

웨어러블 디바이스에서의 3축 가속도 센서를 이용한 고 연령층 수면-각성 감지 알고리즘 개발

박범준¹ · 남보름² · 이종실³ · 김인영^{4*}

Development of A Sleep-Wake Detection Algorithm for The Elderly Using 3-Axis Accelerometer in Wearable Device

Beomjun Bark¹ · Borum Nam² · Jongshill Lee³ · In Young Kim^{4*}

¹Graduate Student, Department of Biomedical Engineering, Hanyang University, Seoul, 04763 Korea

²Graduate Student, Department of Electronic Engineering, Hanyang University, Seoul, 04763 Korea

³PhD, Department of Biomedical Engineering, Hanyang University, Seoul, 04763 Korea

^{4*}Professor, MD, PhD, Department of Biomedical Engineering, Hanyang University, Seoul, 04763 Korea

요 약

수면은 인간에게 필수적인 요소로, 수면다원검사의 수면단계분류를 통해 평가된다. 그러나 이 수면다원검사는 First night effect와 같은 한계점을 가지고 있기 때문에, 웨어러블 디바이스를 기반으로 하는 수면단계분류 알고리즘이 연구되어왔다. 하지만, 이 알고리즘들은 고 연령층에서의 성능이 떨어진다는 한계점이 존재한다. 본 연구에서는 평균 나이 72.6세의 고 연령층을 대상으로 한 웨어러블 디바이스의 가속도 센서를 활용하여 수면 중 각성 감지 알고리즘을 설계하였다. 결과적으로, 고 연령층의 수면 특성을 고려한 특징점을 딥러닝 네트워크에 적용 하였을 때 성능 향상을 얻었으며, 수면의 질을 판별 할 수 있는 4가지 파라미터의 추정 중 3가지 파라미터에서 유의한 결과를 얻을 수 있었다. 본 연구는 고 연령층에 적합한 수면 중 각성 감지 네트워크를 처음으로 제시한 연구로, 간편한 수면 모니터링과 개인 맞춤형 수면 질 평가에 도움이 될 것으로 기대된다.

ABSTRACT

Sleep is an essential factor for humans, and it is evaluated through sleep staging in polysomnography. However, polysomnography has limitations such as the first night effect. Therefore, sleep staging algorithms based on wearable devices have been studied. However, these algorithms have limitations in performance for the elderly. In this study, Sleep-wake detection algorithm was designed using the accelerometer sensor of a wearable device, targeting the elderly with an average age of 72.6 years. As a result, when applying features that consider the sleep characteristics of the elderly to the deep learning network, better performance was achieved. Three of the four estimated parameters to assess sleep quality were significant. This study is the first to propose a sleep-wake detection network specifically designed for the elderly, and it is expected to contribute to convenient sleep monitoring and personalized sleep quality assessment.

키워드 : 수면 질 평가, 웨어러블 디바이스, 딥러닝, 시계열 데이터 분석, 가속도 센서

Keywords : Sleep quality assessment, Wearable device, Deep learning, Time series data analysis, Accelerometer sensor

Received 2 August 2023, Revised 17 August 2023, Accepted 4 September 2023

* Corresponding Author In Young Kim(E-mail:iykim@hanyang.ac.kr, Tel:+82-2-220-0690)
Professor, Department of Biomedical Engineering, Hanyang University, Seoul, 04763 Korea

Open Access <http://doi.org/10.6109/jkiice.2023.27.9.1060>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

1.1 수면 단계와 수면 모니터링의 중요성

수면은 인간에게 필수적인 요소 중 하나로, 신체와 정신의 회복을 도와주는 핵심적인 행위이다. 수면은 신경계를 휴식 상태로 전환하며, 인체에 필요한 단백질의 생성과 합성을 더 활발하게 할 뿐만 아니라, 스트레스 및 우울증을 완화시킴으로써 인간의 안정과 행복을 유지하는데 필수적인 역할을 한다[1, 2]. 이러한 수면의 역할 때문에 수면의 질에 대한 중요성과 관심도는 지속적으로 증가하고 있다. 이에 따라 수면다원검사의 건수도 증가하고 있는 추세이다[3]. 수면다원검사란, 수면 질 및 장애를 평가하고 진단하기 위해 사용되는 종합적인 검사 방법이다[4]. 검사실에서 하루 수면을 취하면서, 다양한 생체신호(뇌파, 심전도, 안구운동, 근전도, 움직임 등)와 수면 관련 설문지를 취득한다. 이를 통해 과학적 근거를 바탕으로 다양한 수면 장애와 정확한 진단을 통한 개인 맞춤형 치료가 가능하다.

수면다원검사를 통한 수면 질 파악 방법으로는 수면 단계분류가 있다. 수면단계는 뇌전도, 안구전도, 근전도를 기반으로 Rechtschaffen and Kales (R&K)나 American Academy of Sleep Medicine (AASM)의 매뉴얼에 따라 현재 수면의 깊이를 측정한다[5]. AASM을 기준 시, Wake, REM, N1, N2, N3(NREM: N1, N2, N3) 총 5개의 단계로 구성되며, Wake에서 N3로 갈수록 수면을 깊이 취하고 있다고 해석 된다. 수면다원검사는 수면을 모니터링 할 수 있는 가장 효과적인 도구로서 알려져 있다.

그러나 수면다원검사의 한계점도 존재한다. 수면다원검사의 경우 1~3일 가량의 짧은 기간 동안만 측정하기 때문에 평상시 수면의 상태를 판단하기 어렵다[6].

또한 다양한 생체신호를 측정해야하기 때문에, 많은 설비와 비용이 요구되며, 수면의 장애, 수면점수, 상태 등을 기록하기 위해서는 전문가가 직접 수행을 해야 한다[7]. 이러한 이유 때문에 검사자의 수면이 평소와 다르게 영향을 받을 수 있다는 한계점을 가지고 있다[8]. 이에 따라 수면다원검사의 단점을 보완하거나, 검사의 필요성을 알려주는 간편한 수면 모니터링 시스템 개발이 요구되어 왔다[9, 10].

1.2 수면 모니터링 관련 사전 연구

먼저, 다양한 수면 모니터링 시스템 개발 관련 연구 중, 웨어러블 디바이스의 가속도 센서를 기반으로 한 수면단계분류 연구가 활발하게 이루어지고 있다. 특히, [11]은 다양한 인공지능 기법을 이용하여 수면 중 발생하는 각성을 감지하는 가속도 센서 기반 알고리즘의 성능을 평가하였다. 사용한 인공지능 기법으로는 Multi-layer perceptron을 포함하여 CNN, RNN, LSTM 뿐만 아니라, Logistic regression 등 다양한 머신러닝 기법 또한 활용하였다. 성능은 정확도 기준으로 66.07%에서 최고 92.86%의 성능을 가졌으며 CNN에서 가장 좋은 성능을 보였다. 또 다른 관련 연구[12]에서는, CNN-LSTM을 결합한 모델 구조에 심박변이도 데이터를 특징점으로 활용하여 학습 보조 역할을 수행하는 구조를 설계하였다. 이를 통해 수면 중에 효과적으로 각성을 감지하는 수면단계분류를 하고자 하였다. 해당 연구는 정확도와 Geometric mean을 제시하였으며, 각각 95.8%와 0.896의 결과 값을 획득하였다. 더 나아가 해당 연구와 비슷한 구조에 Attention layer를 추가함으로써 학습된 모델이 어떤 부분에 집중하는지 여부를 확인하기 위한 Explainable-AI 모델 또한 연구되었다[13]. 추가적으로 가속도 센서에 다양한 생체 신호(심전도, 광용

Table. 1 Sleep-wake detection algorithm based on wearable device using 3 axis-accelerometers (Acc: Accelerometers, PPG: Photoplethysmography)

Ref	Measure	Method	Subjects	Result
[17]	Actiwatch-L	Acc: Wake-Sleep	21 subjects: 38-64 age	Accuracy: 79.8% Cohen's kappa: 0.67
[18]	Motion logger	Acc: Wake-Sleep	40 subjects: 18-40 age	Accuracy: 85.4% Cohen's kappa: 0.54
[19]	Zulu Watch	Acc: Wake-Sleep	8 subjects: 18-35 age	Accuracy: 90.5%
[20]	Fitbit Charge 2	Acc+PPG: W-REM-Light-Deep	35 subjects: 14-40	Accuracy: 65-89% Cohen's kappa: 0.54

적맥파)를 추가하여 더 구체적인 수면단계분류 연구도 수행되었다[14-16]. 이렇듯 웨어러블 디바이스를 수면 단계분류 연구는 지속적으로 수행되고 있으나, 이전까지의 연구들은 한계점이 존재한다. 표 1은 상용 웨어러블 디바이스 내 수면 중 각성 감지 알고리즘에 대한 성능을 평가한 표이다. 해당 표에 따르면, 수행된 연구 대부분이 고 연령층을 제외한 청년과 중장년층의 연령대로 구성된 연구대상자만이 주로 사용되었다는 것을 확인할 수 있다. [21]에 따르면, Mi band2를 이용하여 젊은 성인 10명, 10명의 건강한 노인, 8명의 치매 노인에 대해서 수면 중 각성 감지 알고리즘에 대한 성능을 평가했을 때, 건강한 노인 3명과 8명에 해당하는 치매 노인 모두의 수면을 정확하게 감지하지 못했음을 확인했다. 이러한 원인으로서는 고 연령층 인구는 젊은 성인에 비해 수면 중 움직임이 많고 치매환자는 움직임이 더 많기 때문이라고 해석하였다. 따라서, 현재까지 개발된 알고리즘은 고 연령층을 제외한 연령층에 대해서만 검증을 수행하였으며, 고 연령층의 수면 특성을 충분히 고려한 알고리즘은 부족하다는 사실을 확인할 수 있다.

이에 따라, 본 연구에서는 고 연령층의 수면 특성을 고려한 수면 중 각성 감지 알고리즘을 통해 추정한 수면 단계를 바탕으로 수면의 질을 추정할 수 있는 인공지능 모델을 개발하고자 한다.

II. 데이터 수집 및 딥러닝 모델 설계

2.1 데이터 수집

본 연구에서 사용된 데이터의 수집은 강동성심병원에서 수행하였으며, 수면다원검사를 수행하는 동시에 상용 웨어러블 디바이스를 이용해 3축 가속도 센서 값을 수집하였다. 사용한 상용 디바이스는 ActiGraph wGT3X-BT (그림1)로 100Hz에 해당하는 3축 가속도 센서 값을 최대 25일간 측정 할 수 있는 디바이스이다 [22]. 우리는 해당 장비를 손목에 착용 한 후 하룻밤 동안 수면을 취하게 했다. 그리고 수면다원검사로부터 측정된 뇌파, 안구전도, 근전도를 기반으로 현재 수면 단계를 전문가가 AASM 매뉴얼(Wake, REM, N1, N2, N3)에 따라 30초에 한 단계로 Annotation 하였다.

본 연구에서는 평균 연령 72.6세 (67-79세)에 해당하는 고 연령층 21명에 대해서 데이터를 취득하였다. 해당

검사 인원은 남성 12명 여성 9명으로 이루어져있으며, 수면시간은 평균 6.7시간으로 나타났다. 타 연구[23]에서 19세에서 30세를 대상으로 취득한 수면 파라미터와 비교한 결과는 표 2와 같다. 비교한 파라미터는 TST(total sleep time), SOL(sleep onset latency), SE(sleep efficiency), WASO(wake after sleep onset)이다. 취득한 데이터에서 타 데이터와 비교되는 고 연령층의 수면 특징이 잘 반영되었음을 확인하였다.

Table. 2 Comparison of sleep parameters for the elderly and non-elderly

Sleep parameter	19-30 year[23]	This study
TST (minutes)	414.9 ± 43.4	402 ± 48.4
SOL (minutes)	25.3 ± 18.9	29.3 ± 13.7
SE (%)	86.4 ± 8.9	74.2 ± 12.4
WASO (minutes)	38.2 ± 29.2	103.2 ± 50.9

모든 데이터 수집은 2021년 9월 28일부터 12월 07일까지 수집하였다. 모든 검사 인원은 연구에 참여하는데 동의하였으며, 충분한 설명을 듣고, 임상연구 대상자 동의서와 개인정보 이용 동의서를 작성하였다. 본 연구는 강동성심병원 임상연구심의위원회의 승인 (IRB: 2021-07-018)을 받아 데이터를 수집하였다.

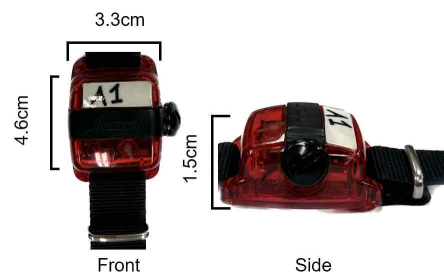


Fig. 1 Wearable device: ActiGraph wGT3X-BT

2.2. 데이터 전처리

본 연구는 모델의 입력 단 데이터를 한 사람에 대해서 812개의 수면단계 (30초 × 812 = 6.7 시간)를 가지도록 데이터의 길이를 조절하였다. 그 후 메모리 절약 및 계산속도 향상을 위해 100Hz로 취득한 신호를 30Hz로 Down-sampling 하였다. 우리는 고 연령층의 잦은 각성

을 효과적으로 감지하기 위해 관련 특징점을 추출하였다. 표3은 추출한 특징점에 대한 항목을 나타내며, 해당 항목들은 Human activity recognition을 수행하기 위해 주로 사용된 특징점을 사용하였다[24-26]. 신호의 변동성을 파악하기 위해 30초 단위로 시간 도메인에서의 3축(X, Y, Z 축)에 해당하는 각각 3가지의 평균, 표준편차와 실효값을 계산하였으며, 주파수 도메인에서의 Fourier transform을 통해 계산된 지배적인 주파수와 평균 전력 스펙트럼 밀도(PSD) 값을 특징점으로 사용했다. 이에 따라 총 15개의 특징점을 추출하였다.

본 연구에서 사용한 데이터는 AASM 매뉴얼에 따라 전문가가 30초에 한 번씩 수면 단계 4단계 (Wake, REM, N1, N2, N3) 중 하나로 annotation 하였다. 본 모델은 수면 중 각성(Wake)을 감지하는 것이 목적이기 때문에 모델의 출력에 해당하는 라벨 값을 0(Wake), 1(수면: REM, N1, N2, N3)로 전처리 하였다.

Table. 3 Domain-specific features in accelerometers

Domain	Feature (15)
Time	Average (3)
	Standard deviation (3)
	Root Mean Square (3)
Frequency	Dominant frequency (3)
	PSD value (3)

2.3.1 딥러닝 모델 구조

한 피험자에 대한 전처리 된 수면데이터와 812개의 각성과 수면으로 이루어진 라벨이 각각 모델의 입력과 출력에 대한 정답으로 사용된다. 기존 시계열 처리에 적합하지 않은 CNN 기반 모델이나 기울기 소실에 의해 긴 시퀀스 데이터 처리에 약점을 보이는 LSTM 모델의

단점을 보완하기 위해 Inception time[27] 기반의 딥러닝 모델 (그림2)을 설계하였으며, 입력단 직후 Attention단을 통해 딥러닝 모델이 어떤 부분에 집중하여 판단을 수행하였는지 시각적 해석을 제공하고자 하였다. Inception time단 이후 Temporal convolutional network (TCN)[28]을 이용하여 시계열을 효과적으로 처리하였다. TCN 단 이후 압축되어진 시계열 데이터는 전처리 과정에서 추출한 특징점 15개와 함께, 1D convolution layers를 통해 데이터를 압축하였다. 출력은 812개의 값에 대해서 시그모이드 활성화 함수를 통한 (812,1) 형태를 획득했다. 최종적으로 812의 길이를 가진 시계열 데이터를 출력하며 라벨과의 차이를 비교하여 성능을 평가하였다.

2.3.2 Inception time과 TCN 네트워크

시계열 데이터 분석에 사용되는 Inception time은 Inception 모듈과 잔차연결로 구성되어 있다. Inception 모듈은 bottleneck 구조 이후 3가지의 서로 다른 크기의 커널을 통해 서로 다른 스케일의 시계열 패턴을 포착한다. 잔차연결은 3개의 Inception block마다 적용되며, 학

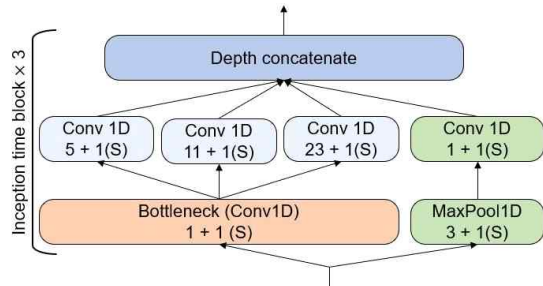


Fig. 3 Structure of Inception time block (S: stride)

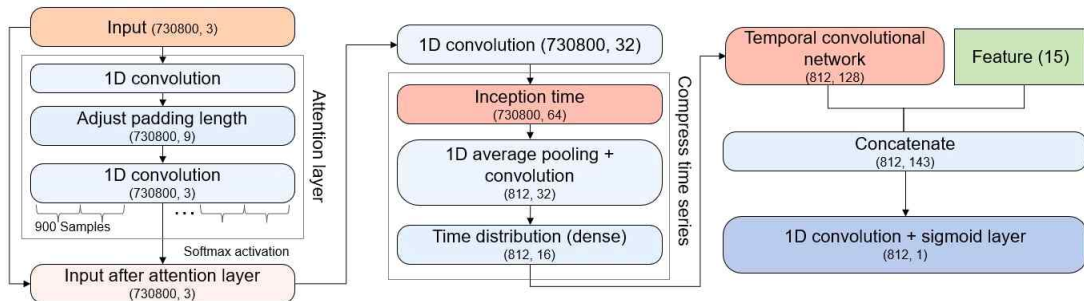


Fig. 2 Deep learning model structure based on inception time and temporal convolutional network

습된 기울기의 흐름을 개선한다. 해당 연구에서 우리는 경험적으로 5, 11, 23의 커널 크기를 채택하여 시계열을 학습하였으며, Inception block 3층을 사용했다. 그림 3은 본 연구에서 사용한 Inception time에 대한 구조 및 계수를 나타낸다.

Inception time와 Time distribution layer에 의해 압축된 특징은 TCN을 이용해 학습했다. TCN은 Dilation convolution 기반의 연산을 통해 동일 계산량으로 더 넓은 수용장을 보는 딥러닝 모델이다. 본 연구에서는 커널 크기 4와 Dilation 계수를 1, 2, 4, 8, 16, 32로 설정한 총 6층의 모델을 설계하였다.

2.2.3 최적화

본 연구에서 설계한 딥러닝 모델은 시그모이드 레이어드의 결과에 대한 값을 기준으로 최적화를 수행한다. 812,1의 시계열 형태를 가진 출력에 대해서 가장 이진 교차 엔트로피 손실 함수를 통해 812번의 누적된 손실 값을 계산하였다. 계산된 손실을 최소화 하는 방향으로 Adam optimization을 적용하였다.

2.2.4 딥러닝 네트워크 파라미터 설정

모델 학습 시 Grid search 기법을 기반으로 최적의 파라미터를 적용하였다. 21명에 대해서 7 fold를 통해 19명을 학습 데이터로, 3명을 검증 데이터로 활용하였으며, 100번의 Epoch을 통해 학습하였다. 학습률은 0.001을 시작으로 장기간 손실이 감소하지 않을 때마다 절반씩 감소시켰다. 해당 데이터의 경우 각성에 비해 수면이 많으므로 각 라벨별 데이터 비율을 기반으로 가중치를 차별화하여 주는 손실을 적용하였다. 배치사이즈는 2로 설정하였다. 이 모델은 Pytorch 프레임워크를 기반으로 설계하였으며, 사용한 CPU는 AMD Ryzen 5 5600X 6-Core Processor, GPU: NVIDIA GeForce RTX 3090, RAM: 64.0GB를 통해 학습되었다.

2.3. 성능평가 지표

설계한 모델은 아래 지표를 통해 성능이 평가되었다. 아래의 TP는 True Positive, TN은 True Negative, FN은 False Negative, FP는 False Positive를 의미한다. 또한, $Pr(a)$ 는 2명의 평가자들의 평가 결과가 일치할 확률이며, $Pr(e)$ 는 2명의 평가자들의 결과가 우연히 일치할 확률이다.

$$\text{Accuracy: } (TP + TN) / (TP + FP + FN + TN) \quad (1)$$

$$\text{Precision: } TP / (TP + FP) \quad (2)$$

$$\text{Recall: } TP / (TP + FN) \quad (3)$$

$$\text{F1-score: } 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

$$\text{Cohen's kappa: } \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (5)$$

위 평가 지표를 바탕으로 전반적인 모델의 성능을 확인함과 동시에, 각성의 F1-score를 통해 수면 상황에서 각성 감지능력을 평가하고자 하였다. 뿐만 아니라 Cohen's kappa 계수를 이용하여 실제 수면다원검사를 실시한 전문가와의 의견 일치도를 확인하여 성능 평가를 하고자 하였다.

III. 평가 및 토의

3.1 Fold별 분류 성능 결과

설계한 모델은 7 fold validation을 통해 평가를 수행하였다. 21명의 검사 인원 중 3명에서의 2,436개의 샘플 (812×3)을 검증을 위해 사용되었다. 표 4는 본 연구에서 사용한 딥러닝 모델의 각 Fold 별 Accuracy, 각성에 대한 F1 score 그리고 전문가와의 일치도를 확인하는 Cohen's kappa 계수의 성능을 평가한 결과이다. 평균적으로 Accuracy는 $82.66 \pm 3.63\%$, Wake F1 score는 0.66 ± 0.04 , Cohen's kappa는 0.54 ± 0.06 의 결과가 나오는 것을 확인하였다. Cohen's kappa 계수의 일치수준[29]을 확인했을 때, 모든 Fold에서 보통(0.41-0.60)에서 상당한(0.61-0.80)한 일치도 점수를 가지고 있는 것을 확인했다.

Table. 4 Classification performance with 7-fold validation

Metric Fold	Accuracy	Wake F1-score	Cohen's kappa
Fold 1	80.18 %	0.72	0.57
Fold 2	81.40 %	0.63	0.51
Fold 3	78.90%	0.59	0.45
Fold 4	83.04%	0.60	0.49
Fold 5	78.65%	0.65	0.51
Fold 6	88.75%	0.67	0.61
Fold 7	86.95%	0.69	0.61
Average	82.66%	0.66	0.54

3.2 제거 실험

본 연구에서는 각성을 효과적으로 감지하기 위해 5가지의 각성 관련 특징점과 한 사람의 하룻밤 전체에 해당하는 수면 데이터를 입력으로 사용했다. 각 특징점이 분류 성능에 미치는 중요도를 평가하기 위해 특징점에 대한 ablation test(제거 실험)를 진행하였다. 7 fold validation을 이용해 검증하였으며, 7개 Fold의 평균에 대해서 표5에 정리하였다. 해당 표에서 Acc는 Accelerometer signal(가속도 신호), Avg는 Average(평균), Std는 Standard deviation(표준편차), RMS는 Root mean square(실효값), Df는 Dominant frequency(지배적 주파수), Pw는 Power(주파수 대역 대 평균 파워)로 정의했다. 시간 영역의 특징점으로는 Avg, Std, RMS이며, 주파수 영역의 특징점은 Df, Pw이다.

Table. 5 Ablation test: Feature importance

Metric Feature	Accuracy	Wake F1-score	Cohen's kappa
Acc	78.41%	0.50	0.37
Acc+Avg	79.10%	0.56	0.43
Acc+Std	80.06%	0.60	0.47
Acc+RMS	78.88%	0.57	0.43
Acc+Df	78.85%	0.56	0.42
Acc+Pw	77.33%	0.54	0.39
Acc+Avg+ Std+RMS	80.49%	0.63	0.50
Acc+Df+Pw	78.93%	0.59	0.45
Acc+Avg+ Std+RMS+Df+Pw	82.66%	0.66	0.54

위 결과를 바탕으로 Standard deviation(표준편차)이 세 가지 성능지표에서 모두 가장 좋은 성능을 보이는 동시에, 성능에 가장 큰 영향을 미친 것을 확인 할 수 있었으며, Power(주파수 대역대의 평균 파워)가 성능에 가장 적은 영향을 미친 것을 확인했다. 뿐만 아니라, 시간 영역 관련 특징점이 주파수 영역 관련 특징점에 비해 상대적으로 모델의 성능에 영향을 많이 끼치는 것 또한 확인 할 수 있었다. 특징점을 사용하지 않고 가속도 신호만을 이용해서 학습했을 때와 모든 특징점을 사용하여 학습했을 때를 비교하였을 때, Accuracy, 각성 F1-score와 Cohen's kappa가 각각 4.3%, 0.16, 0.17 성능이 향상된 것을 확인 할 수 있었다. 결과적으로 본 연구에서 네

트워크 개발에 사용한 특징점이 고 연령층과 같이 수면과 각성의 변화가 빈번한 대상에 효과적인 것을 확인할 수 있었다.

3.3 딥러닝 네트워크 유의성 검증

본 연구에서는 시계열 데이터를 효과적으로 처리하기 위해 Inception time과 TCN의 조합을 이용하였다. 위 구조가 효과적인지 확인하기 위해 두 네트워크 대신 이전 연구에서 사용된 LSTM, CNN-LSTM과 1D CNN을 통해 취득한 데이터에 적용하는 Ablation test를 수행했다. 7 fold validation을 수행하였으며, 특징점을 모두 사용하여 결과를 확인하였다. 표 6의 결과를 바탕으로, 본 연구에서 구현된 네트워크가 기존 다른 네트워크에 비해 제시한 세 가지 지표 모두에서 성능 향상이 보인 것을 확인 할 수 있었으며, 시계열을 더 효과적으로 처리한 것을 검증하였다.

Table. 6 Ablation test: Classification performance when using 1D CNN, LSTM, CNN-LSTM instead of Inception time and TCN

Metric Model	Accuracy	Wake F1-score	Cohen's kappa
1D CNN	72.37%	0.49	0.30
LSTM	74.88%	0.57	0.39
CNN-LSTM	75.99%	0.57	0.40
This study	82.66%	0.66	0.54

3.4 딥러닝 모델 기반 수면 파라미터 추정

본 연구에서는 수면 중 각성을 감지 할 수 있는 분류 모델을 개발했다. 실 적용 시 수면 중 각성 감지 결과를 기반으로 수면 파라미터를 추출하여 수면의 질을 평가

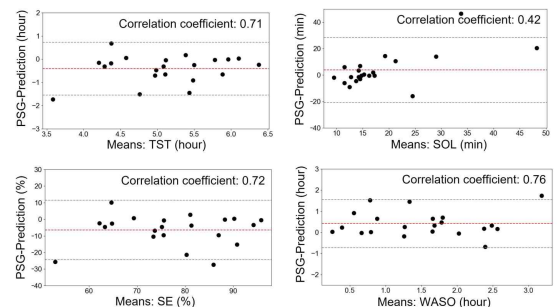


Fig. 4 Bland-Altman plot between polysomnography-derived sleep parameters and estimated sleep parameters

할 수 있다. 이에 따라 이 단락에서는 해당 모델을 기반으로 수면의 질을 계산하였다. 추정한 파라미터는 총 4가지로, TST, SOL, SE, WASO를 평가했다. 전문가의 소견에 해당하는 수면다원검사로 추정한 4가지 수면 파라미터 결과와 본 연구에서 개발한 모델이 추정한 수면 파라미터의 차이를 Bland-Altman 그래프를 통해 확인했다. 테스트인 3명의 피험자에 대해서 7 fold를 적용하여 전체 피험자 21명에 대한 수면 파라미터 추정 결과를 확인하였으며, 추정한 결과와 전문가가 측정한 값과의 상관관계를 분석하였다. 해당 결과는 그림 4와 같다.

그림 4를 통해, SOL을 제외한 나머지 수면 파라미터에서 0.7 이상의 유의한 상관관계를 보이는 것을 확인할 수 있었으며, 본 연구에서 개발한 모델을 통해 수면 질 평가 수단으로서 활용 할 수 있을 것으로 판단된다.

3.5 Attention 기반 시각화

설계한 딥러닝 모델의 입력 직후 1D convolution과 소프트맥스 활성화 함수로 구성되어 있는 Attention 단을 추가하여 시각적, 해석적으로 접근 할 수 있도록 하였다. Attention 값은 0에서 1까지 나타나며 0 일수록 낮은 집중, 1에 가까울수록 높은 집중을 Attention heat map으로 표현하였다. 이에 대한 결과는 그림 5에 나타났다. 수면 중 기상이 발생하게 되면 신체 활성화에 의해 많은 경우에서 가속도 센서 신호의 변화가 발생한다. 그림 5의 사례는 실제 그러한 변화에 집중하여 각성을 감지한 것을 확인할 수 있었으며, 수면 중인 경우에는 낮은 집중도를 보였다. 이러한 Attention 결과를 통해 모델이 입력의 어떤 부분에 주목하여 판단하였는지 확인할 수 있다.

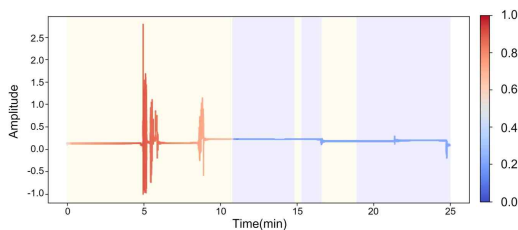


Fig. 5 Visualize model results using Attention heat map

IV. 결 론

본 연구에서는 상용 웨어러블 디바이스의 3축 가속도 센서를 기반으로 고 연령층을 대상으로 한 수면 중 각성을 감지하는 알고리즘을 개발하였다. 고 연령층은 타 연령층에 비해 수면 중 잦은 움직임이 있기 때문에 이를 반영 할 수 있는 수면 중 움직임 관련 특징점을 기반으로 이를 보완하고자 하였다.

설계한 네트워크를 7 fold validation을 통해 평가 하였으며, 모든 Fold의 결과에서 전문가와의 일치도를 확인하는 Cohen's kappa 계수가 적절하거나 충분한 일치도를 가지는 것을 확인할 수 있었다.

사용한 특징점 5종에 대해서 중요도를 파악하는 Ablation test를 수행 하였을 때, Standard deviation(표준편차)을 적용 하였을 때가 가장 큰 성능 향상이 있었으며, 시간 영역의 특징점이 주파수 영역의 특징점에 비해 더 높은 성능 향상을 보인 것을 확인할 수 있었다. 또한 본 연구에서는 시계열 데이터를 효과적으로 다루기 위해 Inception time과 TCN을 조합한 네트워크를 설계하였다. 이에 대한 유효성을 검증하기 위해 기존 연구에서 사용된 네트워크인 LSTM과 CNN-LSTM, 그리고 가장 기초적 네트워크인 1D CNN을 이용해 비교하였다. 결과적으로 본 연구에서 제시한 모든 성능 평가 지표에서 설계한 네트워크가 가장 좋은 성능을 보인다는 것을 확인하였다. 이를 통해 본 연구에서 설계한 딥러닝 모델이 수면 중 움직임이 많은 노인과 같은 인구에게 이전에 개발된 알고리즘 구조보다 안정적으로 적용할 수 있음을 검증하였다.

검증한 모델을 바탕으로 실제 수면 중 각성 감지를 통한 수면 질 평가가 가능한지 여부를 확인하기 위해 수면 파라미터 4종(TST, SOL, SE, WASO)을 추정하여 전문가가 평가한 결과와 비교 분석을 진행하였다. 결과적으로, SOL을 제외한 나머지 파라미터에서 유의한 상관관계를 가진다는 것을 확인함으로써 본 연구의 네트워크가 고 연령층의 사람들에게 실 적용 할 수 있는 가능성 또한 확인했다.

본 연구는 웨어러블 디바이스만을 이용하여 고 연령층을 대상으로 한 수면 중 각성을 감지하는 네트워크를 통해 실 적용 가능성을 보여준 첫 연구이다. 개발한 알고리즘으로 고 연령층을 평가 했음에도 전문가의 의견 일치도에 해당하는 Cohen's kappa 계수에서 이전 연구 및

개발된 알고리즘에 비해 성능이 준하거나 더 우수한 것을 확인할 수 있었다. 그러나 제시한 네트워크는 고 연령층으로만 구성된 데이터만을 이용하였기 때문에 모든 연령층에 사용할 수 있을지는 추가적인 검증이 필요하다. 이에 따라 고 연령층이 아닌 타 연령층에도 해당 딥러닝 네트워크를 적용 하는 것에 대한 후속 연구가 요구된다. 전 연령대에 대한 실적용 여부가 검증되어진다면, 손목에 착용하는 웨어러블 디바이스만으로도 실시간 수면 모니터링과 개인 맞춤형 수면 질 평가를 통해 수면 관리를 제공하는 데 도움을 줄 수 있을 것으로 기대된다.

ACKNOWLEDGEMENTS

(1) This work was supported by 'Smart HealthCare Program(www.kipot.or.kr)' funded by the Korean National Police Agency (KNPA, Korea). [Project Name: Development of wearable system for acquiring lifelog data and customized healthcare service for police officers/ Project Number: 220222M04]

(2) This work was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. NRF-2021M3E5D2A01022397)

REFERENCES

- [1] M. R. Fox, "The importance of sleep," *Nursing Standard (through 2013)*, vol. 13, no. 24, pp. 44, Mar. 1999. DOI: 10.7748/ns1999.03.13.24.44.c7485.
- [2] C. Baglioni, G. Battagliese, B. Feige, K. Spiegelhalter, C. Nissen, U. Voderholzer, C. Lombardo, and D. Riemann, "Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies," *Journal of affective disorders*, vol. 135, no. 1-3, pp. 10-19, Dec. 2011. DOI: 10.1016/j.jad.2011.01.011.
- [3] A. Tuohuti, Z. Lin, J. Cai, and X. Chen, "Can portable sleep monitors replace polysomnography for diagnosis of pediatric OSA: a systematic review and meta-analysis," *European Archives of Oto-Rhino-Laryngology*, vol. 280, no. 8, pp. 2181-2189, Jul. 2023. DOI: 10.1007/s00405-023-08095-6.
- [4] J.V. Rundo and R. Downey III, "Polysomnography," *Handbook of clinical neurology*, vol. 160, pp. 381-392, Jul. 2019. DOI: 10.1016/B978-0-444-64032-1.00025-4.
- [5] D. Moser, P. Anderer, G. Gruber, S. Parapatics, E. Loretz, M. Boeck, G. Kloesch, E. Heller, and G. Dorffner, "Sleep classification according to AASM and Rechtschaffen & Kales: effects on sleep scoring parameters," *Sleep*, vol. 32, no. 2, pp. 139-149, Feb. 2009. DOI: 10.1093/sleep/32.2.139.
- [6] M. Partinen, A. Jamieson, and C. Guilleminault, "Long-term outcome for obstructive sleep apnea syndrome patients: mortality," *Chest*, vol. 94, no. 6, pp. 1200-1204, Dec. 1988. DOI: 10.1378/chest.94.6.1200.
- [7] H. Loewen, L. Korngut, K. Rimmer, O. Damji, T. C. Turin, and P. J. Hanly, "Limitations of split-night polysomnography for the diagnosis of nocturnal hypoventilation and titration of non-invasive positive pressure ventilation in amyotrophic lateral sclerosis," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 15, no. 7-8, pp. 494-498, Aug. 2014. DOI: 10.3109/21678421.2014.942866.
- [8] L. C. Markun and A. Sampat, "Clinician-focused overview and developments in polysomnography," *Current sleep medicine reports*, vol. 6, pp. 309-321, Nov. 2020. DOI: 10.1007/s40675-020-00197-5.
- [9] E. Guillo, C. Lemey, M. Simonnet, M. Walter, E. Baca-Garcia, V. Masetti, S. Moga, M. Larsen, J. Ropars, and S. Berrouguet, "Clinical applications of mobile health wearable - based sleep monitoring: systematic review," *JMIR mHealth and uHealth*, vol. 8, no. 4, p. e10733, Apr. 2020. DOI: 10.2196/10733.
- [10] S. Kwon, H. Kim, and W.-H. Yeo, "Recent advances in wearable sensors and portable electronics for sleep monitoring," *iScience*, vol. 24, no. 5, pp. 102461, May 2021. DOI: 10.1016/j.isci.2021.102461.
- [11] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Ofli, J. Srivastava, A. Elmagarmid, T. Arora and S. Taheri, "Sleep quality prediction from wearable data using deep learning," *JMIR mHealth and uHealth*, vol. 4, no. 4, pp. e6562, Apr. 2016. DOI: 10.2196/mhealth.6562.
- [12] Z. Chen, M. Wu, K. Gao, J. Wu, J. Ding, Z. Zeng and X. Li, "A novel ensemble deep learning approach for sleep-wake detection using heart rate variability and acceleration," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 5, pp. 803-812, Oct. 2020. DOI: 10.1109/TETCI.2020.2996943.

- [13] Z. Chen, M. Wu, W. Cui, C. Liu, and X. Li, "An attention based CNN-LSTM approach for sleep-wake detection with heterogeneous sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3270-3277, Sep. 2020. DOI: 10.1109/JBHI.2020.3006145.
- [14] S. Kudo, Z. Chen, N. Ono, M. Altaf-Ul-Amin, S. Kanaya, and M. Huang, "Deep Learning-Based Sleep Staging with Acceleration and Heart Rate Data of a Consumer Wearable Device," in *2022 IEEE 4th Global Conference on Life Sciences and Technologies (LifeTech)*, Osaka: Japan, pp. 305-307, 2022. DOI: 10.1109/LifeTech53646.2022.9754876.
- [15] Z. Beattie, Y. Oyang, A. Statan, A. Ghoreyshi, A. Pantelopoulous, A. Russel, and C. Heneghan, "Estimation of sleep stages in a healthy adult population from optical plethysmography and accelerometer signals," *Physiological measurement*, vol. 38, no. 11, pp. 1968-1979, Oct. 2017. DOI: 10.1088/1361-6579/aa9047.
- [16] M. Olsen, J. M. Zeitzer, R. N. Richardson, and E. Mignot, "A flexible deep learning architecture for temporal sleep stage classification using accelerometry and photoplethysmography," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 1, pp. 228-237, Jan. 2022. DOI: 10.1109/TBME.2022.3187945.
- [17] J. Razjouyan, H. Lee, S. Parthasarathy, J. Mohler, A. Sharafkhaneh, and B. Najafi, "Improving sleep quality assessment using wearable sensors by including information from postural/sleep position changes and body acceleration: a comparison of chest-worn sensors, wrist actigraphy, and polysomnography," *Journal of Clinical Sleep Medicine*, vol. 13, no. 11, pp. 1301-1310, Nov. 2017. DOI: 10.5664/jcsm.6802.
- [18] S. Haghayegh, S. Khoshnevis, M. H. Smolensky, K. R. Diller, and R. J. Castriotta, "Performance comparison of different interpretative algorithms utilized to derive sleep parameters from wrist actigraphy data," *Chronobiology international*, vol. 36, no. 12, pp. 1752-1760, Oct. 2019. DOI: 10.1080/07420528.2019.1679826.
- [19] J. K. Devine, E. D. Chinoy, R. R. Markwald, L. P. Schwartz, and S. R. Hursh, "Validation of Zulu watch against polysomnography and actigraphy for on-wrist sleep-wake determination and sleep-depth estimation," *Sensors*, vol. 21, no. 1, pp. 1-17, Dec. 2020. DOI: 10.3390/s21010076.
- [20] S. Haghayegh, S. Khoshnevis, M. H. Smolensky, K. R. Diller, and R. J. Castriotta, "Performance assessment of new-generation Fitbit technology in deriving sleep parameters and stages," *Chronobiology international*, vol. 37, no. 1, pp. 47-59, Nov. 2020. DOI: 10.1080/07420528.2019.1682006.
- [21] J. Wei and J. Boger, "Sleep detection for younger adults, healthy older adults, and older adults living with dementia using wrist temperature and actigraphy: prototype testing and case study analysis," *JMIR mHealth and uHealth*, vol. 9, no. 6, pp. e26462, Jun. 2021. DOI: 10.2196/26462.
- [22] ActiGraph. wGT3X-BT [Internet]. Available: <https://theactigraph.com/actigraph-wgt3x-bt>.
- [23] E. Hertenstein, M. Feige, S. Gmeiner, T. Kienzler, A. Spiegelhalder, C. Voderholzer, and D. Riemann, "Reference data for polysomnography-measured and subjective sleep in healthy adults," *Journal of clinical sleep medicine*, vol. 14, no. 4, pp. 523-532, Apr. 2018. DOI: 10.5664/jcsm.7036.
- [24] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Procedia Computer Science*, vol. 34, pp. 450-457, Aug. 2014. DOI: 10.1016/j.procs.2014.07.009.
- [25] P. Casale, O. Pujol, and P. Radeva, "Human activity recognition from accelerometer data using a wearable device," in *Pattern Recognition and Image Analysis: 5th Iberian Conference*, Las Palmas de Gran Canaria: Spain, pp. 289-296, 2011. DOI: 10.1007/978-3-.
- [26] X. Li, Y. Wang, B. Zhang, and J. Ma, "PSDRNN: An efficient and effective HAR scheme based on feature extraction and deep learning," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6703-6713, Jan. 2020. DOI: 10.3390/s21062141
- [27] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. Muller, "Inceptiontime: Finding alexnet for time series classification," *Data Mining and Knowledge Discovery*, vol. 34, no. 6, pp. 1936-1962, Sep. 2020. DOI: 10.1007/s10618-020-00710-y.
- [28] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu: Hawaii, pp. 156-165, 2017. DOI: 10.1190/CVPR.2017.113.
- [29] F. C. Rau, A. Wigger, B. Tellhelm, M. Zwick, S. Klumpp, A. Neumann, B. Oltersdorf, K. Amort and M. Kramer, "Observer variability and sensitivity of radiographic diagnosis of canine medial coronoid disease," *Tierarztl Prax Ausg K Kleintiere Heimtiere*, vol. 39, no. 5, pp. 313-322, 2011. DOI: 10.1055/s-0038-1623595.



박범준(Beomjun Bark)

울산대학교 전기공학부 의공학전공 학사
한양대학교 생체의공학과 석사과정

※ 관심분야 : Time series data analysis, Artificial intelligence, Uncertainty, Sleep analysis



남보름(Borum Nam)

연세대학교 컴퓨터공학과 학사
한양대학교 융합전자공학과 석박통합과정

※ 관심분야 : Deep learning network, Explainable AI, Medical informatics, Facial expression analysis



이종실(Jongshill Lee)

인하대학교 전자공학과 석사
인하대학교 전자공학과 박사

※ 관심분야 : Biometric Authentication, Biomedical Signal Processing, Cardiovascular system



김인영(In Young Kim)

서울대학교 의학과 석사
서울대학교 의학과 의공학전공 박사
삼성종합기술원 의료기기연구팀 의료전자랩 수석연구원
한양대학교 의과대학 의공학교실 주임교수

※ 관심분야 : Bio-engineering, Bio-sensor, Medical informatics, Bio-Signal Acquisition and Analysis