

睡眠ステージ判定のためのスタイル変換に基づく 脳波信号のノイズ・個人差除去

大宮 直樹[†] 堀江 和正^{††} 北川 博之^{†††}

[†] 筑波大学大学院システム情報工学研究群情報理工学位プログラム 〒305-8577 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 〒305-8577 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 〒305-8575 茨城県つくば市天王台1丁目1-1

E-mail: [†]tomiya@kde.cs.tsukuba.ac.jp, ^{††}horie@cs.tsukuba.ac.jp, ^{†††}kitagawa@cs.tsukuba.ac.jp

あらまし 睡眠臨床における必須検査の1つである睡眠ステージ判定を自動的に行う深層学習モデルが複数提案されている。これらは医師・技師とほぼ同等の判定精度を有しているが、生体信号に含まれるノイズや個人差成分によって波形が変化すると、判定精度が低下する傾向にある。本研究では、スタイル変換技術をベースに、生体信号に含まれるノイズや個人差成分を除去するとともにステージ判定に必要な波形を強調するような深層学習モデルの開発に取り組んだ。また、本提案手法が出力した波形を目視で確認するとともに、既存の睡眠ステージ判定モデルに適用した場合の有効性について検討した。

キーワード 生体信号処理, 医療・ヘルスケア, オートエンコーダ, 睡眠ステージ判定, ノイズ・個人差除去

1 序 論

睡眠ステージとは、睡眠の状態を表す名義尺度であり、覚醒(W)レム睡眠(REM)と3種類のノンレム睡眠(N1~N3)の5つで構成される。一晩の睡眠に各ステージがどの程度含まれるか、どのように遷移するかは睡眠の質や睡眠障害の有無を判断するための重要な基準の一つとなっている。脳波を含む生体信号から睡眠ステージを決定する「睡眠ステージ判定」は、睡眠臨床における重要な検査の一つであるものの、現在は生体信号の目視で行われているため、多大な時間と労力を必要としていた。

これらの負担を軽減すべく、近年では深層学習モデルによる自動化が試みられている。例えば Perslev らが提案した U-Time [1] は、睡眠ステージ判定のための時系列セグメンテーションを行う深層学習モデルである。その F1 スコアは 0.79 であり、ヨーロッパの 8 つの睡眠研究所における技師間一致率 0.68 [2] と比べても、十分に実用に足る性能だといえる。

しかしながら、一般的な自動睡眠ステージ判定手法では、ノイズや個人差成分に由来する生体信号の波形変化に弱いという問題を抱えている。例えば、U-time の場合判定精度は 60% 程度に落ちてしまう場合があった。ここで個人差成分とは、被験者の年齢に由来する特定の周波数成分の欠落等を指す。ノイズや個人差を含んだ信号に対して自動睡眠ステージ判定が適用できない場合、人手によって再度判定を行う必要があるため、上記のような自動判定手法の実用性が損なわれてしまう。

本課題を解決する方法としては (1) モデルの生体信号の変化に対する頑健性を高める (2) 事前に生体信号からノイズ・個人差を取り除く、の 2 つが考えられる。本研究では、教師信号作成の手間が少なく、手動のステージ判定にも応用が利く (2)

の方法に着目した。時系列信号から特定のノイズを削除する深層学習モデルは既に存在しているが、これらは理想的な入出力のペアを学習サンプルとして必要とする。ペアを必要としないモデルも提案されているものの、対応できる SN 比の制限が厳しかった。そこで本研究では、ノイズや個人差を特定して除去するのではなく、ステージ判定に必要な波形(特徴波)のみ残し、それ以外の成分を除くアプローチを考えた。

具体的には Gatys らの画像スタイル変換 [3] から着想を得た。スタイル変換とは、画像に映るオブジェクトの概形、配置といった重要な情報である「コンテンツ」をそのままに、色使い、質感等「スタイル」を別のそれに変化させる手法である。本研究におけるスタイル変換は、睡眠ステージ判定に必要な特徴波を「コンテンツ」、ノイズや個人差成分を「スタイル」と見做し、そのスタイルを 0 に近づけることを目的とした。

2 特徴波

アメリカ睡眠医学会(AASM)のマニュアルによると、臨床現場での睡眠ステージの判定は脳波、顎筋電位、眼電図の生体信号が必要となる [4] [5]。これらを用いて、医師や臨床検査技師が 1 エポック (30 秒) 単位で睡眠ステージのラベル付けを行う。その判定根拠となる特徴的な波形を特徴波と呼び、10 数種類の特徴波が定義されている。複数の生体信号を用いた自動睡眠ステージ判定手法が提案されてきたが、判定の計算量、拡張性を考慮した脳波のみの判定手法も多く提案されている。そのため、本研究では脳波単チャンネルからの睡眠ステージ判定に着目し、取り扱う特徴波を表 1 に示す。

行う。

4.2 信号再現部

信号再現部では潜在表現ベクトルの「特徴波以外に関連する潜在表現ベクトル」を0に近づけてデコードする。後述するステージ判定部によって潜在ベクトルの重要度評価を行い、重要度の低い潜在表現ベクトルを「特徴波以外に関連する潜在表現ベクトル」と見做し、それらを0に近づけて信号の再構成を行う。この操作によって特徴波が残り、それ以外の成分が取り除かれた信号が出力されることを期待する。

4.3 ステージ判定部

ステージ判定部では潜在表現ベクトルの一部を用いて睡眠ステージ判定を行い、潜在ベクトルの重要性評価＝「特徴波のベクトル」と「特徴波以外に関連するベクトル」を明らかにする。モデルの学習時は、このモジュールに入力している潜在表現ベクトルが特徴波の情報を持つように訓練されつつ、入力しない潜在表現ベクトルにはノイズ・個人差成分の情報が残ることを期待する。また、モデルの予測時（信号再現時）は、重要度の低い潜在表現ベクトルを取り除く役割を担う。

重要度の求め方は、デクリメンタルな特徴選択を行い、ステージ判定部の判定精度が大きく下がった潜在表現ベクトルほど特徴波情報を有していると仮定し、重要度の高いものと見做す。重要度の高い潜在表現ベクトルだけを信号再現に使用することで、特徴波のみを残した再現信号が得られることを期待する。

4.4 誤差関数

提案モデルは (1) 入力信号の大まかな波形を保持すること (2) 特徴波情報を抜かりなく保持すること (3) 潜在空間上で特徴波とそれ以外を明確に分離すること、の3つの要件を満たす必要がある。そのため、学習に使用する誤差関数は (1) 信号再現誤差 (2) 睡眠ステージ判定誤差 (3) 潜在表現ベクトル誤差、の3つの誤差の線形和を誤差関数として学習に使用した (式 (1))。ここで α, β はそれぞれ実験的に 10, 5 という値に設定した。

$$\begin{aligned} \text{Error} = & L_{\text{rec}}(x_{\text{in}}, x_{\text{out}}) \\ & + \alpha \cdot L_{\text{score}}(y, \hat{y}) + \beta \cdot L_{\text{latent}}(x_{\text{in}}) \end{aligned} \quad (1)$$

4.4.1 信号再現誤差

再構成の過程で、特徴波を含めた入力信号の大まかな波形を保持するため、式 (2) に示す RMSE を使用した。

$$L_{\text{rec}}(x_{\text{in}}, x_{\text{out}}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{\text{in}}^{(i)} - x_{\text{out}}^{(i)})^2} \quad (2)$$

4.4.2 睡眠ステージ判定誤差

潜在表現ベクトルの一部を使って睡眠ステージ判定を行い、それらに特徴波情報を集めるための Categorical cross entropy を式 (3) に示す。

$$L_{\text{score}}(y, \hat{y}) = - \sum_{i=1}^n y^{(i)} \log \hat{y}^{(i)} \quad (3)$$

4.4.3 潜在表現ベクトル誤差

潜在表現ベクトルから信号再現するとき、提案モデルは一つの特徴波を複数の潜在表現ベクトルで表現するケースが見られた。この状況では、微かな特徴波情報がのった潜在表現ベクトルが取り除かれることによって、再構成された特徴波の波形が崩れる恐れがある。そのため、この現象は提案手法が目的を達成するうえで不適切であり、潜在表現ベクトル同士が似た情報を持たないように新たな誤差を定義する必要があると考えた。ある潜在表現ベクトルの、それ以外のベクトルに対する相関を下げるため、アルゴリズム (1) に示す Bottleneck 誤差を導入する。

Algorithm 1 Latent Loss Algorithm

Require: Latent vectors of input signal L

Input: $L \in \mathcal{R}^{t \times c}$ { t : length of time axis, c : # of channels}

Output: $\text{LatentLoss} \in \mathcal{R}$

- 1: $A, B = \text{Split}(L, s)$ { $A \in \mathcal{R}^{t \times s}$ is for Sleep stage scoring, $B \in \mathcal{R}^{t \times (c-s)}$ is for signal reconstruction}
 - 2: $\text{Loss}_a = \text{Sum}(A^T A) / \text{Sum}(\text{Diag}(A))$
 - 3: $\text{Loss}_b = \text{Sum}(B^T B) / \text{Sum}(\text{Diag}(B))$
 - 4: $\text{LatentLoss} = \text{Loss}_a + \frac{s}{c-s} \cdot \text{Loss}_b$
-

4.5 学習方法

最適化手法は Adam($\eta = 0.001$) [13] を使用し、バッチサイズ 128 で 100 回の学習を行なった。検証用データの誤差が 5 回連続で下がらなかったときに早期終了させ、30 回で学習が完了した。

5 評価実験

実験にはアメリカ国立心肺血液研究所が公開する SHHS1 データセット [14] [15] を使用する。睡眠障害を持たない 40 歳以上の被験者から睡眠中の脳波、筋電位、眼電位等を記録したもので、本実験では脳波信号 (C4-A1 チャンネル、サンプリング周波数 125Hz、ハイパスフィルター 0.15Hz) 約 100 万件を使用する。さらに、前処理として全ての信号をサンプリング周波数 100Hz でリサンプリングした。

5.1 提案モデルが特徴波成分を独立して保持しているか定性評価

エンコーダ部から潜在表現ベクトルを得て、そのチャンネルごとに独立させて信号再現を行った結果を図 2 に示す。提案モデルは入力信号を複数の信号片に分解していることが確認できるが、その分布は偏っている。入力信号中央 (13-14 秒付近) には睡眠ステージ N2 特徴波である K 複合が認められるが、この情報はいくつかの信号片が保持している様に見える。複数の潜在表現ベクトルで信号を再現してしまうと、ステージ判定部によってある信号片が重要でないと思われたときに、その潜在表現ベクトルが取り除かれることで、再現信号が歪んでしまう可能性がある。

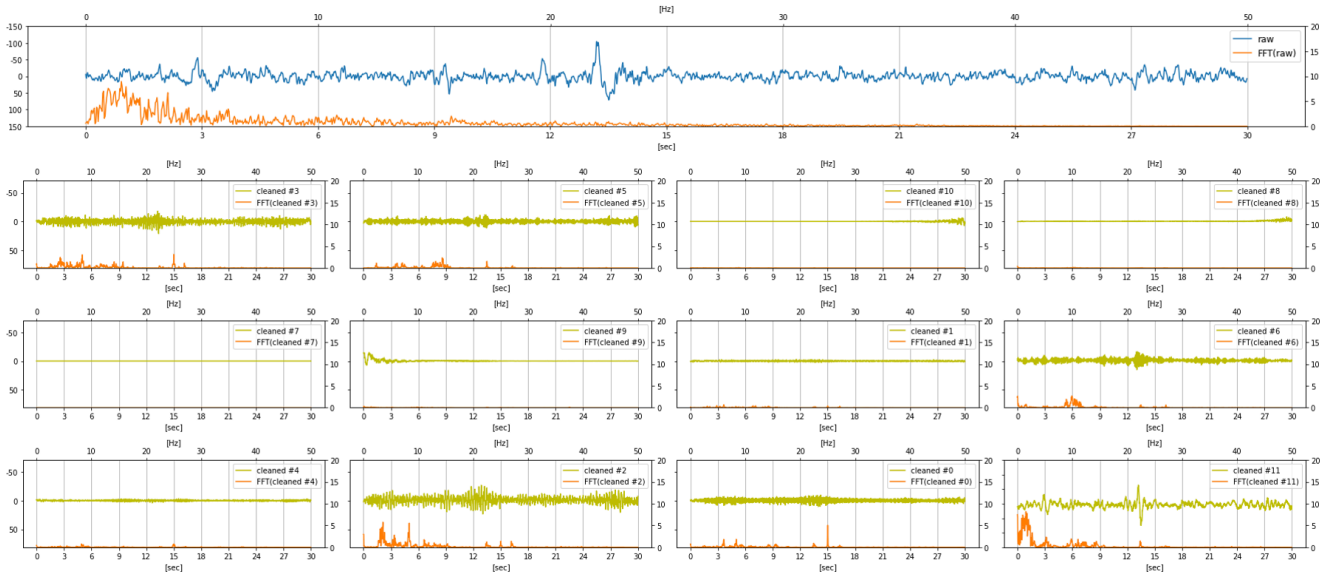


図2 13-14秒付近にK複合が見られる脳波。睡眠ステージN2。上：元の脳波（青）と周波数構成（橙），下：潜在表現ベクトルを一つずつ再現したもの（黄）と周波数構成（橙）。横軸上：周波数 [Hz]，横軸下：時間 [sec]，縦軸左：振幅 [μ V]，縦軸右：FFT 成分の強度

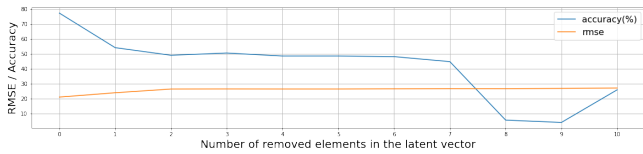


図3 横軸：取り除いた潜在ベクトルの数，縦軸：提案モデルの Accuracy と RMSE の値

5.2 変換した信号を既存モデルに入力して判定精度の評価

ステージ判定部によって重要度が低いと見做された潜在表現ベクトルから順番に取り除いて睡眠ステージ判定を行なった時の Accuracy と、入力信号と再現信号の RMSE の値の遷移を図3に示す。図より8番目までの潜在表現ベクトルを取り除いたときに Accuracy が大きく下がることから、その時点まで残った潜在表現ベクトルは睡眠ステージ判定に必要な情報を有していると仮定した。7番目の潜在表現ベクトルを取り除いた時点での再現信号を図4に示す。さらに、再現信号を用いて U-Time [1] で判定精度がどう変化するかを調査した。

元信号を U-Time に入力して睡眠ステージ判定を行なったときの混同行列を表2に、提案モデルが変換した信号を入力したときの混同行列を表3に示す。変換前後それぞれの判定結果の Accuracy は 0.91, 0.83 であった。この結果から、変換によって一部の特徴波情報が抜け落ちてしまったことが考えられる。

6 考察

再構成した信号は、特徴量である K 複合の形状がある程度維持されつつも、その他の部分の振幅が明確に減少しており、「ノイズ除去・個人差成分の除去」について一定の成果があったと思われる。しかし、取り除いた潜在表現ベクトルの多くは信号再現に大きく寄与しないものがほとんどであったため、潜在表

表2 元信号を入力した U-Time の混同行列

		予測ラベル				
		W	N1	N2	N3	R
正解ラベル	W	221	3	4	0	18
	N1	6	28	5	0	22
	N2	2	1	536	17	12
	N3	0	0	6	24	0
	R	0	0	3	0	161

表3 再構成脳波を入力した U-Time の混同行列

		予測ラベル				
		W	N1	N2	N3	R
正解ラベル	W	229	4	2	0	11
	N1	10	16	8	0	27
	N2	15	15	518	2	18
	N3	0	0	26	4	0
	R	33	1	5	0	125

現ベクトル獲得の過程が改善されれば結果はより良いものになるだろう。

7 結論

スタイル変換に基づくノイズ・個人差を除去・軽減する信号変換モデルを開発した。潜在表現ベクトルを独立して再現する実験から、提案モデルは一つの特徴波を複数の信号片で表現してしまっていることを確認した。そのため、適当な潜在表現ベクトルを取り除くだけでは、睡眠ステージ判定に不要な波形成分を取り除くことができないことがわかった。

今後の課題として、誤差関数を再調整する等、潜在表現ベクトル同士が似通わないような工夫を取り入れることで、特徴波成分とノイズ・個人差成分を明確に分離するモデルの開発に取

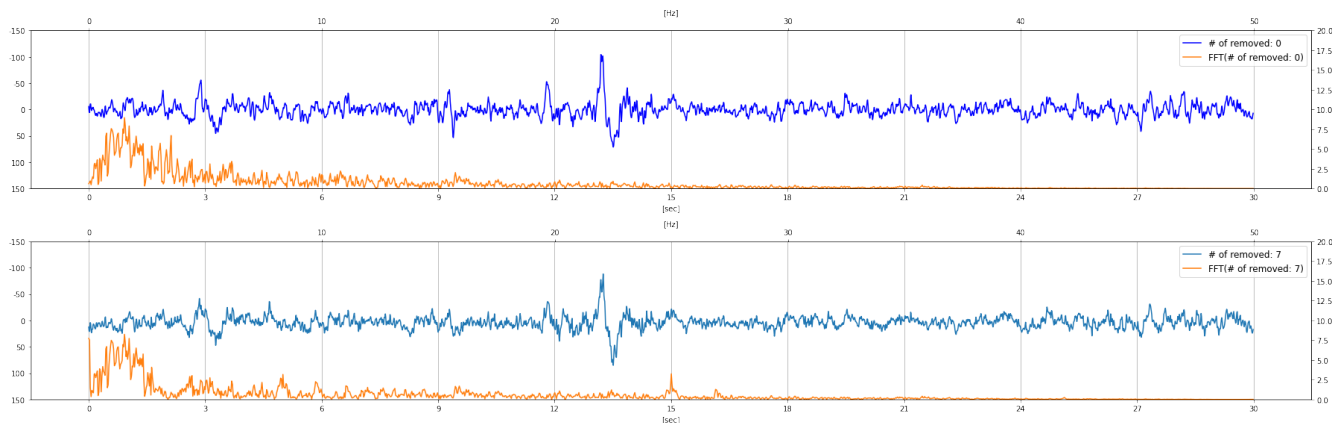


図4 元信号(上)と7つの低重要度の潜在表現ベクトルを取り除いた時点での再現信号。(下)
 横軸: 時間 [sec], 縦軸左: 振幅 [μ V], 縦軸右: FFT 成分の強度

り組む。

8 謝 辞

本研究の一部は, JST ムーンショット (JP21zf0127005), JSPS 科研費 (JP22K19802, 19156917, 19K20334), (株) S'UIMIN 共同研究による。

文 献

- [1] Perslev, Mathias, et al. "U-Time: A fully convolutional network for time series segmentation applied to sleep staging." arXiv preprint arXiv:1910.11162 (2019).
- [2] Danker-Hopfe H, Kunz D, Gruber G, Klösch G, Lorenzo JL, Himanen SL, Kemp B, Penzel T, Röschke J, Dorn H, Schlögl A, Trenker E, Dorffner G. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J Sleep Res.* 2004 Mar;13(1):63-9. doi: 10.1046/j.1365-2869.2003.00375.x. PMID: 14996037.
- [3] Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2016.
- [4] aC, Ancoil-Israel S, Chesson AL et al. "The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications." *American Academy of Sleep Medicine, Westchester, IL, 2007.*
- [5] 日本睡眠学会, 改訂版 臨床睡眠検査マニュアル, (株) ライフサイエンス, 2016.
- [6] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention.* Springer, Cham, 2015.
- [7] Jadhav, Pankaj, et al. "Automatic sleep stage classification using time - frequency images of CWT and transfer learning using convolution neural network." *Biocybernetics and Biomedical Engineering* 40.1 (2020): 494-504.
- [8] Jiang, Dihong, et al. "Robust sleep stage classification with single-channel EEG signals using multimodal decomposition and HMM-based refinement." *Expert Systems with Applications* 121 (2019): 188-203.
- [9] Verma, Prateek, and Julius O. Smith. "Neural style transfer for audio spectrograms." arXiv preprint arXiv:1801.01589 (2018).
- [10] Grinstein, Eric, et al. "Audio style transfer." 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2018.
- [11] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Communications of the ACM* 60.6 (2017): 84-90.
- [12] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
- [13] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).
- [14] Zhang GQ, Cui L, Mueller R, Tao S, Kim M, Rueschman M, Mariani S, Mobley D, Redline S. The National Sleep Research Resource: towards a sleep data commons. *J Am Med Inform Assoc.* 2018 Oct 1;25(10):1351-1358. doi: 10.1093/jamia/ocy064. PMID: 29860441; PMCID: PMC6188513.
- [15] Quan SF, Howard BV, Iber C, Kiley JP, Nieto FJ, O'Connor GT, Rapoport DM, Redline S, Robbins J, Samet JM, Wahl PW. The Sleep Heart Health Study: design, rationale, and methods. *Sleep.* 1997 Dec;20(12):1077-85. PMID: 9493915.