

# 寝姿圧力動画からの睡眠深度分類の検討

岩瀬 慶<sup>1,a)</sup> 榎堀 優<sup>1</sup> 小野瀬 良佑<sup>1</sup> 吉田 直人<sup>2</sup> 間瀬 健二<sup>1</sup>

**概要：**高齢者の介護時の重大な事故の一つとしてベッドからの起床時の転倒がある。転倒予防のために、睡眠深度を判定し、睡眠が浅い状態の通知により起床行動に対応できるようにする。脳波から睡眠深度を分類する手法が多く高精度であるが、脳波の計測は日々の睡眠のモニタリングには向かない。そこで、日常的な計測が可能なシート型布圧力センサを用いて、睡眠時ベッド面にかかる圧力の変化から睡眠深度を分類する手法を検討した。本研究では1次元の脳波信号を入力とする深度分類手法の一つである U-Time を基礎としつつ、U-Time 入力前の画像特徴抽出、ならびに、U-Time 出力後の Bi-LSTM による双方向時系列考慮を追加した。被験者1名72日分の睡眠データで4段階の睡眠深度を分類した結果、正解率は0.644、F値のマクロ平均は0.637の精度となった。U-Time 単体の利用に比べ、正解率は0.029、F値のマクロ平均は0.053改善した。

**キーワード：**寝姿勢、圧力画像、睡眠深度分類、画像特徴抽出、双方向時系列考慮

## 1. はじめに

高齢者の介護時の重大な事故の一つとしてベッドからの起床時の転倒がある。介護施設にて、夜間は見守りを行う人員が減少するため、高齢者が一人の起床時に転倒などの事故が発生してしまう場合がある。実際に介護者の見守りがいない状況における高齢者の転倒、転落事故がベッド周辺にて多数報告されている [1]。ベッドからの起床時の転倒予防として、様々なセンサを用いたベッド上の動作検知が研究されている [2][3]。しかし、その多くは起床動作の検知であり、介護者が駆けつけるまでの時間が考慮されていない。転倒を防ぐためには、高齢者が起床時に介護者が傍に居り、いつでも対処できる状態であることが求められる。しかし、起床か否かの情報のみでは、長い睡眠時間の中から短い起床行動を予測するのは難しい。そこで、睡眠深度を判定し、睡眠が浅い状態の通知により高齢者の起床行動に対応できるようにする技術が求められている。睡眠は深い眠りと浅い眠りを周期的に繰り返すことが知られており、起床行動の多くは浅い眠りの状態から発生する。判定した睡眠深度から、深い眠りから浅い眠りへの変化を検知し介護者へ通知することにより、高齢者の起床行動を予測し事前の対応が可能となる。

睡眠深度は、睡眠の質の評価や多くの睡眠に関連した疾病を診断する要素となっており、広く睡眠深度の分類が研究されている。従来の睡眠深度を分類する研究では、睡眠時の脳波から分類するものが多く高精度であることが知られている [4][5]。しかし、脳波の計測には専用の機器を頭に装着し続ける必要があり、日々の睡眠のモニタリングには向かない。そこで、近年ではウェアラブルセンサやレーザーなどの非接触センサで計測した、体動や呼吸から睡眠深度を分類する手法が注目されている [6][7]。

我々は、布圧力センサで構成したシート型の2D圧力センサ（以下、シート型布圧力センサ）[8]により計測される、ベッド面から得られる圧力値を用いた高齢者の見守りについて研究している。シート型布圧力センサは、カメラやウェアラブルセンサなど他のセンサによる見守りに比べ、プライバシー問題や布団などによるオクルージョン問題が少なく、非侵襲で計測が可能なため、高齢者の見守りに適している。また、ベッド面から得られる圧力値の変化から睡眠時の体動等を計測できるため、シート型布圧力センサで計測した時系列の寝姿勢の圧力画像（以降、寝姿圧力動画）から睡眠深度の分類が可能と思われる。そこで本研究では、シート型布圧力センサで計測した寝姿圧力動画からの睡眠深度分類を目指す。

従来の睡眠深度を分類する手法は、単一または複数の1次元信号を入力とするものが多い。それに対し本研究では時系列の2次元圧力値を入力とする。睡眠深度の分類には体動の大きさが関係しているため、画像特徴により体動の

<sup>1</sup> 名古屋大学 情報学研究科  
Graduate School of Informatics, Nagoya university

<sup>2</sup> 名古屋大学 未来社会創造機構  
Institute of Innovation for Future Society, Nagoya university

a) iwase@cmc.is.i.nagoya-u.ac.jp

大きさや場所を考慮し、分類精度を向上できる可能性がある。そこで、1次元の脳波信号を入力とし睡眠深度を分類する手法の一つである U-Time[5] に、畳み込み層による画像の特徴抽出と、双方向の時系列考慮手法である BiLSTM[9] を追加で、時系列圧力画像から高精度に睡眠深度を分類する手法を提案する。

本稿の構成は以下のとおりである。第2節にて、関連研究を述べる。第3節にて、寝姿圧力動画からの睡眠深度の分類手法を述べる。第4節にて、提案手法の睡眠深度の分類精度評価実験を述べる。最後に、第5節にてまとめと今後の展開を述べる。

## 2. 関連研究

人物の睡眠をセンシングし、睡眠状態を把握する研究は広くされている。本節では既存の研究を述べ、本研究の立ち位置を示す。

睡眠深度は、睡眠時の脳波信号に各段階の特徴がみられるとして、脳波信号から専門家により手動で判別がされてきた。その後、情報技術による自動化が注目され、機械学習や深層学習により脳波信号から睡眠深度を自動で分類する研究が多くされた。Akaraらは、異なる幅のCNNによる時系列考慮とBiLSTMで脳波信号から5段階の睡眠深度を0.77のF値で睡眠深度を分類した[4]。Perslevらは、畳み込み層のみのEncoder-Decoderモデルで脳波から5段階の睡眠深度を0.79値で睡眠深度を分類した[5]。しかし、脳波の計測には専用の機器を必要とし、日々の睡眠のモニタリングには向かない。

睡眠深度を判別する脳波以外の要素として、睡眠時の体動や心拍、呼吸の特徴がある。日々の睡眠時にも計測可能なウェアラブルデバイスを用いた睡眠深度の分類が研究されている。Fujimotoらはスマートウォッチから計測した体動と脈拍間隔(PPI)から4段階の睡眠深度を0.62のF値で分類した[6]。また、睡眠時のデバイス装着による不快感やデバイス管理などの負担を少なくするため、非装着型や非接触型センサによる睡眠深度の分類が研究されている。HongらはCWドップラーレーダーで呼吸、心拍、体動を非接触で計測し4段階の睡眠深度を0.81の正解率で分類した[7]。

シーツ型布圧力センサは、プライバシー問題や布団などによるオクルージョン問題が少なく、非侵襲で計測が可能なため高齢者のベッド環境の見守りに適しており、睡眠深度を判定する要因となる体動等を計測できるため、本研究ではシーツ型布圧力センサで計測した寝姿圧力動画から睡眠深度を分類する。

## 3. 画像特徴抽出と双方向の時系列考慮による寝姿圧力動画からの睡眠深度分類

本節では、1次元の脳波信号から睡眠深度を分類する

U-Time[5]に、画像特徴の抽出とBiLSTMによる双方向の時系列考慮の2つの拡張を追加した、寝姿圧力動画からの睡眠深度分類手法を述べる。ネットワーク構造の概要図を図1に示す。

提案手法の拡張元のU-Time[5]は、シングルチャネルの脳波信号から睡眠深度を分類する手法の一つである。ネットワーク構造は画像の領域分類手法の一つであるU-Net[10]をベースとし、時系列方向に対して畳み込みを行うEncoder-Decoderモデルである。すべて畳み込み層で構成されており、他のRNNによる睡眠深度分類手法に比べてデータ毎のネットワーク構造の最適化の必要がないため、我々が用いる時系列圧力値に対してもネットワーク構造を大きく変えることなく分類が可能であると考えた。本手法では、エンコーダとデコーダの各畳み込みの幅を5、プーリング幅を1、2段目を4、3、4段目を2とした。

U-Timeで寝姿圧力動画から睡眠深度を分類した結果、一定の精度による分類は可能であった。しかし、U-Timeは1次元の時系列データを入力としているため、画像特徴を考慮していない。睡眠深度の分類には体動の大きさが関係しているため、画像特徴により体動の大きさや場所を考慮し、分類精度を向上させることが出来る可能性がある。加えて、既存の睡眠深度分類手法[4]より、畳み込み層による近接の時系列考慮に、BiLSTMによる双方向の時系列考慮の追加による精度の向上が示されており、U-Timeは畳み込み層のみで構成されたネットワークのため、同様に精度が向上が期待できる。そこで、U-Timeによる時系列考慮前に画像方向の特徴量抽出を追加し、U-Time後にBiLSTMによる双方向の時系列考慮を追加した。以降にて、画像特徴抽出とBiLSTMによる双方向の時系列考慮の追加手法の詳細を述べる。

### 3.1 画像特徴の抽出

画像特徴を考慮するために、U-Timeのエンコーダとデコーダの畳み込み層で時系列の特徴を抽出する前に、画像の特徴を抽出する畳み込みネットワークを追加した。ネットワークの詳細を図2に示す。画像方向に畳み込む、幅 $3 \times 3$ の畳み込み層を2回と幅 $2 \times 2$ のプーリング層を1回を1ブロックとし、これを3回繰り返した。その後、幅 $3 \times 3$ の畳み込み層を2回した後、画像方向に平滑化しU-Timeのエンコーダの入力とした。U-Timeのエンコーダとデコーダでは、時系列方向にのみ畳み込み、画像方向の畳み込みはしない。

### 3.2 BiLSTMによる双方向の時系列考慮

畳み込み層による近接の時系列考慮とBiLSTMによる双方向の時系列考慮を組み合わせるために、U-TimeのDecoderの出力後の畳み込み層の間にBiLSTM層を追加した。ネットワークの詳細を図3に示す。U-Timeではデ

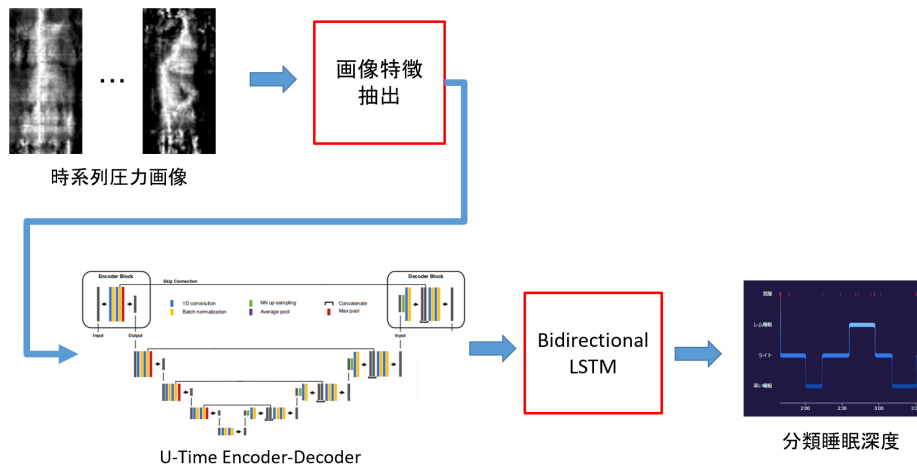


図 1 睡眠深度分類手法概要 (一部 [5] 図 S.2 より引用)

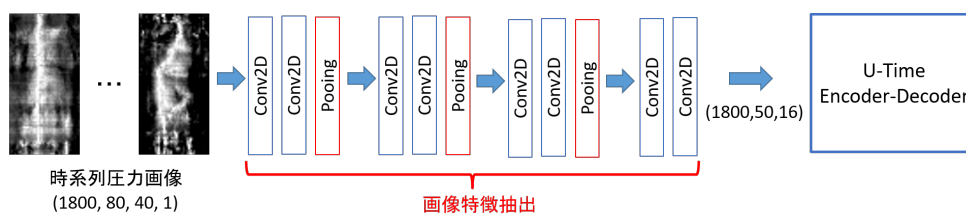


図 2 U-Time への画像特徴抽出の追加

コーダの出力後に、畳み込み層で睡眠深度の段階数の特徴量に変換，出力深度ラベル数になるように時系列方向にプーリング層，もう一度畳み込み層を通して睡眠深度段階を出力している。本手法では、畳み込み層で睡眠深度の段階数の特徴量に変換後，時系列方向にのみプーリングし画像方向にはプーリングせず BiLSTM へ入力した。BiLSTM の出力後に U-Time と同様にもう一度畳み込み層を通して睡眠深度段階を出力とした。

#### 4. 精度評価実験

本節では、第 3 節にて提案した睡眠深度分類手法の精度評価を行った実験を述べる。まず、被験者 1 名の長期データで提案手法の効果を検証した。次に、複数被験者のデータで複数の人物のデータに対して提案手法で睡眠深度が分類可能か検証した。第 4.1 節にて精度評価に用いたデータセットを述べる。第 4.2 節にて被験者 1 名の長期データを用いた評価実験，第 4.3 節にて複数被験者のデータを用いた評価実験を述べる。

##### 4.1 データセット

提案手法における睡眠深度分類精度を評価するために、睡眠時の圧力値及び睡眠深度データを収集した。その計測及び作成方法を以降にて述べる。

##### 4.1.1 計測方法

本計測では、睡眠時のベッド面にかかる圧力値と睡眠深度を計測した。計測環境の模式図を図 4 に示す。圧力値は

先行研究 [8] と同様のシート型布圧力センサで測定した。センサの計測点は 3200 (40 × 80) 点，大きさ約 90 × 180 cm である。計測に用いたセンサのサンプリング周期は多少の変動があるため，約 6Hz で計測した後，1 Hz でリサンプリングした。シート型布圧力センサをマットレスの上に設置し，その上に一般的な寝具用のシーツを敷き，その上で睡眠をとった。睡眠深度はスマートウォッチの Fitbit inspire2[11] で計測した。Fitbit は睡眠時に手首に装着することで自動で入眠と起床を検出し，30 秒ごとの睡眠深度を 4 段階（深い眠り：Deep，浅い眠り：Light，レム睡眠：Rem，覚醒：Wake）で判定する。睡眠深度の判定には睡眠時の体動や心拍などを用いている。

被験者は健康な男子大学生 3 名，成人男性 1 名，成人女性 1 名の 5 名 (S1-S5) を対象とした。計測場所は各被験者の自宅であり，マットレスや枕，かけ布団などの寝具は各被験者が日常使用しているものを用いた。

##### 4.1.2 データセット作成方法

計測したデータの処理方法を述べる。計測した圧力値と睡眠深度は，スマートウォッチの検出した入眠から起床時刻までのデータを用いた。圧力値は 1 Hz でリサンプリングした後，一つ前の圧力画像との差分を取り，時系列の圧力差分画像を生成した。その後，計測日毎の一連のデータで正規化した。睡眠深度は計測した 4 段階をそれぞれ 4 つのラベルに割り当てた。

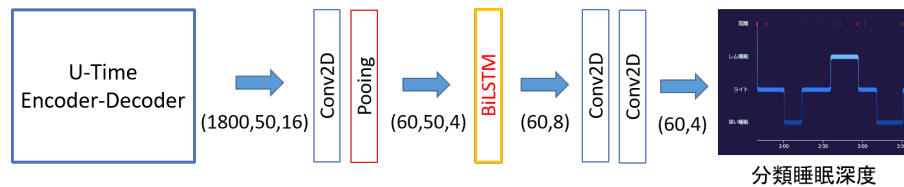


図 3 U-Time への BiLSTM の追加



図 4 計測環境

表 1 各睡眠深度のラベル数

	Deep	Light	Rem	Wake	合計
ラベル数	7890	23534	11798	5258	48480

#### 4.2 被験者 1 名のデータによる精度評価実験

本節では、提案手法の効果を検証するための、被験者 1 名の長期データによる睡眠深度の分類精度の評価実験を述べる。計測データは健康な男子大学生 1 名 (S1) の合計 72 日分の睡眠データを用いた。72 日分の各睡眠深度毎のラベル数及び合計は表 1 となった。

##### 4.2.1 評価方法

本評価では、データ入力を時系列圧力差分画像 1800 枚 (1800 秒)、出力を 4 段階の睡眠深度 60 ラベル (1800 秒) とした。学習の Loss は正解率を向上させる交差エントロピー誤差 (Cross Entropy Loss) と F 値を向上させる Dice 係数を用いた Loss (Dice Loss) とした。

比較条件について。まず、分類ネットワークの構造は、U-Time に圧力画像特徴抽出 (IFE: Image Feature Extraction) と BiLSTM の追加の有無の 4 条件で精度を比較した。時系列特徴を抽出しない場合は、入力の時系列圧力画像を時系列の 1 次元圧力値に変形し、U-Time の入力とした。次に、U-Time に BiLSTM を追加したモデルにて、圧力画像特徴抽出の畳み込み回数を比較をした。4.3 節にて述べた画像特徴抽出における畳み込みブロック (畳み込み層 2 回、プーリング層 1 回) の回数を、1 から 4 回繰り返すモデルで精度を比較をした。最後に、U-Time に 2 つ拡張を加えたモデルにて、Loss が交差エントロピー誤差のみ、交差エントロピーと Dice 係数の Loss の場合で精度を比較した。

本評価実験の学習・評価時は、計測日単位で 72 日分のデータを 8 分割し (d1-d8)、表 2 のデータ選択方法に従い学習・評価・テストのデータグループを選択した。グルー

表 2 データ選択方法

Pattern	Training data	Validation data	Test data
$D_{t8}^{v7}$	d1-d6	d7	d8
$D_{t7}^{v6}$	d1-d5,d8	d6	d7
$D_{t6}^{v5}$	d1-d4,d7-d8	d5	d6
$D_{t5}^{v4}$	d1-d3,d6-d8	d4	d5
$D_{t4}^{v3}$	d1-d2,d5-d8	d3	d4
$D_{t3}^{v2}$	d1,d4-d8	d2	d3
$D_{t2}^{v1}$	d3-d8	d1	d2
$D_{t1}^{v8}$	d2-d7	d8	d1

プ分割の際は、各グループ間で睡眠時間の合計の差が大きく生じないように、1 日あたりの平均睡眠時間がおよそ同じに分割した。学習用データを 54 日分、検証用データを 9 日分、テスト用データを 9 日分とした。以降データの選択パターンを  $D_{t8}^{v7}$  のように表記する。これは d7 を検証データ、d8 をテストデータ、残りを学習データとしたパターンを示す。睡眠深度の精度評価指標は正解率 (accuracy)、各睡眠深度の F 値 (F1-score)、F 値のマクロ平均 (macro mean) を用いた。

##### 4.2.2 評価結果

まず、画像特徴抽出、BiLSTM の追加の有無の 4 条件の精度比較結果について。正解率及び F 値による評価結果を表 3、各睡眠深度の正解ラベル数で正規化した結果で色づけした混同行列を図 5 に示す。画像特徴抽出 (IFE) を追加した場合、U-Time のみの場合に比べ、正解率は 0.015、F 値のマクロ平均は 0.040 向上した。各睡眠深度段階について、深い眠り、レム睡眠、覚醒の F 値が向上したが、浅い眠りのみ低下した。BiLSTM を追加した場合、U-Time のみの場合に比べ、正解率、F 値ともに低下した。各睡眠深度段階についても、すべての段階で精度が低下した。画像特徴抽出と BiLSTM の両方の追加により、U-Time のみに比べ、正解率は 0.029、F 値のマクロ平均は 0.053 向上し、正解率が 0.644、F 値のマクロ平均が 0.637 で分類でき、他の条件と比較して最も良い精度となった。また、すべての睡眠深度で U-Time のみの場合に比べ精度が向上し、覚醒以外の深度は他の条件と比較して最も良い精度となった。U-Time のみと提案手法による分類結果例を図 6 に示す。図 6 より、睡眠深度をより正確に分類していることが分かる。

次に画像特徴抽出の畳み込みブロックの回数の違いの比較について。正解率及び F 値による評価結果を表 4 に示

表 3 各追加手法条件の比較結果

method	accuracy	F1-score				
		macro mean	Deep	Light	Rem	Wake
U-Time	0.615	0.584	0.653	0.691	0.343	0.649
U-Time+IFE	0.630	0.624	0.666	0.679	0.446	<b>0.705</b>
U-Time+BiLSTM	0.551	0.500	0.451	0.653	0.329	0.568
U-Time+IFE+BiLSTM	<b>0.644</b>	<b>0.637</b>	<b>0.678</b>	<b>0.692</b>	<b>0.475</b>	0.703

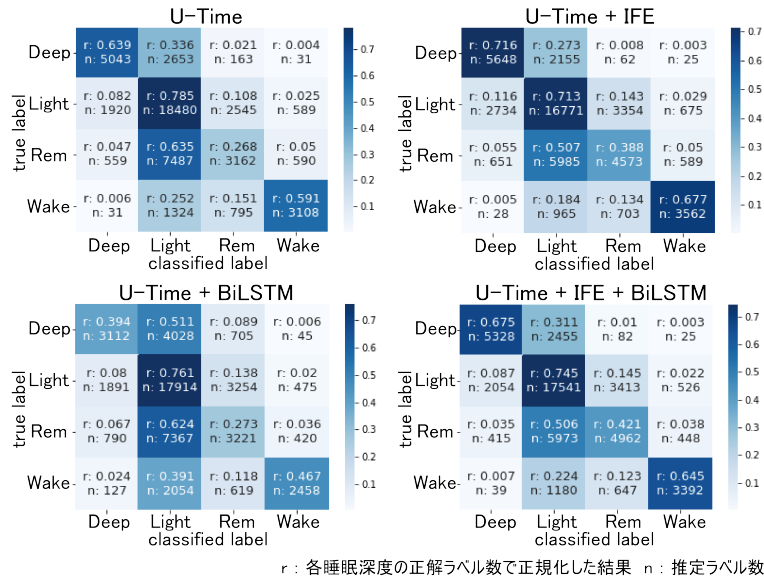


図 5 各追加手法条件の混同行列

表 4 畳み込みブロック回数比較結果

Conv Block	accuracy	macro mean	F1-score			
			Deep	Light	Rem	Wake
0	0.551	0.500	0.451	0.653	0.329	0.568
1	0.584	0.552	0.521	0.666	0.398	0.625
2	0.609	0.588	0.612	0.678	0.396	0.665
3	<b>0.644</b>	<b>0.637</b>	0.678	0.692	<b>0.475</b>	0.703
4	0.637	0.620	<b>0.682</b>	<b>0.695</b>	0.398	<b>0.705</b>

表 5 Loss 条件比較結果

Loss	accuracy	macro mean	F1-score			
			Deep	Light	Rem	Wake
CE	0.642	0.631	0.650	<b>0.696</b>	0.470	<b>0.707</b>
CE+DL	<b>0.644</b>	<b>0.637</b>	<b>0.678</b>	0.692	<b>0.475</b>	0.703

す。正解率と F 値のマクロ平均ともに、畳み込みブロックは 3 回まで追加すればするほど精度が良くなった。4 回追加した場合、深い眠り、浅い眠り、覚醒の F 値は最良となり、3 回追加した場合、正解率、F 値のマクロ平均、レム睡眠の F 値は最良精度となった。畳み込みブロックを 1 回と 2 回追加した場合は、U-Time のみに比べ精度が劣り、3 回と 4 回の場合は U-Time のみより良い精度となった。

最後に、交差エントロピー誤差のみ (CE) と交差エントロピー誤差と Dice 係数の Loss の場合 (CE + DL) の比較について。正解率及び F 値による評価結果を表 5 に示す。

交差エントロピー誤差と Dice 係数の Loss を用いた場合、交差エントロピー誤差のみの場合と比べ正解率は 0.002、F 値のマクロ平均は 0.007 向上した。各睡眠深度段階について、深い眠り、レム睡眠の F 値は Dice 係数の Loss を追加した方が良く、浅い眠り、覚醒は交差エントロピー誤差のみの方が良い。

#### 4.2.3 考察

画像特徴抽出について。精度評価結果より、正解率及び F 値のマクロ平均が大きく向上したため分類精度の向上効果があると分かった。特にレム睡眠の F 値が大きく改善しており、図 5 よりレム睡眠を浅い眠りと誤分類するのが少なくなっている。画像特徴の考慮により、体動の大きさや姿勢の変更の仕方を考慮し、覚醒度合いをより細かく判断できたため、精度が向上したと思われる。また、画像特徴抽出の畳み込みの回数について、本手法では畳み込みブロックを 3 回繰り返す場合が最適であった。プーリング層を 3 回するため、80 × 40 の大きさの寝姿体圧画像を 10 × 5 まで畳み込み、各ブロックの体動を考慮するのが睡眠深度の分類に最適であることが分かった。

BiLSTM による双方向の時系列考慮の追加について。画像特徴抽出を追加せず BiLSTM を追加した場合には、U-Time のみに比べ精度が低下した。また、画像特徴抽出の畳み込みブロックの繰り返し回数が 2 回以下の場合に

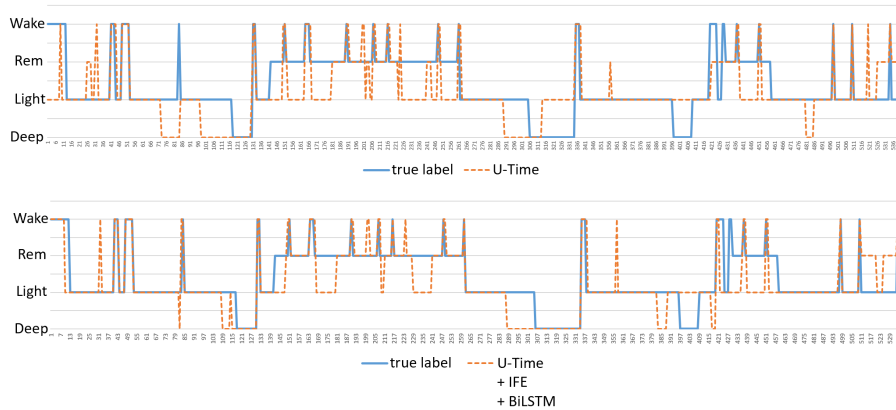


図 6 分類結果例

は、U-Time のみに比べ精度が低下し、3 回以上の場合には精度が向上した。U-Time 後の BiLSTM の入力特徴は、時系列方向にのみプーリングし、画像方向に平滑化したものである。そのため画像特徴抽出が不十分の場合、BiLSTM による時系列考慮の際に不要な特徴が多く含まれ、逆効果となったと考える。画像特徴抽出が十分であれば、先行研究 [4] と同様に、BiLSTM の追加で精度向上が見られた。U-Time における畳み込み層でも時系列考慮がなされているが、BiLSTM の追加により全体的かつ双方向の睡眠深度の関係性を考慮でき精度が改善できたと思われる。

Dice 係数の Loss の効果について、表 5 より交差エントロピー誤差に Dice 係数の Loss の追加で正解率は少し向上し、F 値も向上が見られた。本実験に用いたデータは睡眠深度のラベル数に偏りがあるため、正解率のみ向上するように学習するよりも、F 値も向上するように学習することで、精度の向上ができたと思われる。しかし、本実験にて Dice 係数の Loss のみでは学習が進まず分類モデルを作成できなかった。Dice 係数の Loss は全体の F 値が向上するように学習を進めるため、学習の初期段階では F 値の変動が小さくうまく学習が進まなかったと考えられる。学習の経過とともに各 Loss の重みの比率を変化させることにより、精度が向上する可能性がある。

本実験では、被験者 1 名 (S1) のみのデータで精度を評価した。そこで次節では、複数の人物の睡眠データに対して同様に提案手法により睡眠深度が分類できるか検証した。

#### 4.3 複数被験者データによる評価実験

本節では、複数の人物の睡眠データに対して提案手法で睡眠深度が分類できるか検証するための、被験者 5 名分のデータによる提案手法の睡眠深度の分類精度の評価実験を述べる。データは被験者 5 名の各 20 日分、合計 100 日分の睡眠データを用いた。被験者 S1 のデータは、より安定した睡眠のデータを用いるために、72 日のデータから睡眠時間が長い順に 20 日分選択した。各被験者 (S1-S5) の睡眠深度毎のラベル数及び合計は表 6 となった。

表 6 各睡眠深度のラベル数

被験者	Deep	Light	Rem	Wake	合計
S1	2475	7602	4047	1596	15720
S2	2992	7058	3867	1893	15810
S3	1890	8849	3560	1571	15870
S4	1787	10437	3109	2547	17880
S5	3697	7974	4035	1544	17250

表 7 人物毎に最適化する場合のデータ選択方法

Pattern	Training data	Validation data	Test data
$D2_{t5}^4$	d1-d3	d4	d5
$D2_{t4}^3$	d1,d2,d5	d3	d4
$D2_{t3}^2$	d1,d4,d5	d2	d3
$D2_{t2}^1$	d3-d5	d1	d2
$D2_{t1}^5$	d2-d4	d5	d1

表 8 未知の人物データに対する評価時のデータ選択方法

Pattern	Training data	Validation data	Test data
$D3_{t5}^4$	S1-S3	S4	S5
$D3_{t4}^3$	S1,S2,S5	S3	S4
$D3_{t3}^2$	S1,S4,S5	S2	S3
$D3_{t2}^1$	S3-S5	S1	S2
$D3_{t1}^5$	S2-S4	S5	S1

#### 4.3.1 評価方法

まず、被験者毎にそれぞれのデータで学習・評価し、人物毎に分類モデルを最適化する場合の提案手法の汎用性を検証した。学習・評価は、各被験者の計測日単位で 20 日分のデータを 5 分割し (d1-d5)、被験者毎にそれぞれ表 7 のデータ選択方法に従い学習・評価・テストのデータグループを選択した。グループ分割の際は、各グループ間で睡眠時間の合計の差が大きく生じないように、各グループの 1 日あたりの平均睡眠時間がおよそ同じに分割した。学習用データを 12 日分、検証用データを 4 日分、テスト用データを 4 日分とした。以降データの選択パターンを  $D2_{t5}^4$  のように表記する。これは d4 のデータを検証データ、d5 のデータをテストデータ、残りを学習データとしたパターンを示す。

表 9 被験者毎に最適化した場合の評価結果

被験者	accuracy	macro F1-score				
		mean	Deep	Light	Rem	Wake
S1	0.521	0.428	0.390	0.638	0.309	0.374
S2	0.426	0.178	0.028	0.595	0.091	0.000
S3	0.550	0.260	0.005	0.704	0.053	0.277
S4	0.657	0.384	0.015	0.769	0.013	0.737
S5	0.463	0.274	0.057	0.621	0.125	0.294

表 11 未知の人物データに対する評価結果

被験者	accuracy	macro F1-score				
		mean	Deep	Light	Rem	Wake
S1	0.557	0.425	0.370	0.682	0.002	0.646
S2	0.500	0.398	0.472	0.610	0.002	0.507
S3	0.569	0.468	0.432	0.698	0.098	0.646
S4	0.524	0.458	0.367	0.649	0.191	0.626
S5	0.502	0.360	0.287	0.639	0.000	0.515

表 10 学習済みモデルに対し被験者毎に最適化した場合の評価結果

被験者	accuracy	macro F1-score				
		mean	Deep	Light	Rem	Wake
S2	0.511	0.495	0.577	0.557	0.378	0.469
S3	0.588	0.409	0.242	0.728	0.048	0.619
S4	0.644	0.481	0.269	0.753	0.154	0.746
S5	0.531	0.503	0.595	0.584	0.347	0.487

次に、第 4.2 節の評価実験にて作成した、被験者 S1 の長期データを用いて学習した分類モデルを、被験者毎にそれぞれのデータでファインチューニングし、人物毎に分類モデルを最適化する場合の汎用性を検証した。被験者 S2-S5 の 4 名の各 20 日分のデータを用いて、被験者 S1 の 72 日分のデータで学習済みの分類モデルをファインチューニングした。データの選択方法は、ファインチューニングをしない場合と同様である。

最後に、他者のデータで学習した分類モデルにより評価し、未知の人物のデータに対する汎用性を検証した。学習・評価には、被験者単位でデータを分割し (S1-S5)、表 8 のデータ選択方法に従い学習・評価・テストのデータグループを選択した。学習用データを 3 名分 (60 日分)、検証用データを 1 名分 (20 日分)、テスト用データを 1 名分 (20 日分) とした。以降データの選択パターンを  $D3V_{15}^4$  のように表記する。これは被験者 S4 のデータを検証データ、被験者 S5 のデータをテストデータ、残りを学習データとしたパターンを示す。

本評価実験では、4.4.2 節の評価実験と同様に、データの入力を時系列圧力差分画像 1800 枚 (1800 秒)、出力を 4 段階の睡眠深度 60 ラベル (1800 秒) とし、学習時の Loss には交差エントロピー誤差と Dice 係数の Loss を用いた。ネットワークは、提案した画像特徴抽出と BiLSTM による双方向の時系列考慮を U-Time に追加したものを用いた。睡眠深度の精度評価指標には正解率、各睡眠深度の F 値、F 値のマクロ平均を用いた。

#### 4.3.2 評価結果

まず、被験者 5 名の各 20 日のデータで、被験者毎に表 7 に従って学習・評価し、人物毎に最適化した場合の分類精度を評価した。被験者毎の正解率及び F 値による評価結果を表 9 に示す。表 9 より、第 4.2 節の被験者 1 名による評価精度に比べ、正解率は被験者 S4 を除き低下し、F 値の

マクロ平均は全被験者にて低下した。各睡眠深度に注目すると、被験者 S1 は、第 4.2 節の被験者 1 名による評価精度に比べ、浅い眠りの分類精度はやや劣る。他の深度は他の被験者に比べると分類できているが、第 4.2 節の被験者 1 名による評価精度に比べると劣る。被験者 S2, S3, S5 は、第 4.2 節の被験者 1 名による評価精度に比べ、浅い眠りの分類精度は同等もしくはやや劣り、他の段階の精度は大幅に劣る。被験者 S4 は、第 4.2 節の被験者 1 名による評価精度に比べ、浅い眠りと覚醒の分類精度は高いが、深い眠りとレム睡眠の精度は大幅に劣る。

次に、被験者 S2-S5 の 4 名の各 20 日のデータを用いて、第 4.2 節の評価実験にて作成した被験者 S1 の 72 日のデータで学習した分類モデルを、被験者毎に表 7 に従ってファインチューニングし評価した。被験者毎の正解率及び F 値による評価結果を表 10 に示す。表 10 より、ファインチューニングしない場合に比べ被験者 S2-S5 の全被験者の F 値のマクロ平均が向上した。被験者 S2 と S5 は、第 4.2 節の被験者 1 名による評価精度に比べると精度は劣るものの、特定の深度の精度が極端に低いことがなく、睡眠深度を分類できた。被験者 S3 と S4 は、第 4.2 節の被験者 1 名による評価精度に比べ、浅い眠りと覚醒の精度は同等または良くなったが、深い眠りとレム睡眠は大幅に劣る。

最後に、被験者 S1-S5 の 5 名の各 20 日、合計 100 日のデータを用いて表 8 に従って学習・評価し、未知データに対する分類精度を評価した。被験者毎の正解率及び F 値による評価結果を表 11 に示す。表 11 より、正解率と F 値のマクロ平均が、第 4.2 節の被験者 1 名による評価精度に比べ、全被験者で低くなった。各睡眠深度に注目すると、浅い眠りと覚醒の分類精度は、第 4.2 節の被験者 1 名による評価精度に比べやや劣るが、深い眠りとレム睡眠の精度は大幅に劣る。特にレム睡眠は殆ど正しく分類できなかった。

#### 4.3.3 考察

まず、人物毎に分類モデルを最適化した場合について。表 9 の評価結果より、被験者 S1-S5 の 5 名のいずれの人物でも第 4.2 節における評価実験に比べ精度が劣り、殆ど分類できなかった深度があった。第 4.2 節における評価実験と比べ、同じ被験者 S1 でも、データ数が少ない場合は精度が劣った。本実験では 1 被験者当たり 20 日分のデータを用いており、第 4.2 節にて用いた被験者 S1 のデータ数

の72日より少ない。そのため、学習に用いた20日分のデータでは、データのバリエーションが少なく、分類が出来なかったと思われる。

次に、長期データで学習済みの分類モデルの人物毎のファインチューニングによる最適化について。表10の評価結果より、ファインチューニングにより、しない場合に比べ、正しく睡眠深度を分類できた。従って、長期データで学習した分類モデルを個人に合わせて最適化することにより、分類対象のデータが少ない場合にも分類が出来る可能性が示唆された。しかし、被験者S3やS4の深い眠りやレム睡眠の分類の精度は低いままであった。この要因として、被験者のデータのバリエーションが少ないことや、事前学習のデータとファインチューニングのデータの特徴が異なっていた可能性が挙げられる。

最後に、未知の人物のデータに対する分類について。表11の評価結果より、正解率やF値のマクロ平均は第4.2節における被験者1名の評価結果に比べ大きく劣り、特にレム睡眠の段階の分類が出来なかった。この要因として、被験者毎のデータの日数が足りない点や、被験者毎にデータの特徴が異なり、被験者数が足りない点が挙げられる。各被験者のデータで学習・評価した場合より、各被験者20日分のデータではバリエーションが少ない可能性があり、各被験者のデータの計測日数がより必要な可能性がある。また、各睡眠深度の周期の長さや発生頻度などの睡眠特徴は人物によって異なり、被験者毎に体格や計測環境が異なるため、計測された圧力の変化と睡眠深度の関係性が異なっていた可能性もある。そのため、学習に用いた被験者3名分のデータでは、学習に用いていないデータの特徴を捉えられる程のデータのバリエーションがなく、分類が出来なかったと思われる。

## 5. まとめ

本研究では、介護において重大事故の一つである転倒予防のために、シーツ型布圧力センサで計測した寝姿圧力動画から睡眠深度を分類する手法を提案した。U-Timeに画像特徴抽出、BiLSTMによる双方向の時系列考慮を追加し、寝姿圧力動画からの睡眠深度の分類精度の向上を図った。被験者1名の72日分のデータで精度を検証した結果、U-Timeのみに比べ、正解率は0.029、F値のマクロ平均は0.053向上し、正解率が0.644、F値のマクロ平均が0.637の精度で分類できた。提案手法の複数の人物に対する汎用性を検証するために、被験者5名の各20日分のデータで評価実験した。各被験者のデータで学習・分類した結果、正しく分類できなかったが、被験者1名の72日分のデータで学習したモデルを各被験者の20日分のデータでファインチューニングすることで分類できるようになった場合があり、分類対象の人物のデータが少なくても長期データで学習したモデルがあれば分類できる可能性が示唆された。

未知の人物のデータの分類を評価した結果、被験者5名の各20日分のデータでは正しく分類できなかったため、より多くの日数や被験者数のデータで学習する必要がある。

以上より、寝姿体圧動画から睡眠深度を分類でき、転倒の予防への利用が期待できる。今後は、より多くの日数、被験者のデータによる精度評価や睡眠深度の予測に取り組む予定である。

**謝辞** 本研究はJSPS 科研費21H03481の助成、ならびに、栢森情報科学振興財団の研究助成を受けたものです。

## 参考文献

- [1] 三田寺裕治, 赤澤宏平. 介護保険施設における介護事故の発生に関する分析. 社会医学研究, Vol. 30, No. 2, pp. 123-130, 2021.
- [2] Changchun Lu, Jian Huang, Zhi Lan, and Qiang Wang. Bed exiting monitoring system with fall detection for the elderly living alone. IEEE International Conference on Advanced Robotics and Mechatronics (ICARM), 2021.
- [3] Jose Clemente, Maria Valero, Fangyu Li, Chengliang Wang, and WenZhan Song. Helena: Real-time contact-free monitoring of sleep activities and events around the bed. IEEE International Conference on Pervasive Computing and Communications (PerCom), 2020.
- [4] Supratak, Akara, Dong, Hao, Wu, Chao, and Guo Yike. Deepsleepnet: a model for automatic sleep stage scoring based on raw single-channel eeg. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 25, pp. 1998-2008, 2017.
- [5] Perslev Mathias, Jensen Michael, Darkner Sune Jennum, Poul rgen, and Igel Christian. U-time: A fully convolutional network for time series segmentation applied to sleep staging. Advances in Neural Information Processing System, Vol. 32, pp. 4415-4426, 2019.
- [6] Fujimoto K, Ding Y, and Takahashi E. Sleep stage detection using a wristwatch-type physiological sensing device. Sleep and Biological Rhythms, Vol. 16, pp. 449-456, 2018.
- [7] Hong Hong, Li Zhang, Chen Gu, Yusheng Li, Guangxin Zhou, and Xiaohua Zhu. Noncontact sleep stage estimation using a cw doppler radar. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Vol. 8, pp. 260-270, 2018.
- [8] Yu Enokibori and Kenji Mase. Data augmentation to build high performance dnn for in-bed posture classification. Journal of Information Processing, Vol. 26, pp. 718-727, 2018.
- [9] M. Schuster and K. Paliwal. Bidirectional recurrent neural networks. IEEE Transactions Signal Process, Vol. 45, No. 11, pp. 2673-2681, 1997.
- [10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), Vol. 9351, pp. 234-241, 2015.
- [11] Fitness tracker with heart rate - shop fitbit inspire 2. <https://www.fitbit.com/global/us/products/trackers/inspire2>.