得した動画データに対しては、撮影位置を移動させてデータの拡張を行った。具体的には、動画を四隅の方向に移動させ合計 150 セット分の動画データデータを準備した。

上記の 150 セット分の動画データを PoseNet によって姿勢推定することで得られた 17 個の関節分の x,y 座標と関節に対する予測確率、ポーズ全体としての予測確率の 4 種類のデータを入力データとした。

データセットの概要を図に表したものを図 13 に示す。また、入力データは 68 系列、正解ラベルデータは

2 系列となるが、その 1 系列ごとに 0-1 の範囲に収まるように、1 セットごとに正規化を行った。

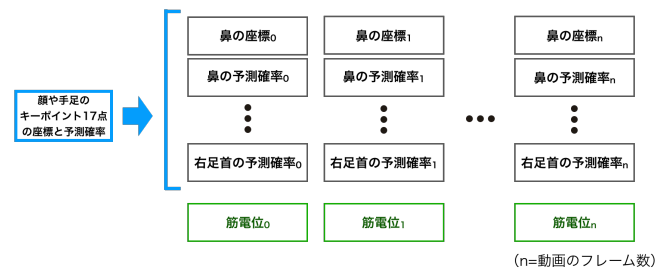


図 13 : データセット概要

全結合層モデルは、全結合層のみであり、最終層の 1 層のユニット数は 1 とした。そして、それ以外の 4 層のユニット数は 128 とした。また、活性化関数は relu 、最適化関数は adam 、誤差計算は平均二乗誤差、バッチサイズは 32 とした。

CNN モデルは、畳み込み層 (Conv-1d) の出力フィルタ数は 64、カーネルサイズは 8 とし、Pooling 層 (MaxPooling) の PoolingSize は 3 とした。全結合層のユニット数は 1 とした。また、活性化関数等は NN モデルと同様とした。

LSTM モデルは、LSTM 層の隠れ層は 300 とし、全結合層のユニット数は CNN モデルと同様に 1 とした。また、活性化関数は linear とし、最適化関数等は他のモデルと同様にした。LSTM モデルは時系列のデータを入力とする必要があるため、LSTM モデルのみ、進み 3 フレーム分のデータを、1 フレームずつずらしたデータセットを新規作成し、入力とした。

上記 3 つのモデル全ての epoch 数は 1000 としたが、過学習を避けるために validation loss が 5 epoch 分学習しても、下がらなければ学習を打ち止めにする設定で学習を行ったため、実際は 3 つのモデル共、1000 epoch 分の学習が行われることは無かった。

モデル作成の後、訓練データに加えていない未知のデータセット (被験者は 2 人とし、内 1 人は訓練データと関係のない別の被験者) の動画データのみから筋電位予測を行った。

4.2. 実験結果

学習データを取得した被験者とは別の被験者の動画データで予測を行った際の各モデルの出力をプロットしたグラフを図 14~16 に示す。ここで、縦軸は筋電位であり、横軸はフレーム数である。なお、実験では両太腿の筋電位を予測しているが、グラフでは右太腿のみの予測筋電位を示している。

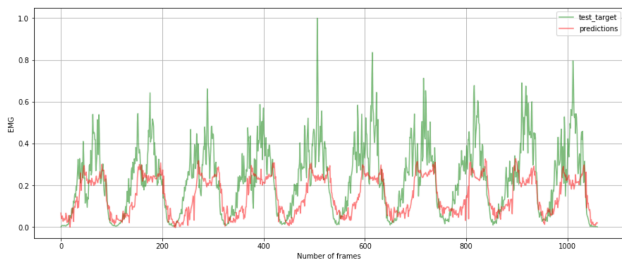


図 14：全結合層モデルでの予測結果

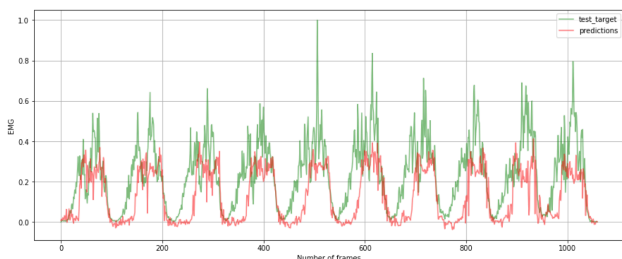


図 15：CNN モデルでの予測結果

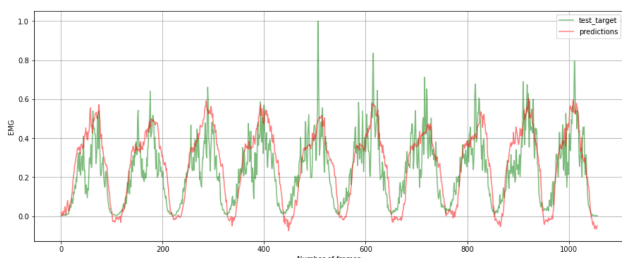


図 16：LSTM モデルでの予測結果

次に、それぞれのモデルにおける動画 1 フレームあたりの実測筋電位と予測筋電位の平均誤差を図 17 に示す。

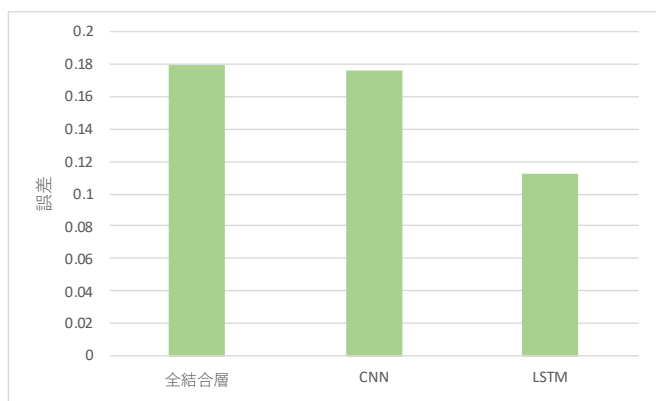


図 17：各モデルの平均誤差

4.3. 考察

予測実験結果より、時系列データとして解析を行った LSTM モデルの平均誤差が他の 2 つのモデルよりも

低いことや、実測筋電位の波形の特徴を大まかに捉えていることなどから、LSTM モデルが実際のサポートアプリに適していることが分かった。また、このことより、提案手法でトレーニング中の筋電位を予測可能であることが解った。

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、未知のデータに対しての予測実験を行い、実験結果より、未知のユーザーの動画データのみから筋電位を予測できることを示した。

今後の課題として、まず、精度を高めるために、関節の座標だけではなく、姿勢画像も訓練データに加えて学習を行うことを考えている。また、複数のセンサ数を装着した状態でデータを採取し、複数のトレーニング種目に対応できるかを検証していく予定である。

また、動画から推定された予測筋電位の可視化、及び、予測筋電位を基にしたトレーニング中の運動量などをユーザーにフィードバックする機能を有するサポートアプリのプロトタイプシステムの制作及び、サポートアプリを使用した状態での運動効果の向上度合いを調査する被験者実験等も行う予定である。

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, “Measec ~筋 (近) 未来サポートシステム メーカー~”, 2020 年 12 月閲覧, <https://www.ncp.co.jp/measec/>
- [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", インタラクション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] Kohei Yamamoto, Ryosuke Hiromatsu, Tomoaki Ohtsuki, "ECG Signal Reconstruction via Doppler Sensor by Hybrid Deep Learning Model With CNN and LSTM", IEEE Access (Volume:8) (p. 130551-130560), 2020
- [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] COCO Common Objects in Context, "COCO 2020 Keypoint Detection Task", 2020,
<https://cocodataset.org/#keypoints-2020>
- [19] 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
- [20] 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