

リアルタイム性を考慮した自動睡眠ステージ判定システムの設計

国生 泰資[†] 山田 空^{††} 堀江 和正^{†††} 阿部 高志^{††††} 北川 博之^{††††}

[†] 筑波大学大学院 理工情報生命学術院 システム情報工学研究群 〒305-8573 茨城県つくば市天王台 1 丁目 1-1

^{††} 筑波大学 情報学群 情報科学類 〒305-8573 茨城県つくば市天王台 1 丁目 1-1

^{†††} 筑波大学 計算科学研究センター 〒305-8573 茨城県つくば市天王台 1 丁目 1-1

^{††††} 筑波大学 国際統合睡眠医科学研究機構 〒305-8573 茨城県つくば市天王台 1 丁目 1-1

E-mail: [†]{kokusho,kyamada,horie}@kde.cs.tsukuba.ac.jp, ^{††}abe.takashi.gp@u.tsukuba.ac.jp,

^{†††}kitagawa@cs.tsukuba.ac.jp

あらまし 本研究では、リアルタイム性を考慮した自動睡眠ステージ判定手法を提案する。これまでの手法では計測済みの睡眠データに対して睡眠ステージを判定するため、計測中に睡眠ステージを判定し、被験者の睡眠状態を把握することは困難であった。しかし、リアルタイムでの睡眠モニタリングや、特定の睡眠ステージにおいて特定の刺激を提示したい等、リアルタイム睡眠ステージ判定に対する睡眠医療、睡眠研究上のニーズが存在する。そこで本研究では、脳波計測器から流れるストリームデータに対し、リアルタイム性を持ったステージ判定を行うシステムを構築する。さらに、システムのパフォーマンスに影響する因子についても検討する。

キーワード ヘルスケア、睡眠ステージ判定、機械学習

1 はじめに

睡眠は人間の健康と活動に大きく関わっている。睡眠障害は情動障害の発生、作業効率の低下、交通事故のリスク増大を引き起こすとされている一方で、未だ明らかになっていない事象が多い。したがって、現在に至るまで睡眠に関する研究が盛んに行われてきた。睡眠の評価手法の一つに PSG (Polysomnophy) 検査がある。これは終夜を通して、被験者の脳波、眼球運動、顎筋電位、心電図、酸素飽和度等を記録し被験者の睡眠の深度や睡眠サイクルを評価する検査である。米国睡眠医学会が示すガイドラインによると睡眠はその深度により 5 つのステージに分けられる [2]。各ステージとその特徴を図 2 に示す。

従来ヒトを対象とした睡眠ステージ判定は、脳波等の生体信号を一定長のエポック (通常 30s 程度) に分割し、専門家が目視によって各エポックに対するステージラベルを決定づけることで行われてきた。しかし、被験者 1 人当たり 1 晩 (数時間) 相当の記録の全エポックを分析することは、高い専門性を必要とする上、作業に多くの労力を要してしまうことが課題となっていた。このような課題に対し、睡眠ステージ判定タスクを自動化する研究がこれまで行われてきた。近年ではニューラルネットワークの発展により、専門家と同等の精度を持つ判定モデルが作成可能となってきている。ちなみに、異なる専門家同士の睡眠ステージ判定結果の一致率は 80% 強とされている。

一方で、従来ほとんどの研究はリアルタイムに睡眠ステージ判定を行うことができない。リアルタイム睡眠ステージ判定には、リアルタイム睡眠モニタリングや、特定の睡眠ステージにおいて特定の刺激を提示したい等、睡眠医療、睡眠研究上のニーズが存在する。

以上を踏まえて、本論文では学習済みの睡眠判定モデル [8] を

組み込んだリアルタイムな判定システムを設計し、判定結果の妥当性や遅延について考察する。

2 前提知識

2.1 PSG 検査

PSG 検査では被験者の頭部や目、顎等の周辺に電極を取り付け、生体電位測定システムを用いて各部位の生体信号を計測する。脳波信号を EEG、眼球運動による信号を EOG、筋電位による信号を EMG と呼ぶ。専門家はこれらの信号を目視で分析することで睡眠ステージ判定を行う。PSG 検査により得られる EEG、EOG、EMG の例を図 1 に示す。

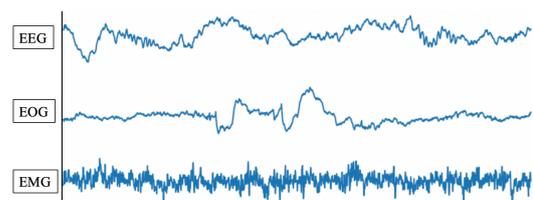


図 1 EEG, EOG, EMG 信号

2.2 睡眠ステージ判定

AASM の新基準 [2] に従い、得られた信号データに "Wake", "N1", "N2", "N3", "REM" のラベリングを行う。専門家によるラベリング作業は図 2 に示されるような特徴や前後エポックの睡眠ステージ等を考慮した上で判定される。一般的にこのラベリング作業は 30s のエポックごとに行われることが多い。

3 関連研究

現在に至るまでさまざまなヒトを対象とした睡眠ステージ判

睡眠ステージ	特徴
Wake	α 波(8-13Hz), まばたき, 急速眼球運動
REM	急速眼球運動, のこぎり波, 低頻筋電位
N1	低速眼球運動, 頭蓋頂鋭波
N2	K-complex波, 紡錘波
N3	低周波数波の活発化(0.5-20Hz)

図2 睡眠ステージと特徴

定モデルが提案されてきたが, その中でもニューラルネットワークを用いたモデルが高い精度を記録しており, 注目が集まっている. Tsinaliset ら [3] により 2 層の畳み込み層を用いた CNN モデルが発表されている. しかし, 1ch 分の EEG のみを入力とする簡素なモデルであるため正答率 0.71-0.76 にとどまっている. Supratak ら [4] により提案された DeepSleepNet は CNN と Bi-directional LSTM を組み合わせたモデルを用い, 前後情報を考慮した推測を可能とした. 結果として正答率は 0.81-0.86 を記録しているが, ステージの判定に 25 エポック分の情報が必要となる. Phan ら [5] により提案された SeqSleepNet についても正答率 0.87 を記録しており, 高精度であるが, 10-30 エポックの情報を必要とする. これらの研究では高い正答率を記録することを重視しており, リアルタイムにステージ判定をする場面を想定していない.

一方で, リアルタイム性を考慮した関連研究として, Koushik ら [6] は信号計測と睡眠ステージ判定処理それぞれの機能をプロセスレベルで分離し, 両者間のデータ転送に CFS 形式のファイルを使用したアーキテクチャを提案しており, およそ 5ms 毎に 1 エポック分のステージ判定を下すことができる. しかし, ある時点でのステージ判定には直後 5 エポック分のステージ判定結果を反映させるモデルであるため, 現実的にはリアルタイムにステージ判定を行うことはできない. Patanaik ら [7] は 1ch の EEG を用いたリアルタイムなステージ判定システムをスマートフォン上で構築している. 1ch の EEG のみであるが正答率は 0.83 と高い精度を記録している. しかし, 多チャンネルの信号に対応できないという問題点がある.

4 提案手法

本研究の目的は生体電位測定器からデータを取得し, 睡眠ステージを判定するリアルタイムシステムを提案することである.

対象とする生体電位測定器は Biosemi 社製 ActiveTwo を想定しており, 図3に示す通り, 計測器と睡眠ステージプログラム間のインターフェースとして TCP 通信を利用する. 本研究では, 実際の対象計測機器を用いたリアルタイム判定システムを構築する前段階として, データ配信サーバを用いたシミュレーションシステムを構築した. すなわち, 計測器を擬似的に模したサーバプログラムを用意し, 睡眠ステージ判定を行うクライアントプログラムとの通信を行うサーバクライアントシステムでシミュレーションを行った.

4.1 入力

本モデルへの入力データは, PSG 検査に用いられる計測器

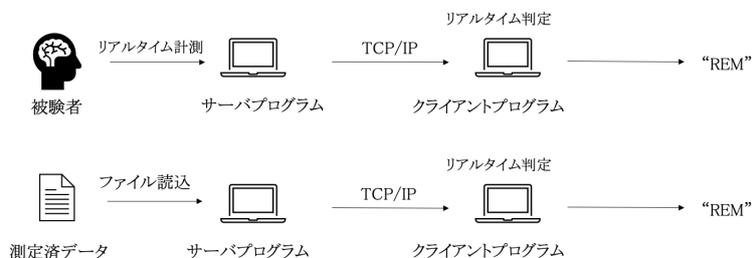


図3 想定アーキテクチャ(上)とシミュレーションアーキテクチャ(下)

PSG1100 で取得した $EEG \times 6, EOG \times 2, EMG \times 1$ の計 9 つの信号から成る. サンプル周波数はそれぞれ 200, 100, 200Hz である. また, 従来の非リアルタイム睡眠ステージ判定は 30s エポックを基本に行うことを踏まえ, 本研究においては直近 30s の入力データをベースとしたリアルタイム判定を行う. 以上を踏まえて, 以下のステージ判定モデルにおける特徴抽出ユニットに対する入力は EEG:(6, 6000, 1), EOG:(2, 3000, 1), EMG:(1, 6000, 1) となる.

4.2 ステージ判定モデル

ステージ判定モデルは特徴抽出ユニットと LSTM ユニットから構成される.

4.2.1 特徴抽出ユニット

このユニットでは我々の研究グループで PSG データ解析用に開発した Sleep-CAM モデル [8] における学習済みの特徴抽出部のみを再利用し, 入力信号の特徴量を LSTM ユニットへの入力とする.

4.2.2 LSTM ユニット

LSTM ユニットは特徴抽出ユニットから得られた直近 30s(1 エポック)に加えて, さらにその前の 30s の入力も考慮した直前 60s 分の特徴量を入力として受け取る. これにより直近 30s の前に至る前の状況を考慮したステージ判定が可能となる. この LSTM ユニットでは直近 30s の各時点についてのステージ確信度をそれぞれ出力する. 本研究におけるモデルでは, $LSTM_len(0 \leq LSTM_len \leq 30s)$ というパラメータを導入し, 直近何秒時点までの判定結果を平均してステージ判定を行うか設定可能にしている. LSTM ユニットの学習には, Sleep-CAM モデル用の 42 夜分のデータを用いた.

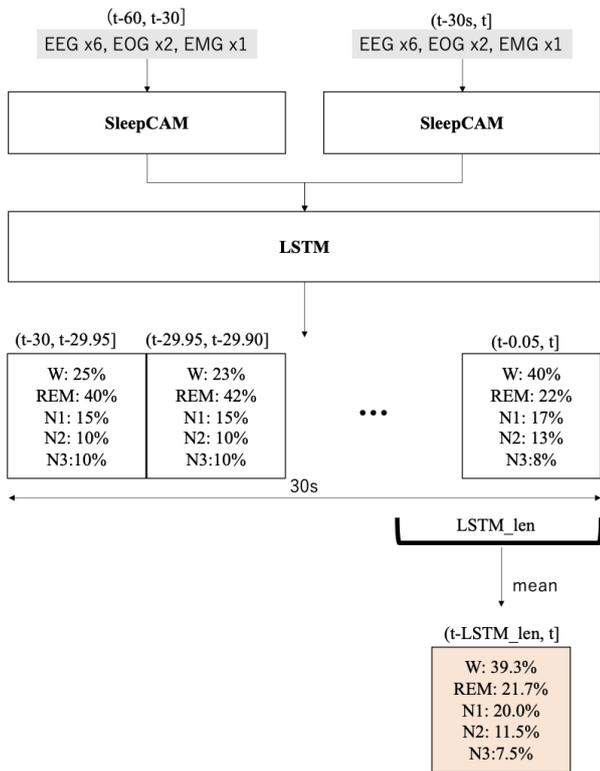


図 4 ステージ判定モデル

4.3 サーバプログラム

サーバプログラムでは実際の計測速度と同等な頻度で信号データをクライアントプログラムに送信する。1 パケットで送信するサイズは 10 サンプル (EOG は 5 サンプル) とした。サンプリング周波数は 200 (EOG は 100) Hz であるためパケットの送信間隔は 0.05s となる。

4.4 クライアントプログラム

本モデルでは、サーバプログラムから送られてくるデータに対して移動窓を用いて上記睡眠ステージ判定モデルで分析し、睡眠ステージ判定を行う。フローチャートを図 5 に示す。SLIDE 変数は移動窓を毎回何秒分移動するかを指定するパラメータであり、例えば SLIDE=10s であれば 10s 相当の信号データが到着するたびに判定を行う。計測開始直後から 60s 分のデータを受け取ると、そのデータに対して特徴量抽出と LSTM によるステージ判定を行う。以降はパラメータ変数 SLIDE に定められた増分サンプルを受け取るごとに特徴抽出、LSTM による睡眠ステージ判定を行う。図 5 ではサーバプログラムより、10 サンプル (0.05s 相当) を受け取り続け、12000 サンプル (60s 相当) まで受け取るとステージ判定を行っている。

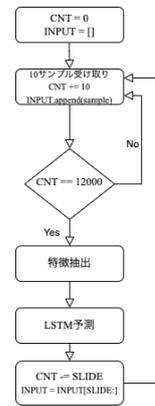


図 5 クライアントプログラム

5 実験

スライド幅 (SLIDE) を 30s, 5s, 1s とした場合それぞれについてリアルタイム計測を行う。この際、判定区間 (LSTM_len) は 30s と設定した。また、判定区間を 30s, 5s, 1s とした場合についてもリアルタイム計測を行う。この際、スライド幅は 1s と設定した。OS は Ubuntu LTS 20.0.4 を使用し、実装は Python の Tensorflow-gpu-2.1.0 を使用した。また、CPU は Intel Core i3-12100, GPU は Nvidia GeForce GTX 1660 SUPER を使用した。

5.1 データセット

実験に用いるデータセットは Sleep-CAM [8] 用に作成された、学習用データとは異なるテスト用の 7 夜分の専門家によるステージラベリング済みのデータを使用する。データセット長は計 62 時間 27 分に及ぶ。なお、テストデータはあらかじめノイズ除去や正規化処理がなされている。

5.2 実験結果

5.2.1 遅延の評価

図 6 に計測結果を示す。計測項目として、移動窓スライド幅分のパケットを受け取り、特徴抽出ユニットおよび LSTM ユニットの判定結果を得るまでの時間を "Iteration Time" とし、さらに Iteration Time の内訳としてパケット受け取りにかかる時間を "TCP Time", ステージ予測にかかる時間を "Scoring Time" としてそれぞれを計測した。

さらに、実際のリアルタイム性の評価指標として、サーバプログラムから該当エポックに対する計測信号パケットを送信し終えた時刻とクライアントプログラムでそのエポックに対するステージ判定を終えた時刻との差分を "Latency" として計測した。

計測結果よりスライド幅 30s, 5s, 1s の場合において、判定間隔である Iteration Time がそれぞれ 30.189, 4.998, 0.978 であるため、ほぼスライド幅で指定した間隔でステージ判定を行えていることがわかる。

どのスライド幅においても、Iteration Time はスライド幅と同等の時間を要していることが分かり、その内訳を見てみると、TCP パケットを受け取る時間が大きな割合を占めていること

がわかる。また、どのスライド幅においても判定モデルによる予測に要する時間は 0.6s ほどであることがわかる。これは今回の実験環境下におけるリアルタイムシステムの遅延時間の下限と考えられる。ただし、本論文では実際の信号生データに対するフィルタ処理やサンプリング等に要する時間を除外しており、今後その部分を含めた検討が必要である。

さらに Latency について見てみるとこの時間には判定エポック終端に相当する最終パケットの送受信と Scoring Time が含まれているが、ほとんど Scoring Time が支配的となっている。この Latency は本システムにおける応答時間を示す値であることから、本システムでは 0.6s 程度の遅延時間を持ったリアルタイム判定が可能であると言える。

	Iteration		TCP		Scoring		Latency	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
SLIDE=30s, LSTM_len=30s	30.189	0.122	29.572	0.085	0.610	0.072	0.618	0.072
SLIDE=5s, LSTM_len=30s	4.998	0.058	4.408	0.043	0.584	0.040	0.631	0.042
SLIDE=1s, LSTM_len=30s	0.978	0.046	0.389	0.033	0.582	0.033	0.619	0.041

図 6 遅延計測結果

5.2.2 正答率

30s ごとに専門家によりラベリングされたステージ正解ラベルと本ステージ判定モデルの判定ラベルの比較を行う。5.1 に記載した通り、実験で使用したデータセットの長さは 62 時間 27 分で、30s 区切りの正解ラベルの数は 7494 エポックとなる。スライド幅が 30s の場合はこの正解ラベルと純粋に比較可能であるが、30s より短い 5s や 1s の場合、より細かな正解ラベルがないため、厳密な判定はできない。そこで今回の実験結果においては 30s ごとに付与された正解ステージラベルを細分化し、1s または 5s 時点ごとに該当する正解ラベルを割り振ることで擬似的に正解ラベルを用意している。

スライド幅が 30s, 5s, 1s の場合についての正答率とカッパ係数を図 7 に示す。同様にして、判定区間が 30s, 5s, 1s の場合についての正答率とカッパ係数を図 8 に示す。計測結果より本モデルでは正答率が 0.88 であることがわかった。30s 以外のスライド幅においては正解ラベルが 30s ごとにしか付与されていないため、正答率が落ちていることが考えられる。そのため、今後、より細かな判定ラベルを用意し、厳密な評価を行う必要がある。しかしながら、どのスライド幅においても判定精度は約 0.83 ほどであり、関連研究と比較しても決して低い値ではないことから、スライド幅を狭めよりリアルタイム性のある判定システムを構築可能であることが示唆される。

また、判定区間を変更した場合については、判定精度に大きな差異は見られなかった。

	ACC	Kappa
SLIDE=30s, LSTM_len=30s	0.8839	0.8393
SLIDE=5s, LSTM_len=30s	0.8318	0.7405
SLIDE=1s, LSTM_len=30s	0.8277	0.7339

図 7 正答率とカッパ係数 (スライド幅を変更)

	ACC	Kappa
SLIDE=1s, LSTM_len=30s	0.8277	0.7339
SLIDE=1s, LSTM_len=5s	0.8327	0.7439
SLIDE=1s, LSTM_len=1s	0.8287	0.7387

図 8 正答率とカッパ係数 (判定区間を変更)

さらにリアルタイム判定したステージの時系列変化を表す、睡眠経過図 (ヒプノグラム) の 1 夜分を図 9 に、一部の拡大図を図 10 に示す。ここで専門家により作成された正解ラベルは 30s のエポックに対しての正解であることに注意されたい。

両図より SLIDE=30s, 5s, 1s の場合それぞれの間隔ごとに判定がなされている様子が確認できる。図 9 よりどのスライド幅においても正解ラベルと概ね同様なステージ遷移を捉えていることがわかる。図 10 よりスライド幅が細かいほど、より細かなステージ判定を行えていることがわかる。

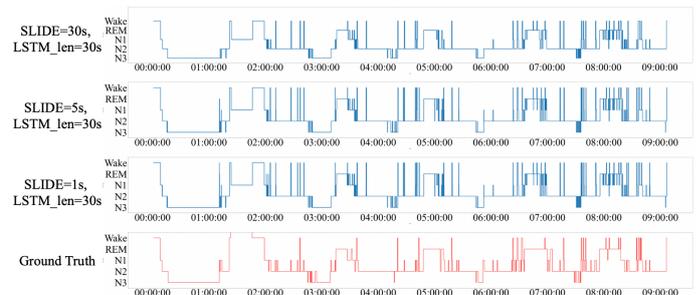


図 9 ヒプノグラム (全体)

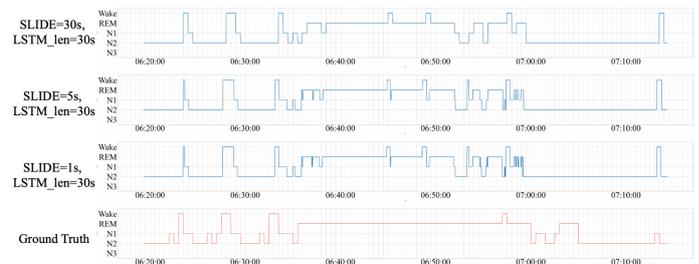


図 10 ヒプノグラム (拡大)

6 まとめ

本論文では学習済みの特徴抽出部を持つ機械学習モデルを組み込んだリアルタイムな睡眠ステージ判定システムを設計し、計測実験を通して遅延や判定結果について検討を行った。結果として多チャンネルを入力とするステージ判定モデルを使用したうえで 1s 間隔程度でステージ判定を行えることが分かった。これは十分にリアルタイム性を持つため、様々な睡眠研究に活用できる可能性が示唆される。

今後の課題として以下が挙げられる。

- リアルタイム判定結果に対する妥当性の評価のため、正解ラベルデータの収集等について検討を行う。
- 他計測器に対応した判定モデルを組み込み同様の計測実験を行い、様々なプラットフォームで本システムが稼働することを確認する。
- フィルタ処理や正規化、リサンプリングなど、生データを想定したシステムの設計を行う。

7 謝 辞

本研究の一部は、JSPS 科研費 JP19H04114, JP22K19802, AMED Grant Number JP21zf0127005, F-MIRAI 共同研究, (株) S'UIMIN 共同研究による。本研究の一部は、JSPS 科研費 JP19H04114, JP22K19802, AMED Grant Number JP21zf0127005, F-MIRAI 共同研究, (株) S'UIMIN 共同研究による。

文 献

- [1] E. F. Codd, "A Relational Model of Data for Large Shared Data Banks," *Communications of the ACM (CACM)*, Vol. 13, No. 6, pp. 377–387, 1970.
- [2] Iber, C., Ancoli-Israel, S., Chesson, A. L., and Quan, S., "The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications," 2007.
- [3] Orestis Tsinalis, Paul M. Matthews, Yike Guo, Stefanos Zafeiriou, *Automatic Sleep Stage Scoring With Single-Channel EEG Using Convolutional Neural Networks*, 2016
- [4] Akara Supratak, Hao Dong, Chao Wu, and Yike Guo, *Deep-sleepnet: a model for automatic sleep stage scoring based on raw single-channel eeg*, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 25, No.11, pp. 1998-2008, 2017.
- [5] Huy Phan, Fernando Andreotti, Navin Cooray, Oliver Y Chen, and Maarten De Vos. *Seqsleepnet: End-to-end hierarchical recurrent neural network for sequence-to-sequence automatic sleep staging*, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 27, No. 3, pp.400-410, 2019.
- [6] Abhay Koushik, Judith Amores, Pattie Maes, *Real-Time Sleep Staging using Deep Learning on a Smartphone for a Wearable EEG*, *IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2019.
- [7] Amiya Patanaik, Ju Lynn Ong, Joshua J Gooley, Sonia Ancoli-Israel and Michael W L Chee, *An end-to-end framework for real-time automatic sleep stage classification*, *Sleep*, Vol. 41, No. 5, 2018.
- [8] Kazumasa Horie, Leo Ota, Ryusuke Miyamoto, Takashi Abe, Yoko Suzuki, Fusae Kawana, Toshio Kokubo, Masashi Yanagisawa and Hiroyuki Kitagawa, *Automated sleep stage scoring employing a reasoning mechanism and evaluation of its explainability*, *Scientific Reports* 12, No. 12799, 2022.