Nature and Science of Sleep

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Brain Age Estimation from Overnight Sleep Electroencephalography with Multi-Flow Sequence Learning

Di Zhang[1,2], Yichong She[1,2], Jinbo Sun[1,2], Yapeng Cui[1,2], Xuejuan Yang[1,2], Xiao Zeng[1,2], Wei Qin[1,2]

[1]Engineering Research Center of Molecular and Neuro Imaging of the Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, 710126, People's Republic of China; [2]Intelligent Non-Invasive Neuromodulation Technology and Transformation Joint Laboratory, Xidian University, Xi'an, People's Republic of China

Correspondence: Xiao Zeng, Xidian University, 266 Xifeng Road, Xinglong, Xi'an, Shaanxi, People's Republic of China, 710126, Email xiaozeng1024@gmail.com

Purpose: This study aims to improve brain age estimation by developing a novel deep learning model utilizing overnight electroencephalography (EEG) data.

Methods: We address limitations in current brain age prediction methods by proposing a model trained and evaluated on multiple cohort data, covering a broad age range. The model employs a one-dimensional Swin Transformer to efficiently extract complex patterns from sleep EEG signals and a convolutional neural network with attentional mechanisms to summarize sleep structural features. A multi-flow learning-based framework attentively merges these two features, employing sleep structural information to direct and augment the EEG features. A post-prediction model is designed to integrate the age-related features throughout the night. Furthermore, we propose a DecadeCE loss function to address the problem of an uneven age distribution.

Results: We utilized 18,767 polysomnograms (PSGs) from 13,616 subjects to develop and evaluate the proposed model. The model achieves a mean absolute error (MAE) of 4.19 and a correlation of 0.97 on the mixed-cohort test set, and an MAE of 6.18 years and a correlation of 0.78 on an independent test set. Our brain age estimation work reduced the error by more than 1 year compared to other studies that also used EEG, achieving the level of neuroimaging. The estimated brain age index demonstrated longitudinal sensitivity and exhibited a significant increase of 1.27 years in individuals with psychiatric or neurological disorders relative to healthy individuals.

Conclusion: The multi-flow deep learning model proposed in this study, based on overnight EEG, represents a more accurate approach for estimating brain age. The utilization of overnight sleep EEG for the prediction of brain age is both cost-effective and adept at capturing dynamic changes. These findings demonstrate the potential of EEG in predicting brain age, presenting a noninvasive and accessible method for assessing brain aging.

Keywords: brain age, sleep polysomnography, electroencephalography, deep learning, swin transformer

Introduction

In recent years, an increasing number of neuropsychiatric disorders have been shown to be associated with increased brain age. These include traumatic brain injury,[1] schizophrenia,[2,3] bipolar disorder,[3] epilepsy,[4] Alzheimer's disease and Parkinson's disease,[5] major depressive disorder,[6,7] dementia[8] and among others.[9] Researches indicate that people with these disorders tend to have changes in brain structure and function, exhibiting accelerated brain aging compared to their chronological age (CA). To better understand brain age acceleration, researchers have developed the brain age index (BAI). It represents the discrepancy between an individual's estimated brain age (BA) and its CA, presenting a potential novel biomarker for age acceleration.[10] Accurately predicting brain age can provide valuable insights into the risk of cognitive decline, neurological disorders, psychiatric conditions, and even mortality.[7,11]

Currently, a common non-invasive approach for estimating BA involves imaging modalities, where machine learning techniques are used to evaluate the extent of brain aging based on imaging features. Traditional approaches often rely on feature-

Received: 10 February 2024
Accepted: 19 June 2024
Published: 1 July 2024

based regression modeling, including the extraction of principal components,[12] cortical thickness and surface curvature,[13] as well as the volumes of gray matter, white matter, and cerebrospinal fluid.[14] These feature extraction methods introduce prior knowledge, which aids in the interpretability and guidance of results. Recently, deep learning approaches have attracted significant interest among researchers. These data-driven methods can extract deep features and reduce the reliance on prior knowledge. Deep neural network models, particularly those based on convolutional neural networks (CNNs), have demonstrated improved accuracy in BA estimation across various studies.[15–17] However, this approach is limited by its sensitivity disadvantages, as structural brain aging is typically slow and gradual, and the high costs restrict its use for tracking and at-home monitoring of dynamic neural changes. Consequently, it is essential to integrate neural data from other modalities in order to provide a more comprehensive assessment of brain age.

As our understanding of aging evolves, new paradigms are being proposed. The analysis of electroencephalogram (EEG) during sleep has become a new method to enhance BAI predictions. This approach focuses on the dynamic and functional aspects of sleep patterns and EEG signals, rather than the structural aspect. This unique approach allows analysis of the intrinsic effects of aging on brain activity. It is also participant-friendly, avoiding the high costs and complexity associated with imaging techniques. Moreover, as EEG and sleep are inherently dynamic, this approach provides a more sensitive assessment of the impact on brain health than neuroimaging, allowing for a timely assessment of the effectiveness and potential risks of interventions.

Researchers have achieved significant results in predicting BA using machine learning and deep learning methods applied to sleep EEG data. Zoubi et al proposed a stack-ensemble learning-based model to predict BA from manually extracted features including amplitude domain, range domain, spectral domain, and fractal dimension domain, which resulted in MAE= 6.87 years.[18] Sun et al employed a publicly available dataset to predict BA from sleep structural features and EEG features, achieving MAEs of 23.3 years and 7.6 years, respectively.[19] The introduction of deep learning methods significantly improved the BA prediction performance. Yook et al introduced a 3D DenseNet model to analyze multichannel scalograms of EEG and achieved an MAE of 5.4 years.[20] Brink-Kjaer et al developed a deep framework for age estimation using multiple datasets. The framework performed short-term age estimation of 5-min physiological signals using a CNN model, and integrated overnight results using a long- and short-term memory network, achieving MAEs of 5.8–14.6 years. The MAE using central EEG alone was 6.77 years.[21]

However, challenges remain when using sleep EEG for brain age prediction. Firstly, most existing methods for calculating BAI are limited in formulation without model optimization for sleep characteristics. Secondly, these methods were developed on relatively small data sizes or single cohorts, limiting their generalizability and applicability. Finally, compared to other perspectives for assessing brain age, methods using EEG still have room for improvement in accuracy.

This study aims to address these challenges by developing a novel multi-flow deep learning model. The model was specifically developed to enhance the accuracy and robustness of brain age estimation from overnight sleep EEGs. The major contributions of this work are presented below.

1) Integration of Sleep EEG for Brain Age Estimation: This study utilizes overnight sleep EEG for predicting brain age, which is a dynamic, cost-effective, and suitable for therapeutic interventions approach rather than traditional neuroimaging methods.

2) Development of the multi-flow model: We present a multi-flow deep learning architecture for the sleep-specific problem. The model utilizes an attention mechanism with two modalities, EEG features and sleep structure information, to enhance accuracy and robustness.

3) Sleep Dynamics Investigation: This study reveals that the stage N2 significantly contributes to accurate brain age prediction and underscores the close relationship between light sleep and accelerated brain aging.

4) High Accuracy and Expanded Applicability: The proposed model achieves prediction accuracy comparable to that of neuroimaging techniques. Furthermore, it is capable of detecting accelerated brain aging in populations with psychiatric or neurological disorders, thereby demonstrating its potential clinical utility.

Materials and Methods

Datasets

This study utilized data from a total of 8 publicly available cohorts, each encompassing diverse research objectives or populations:

- The Apnea, Bariatric surgery, and CPAP (ABC) study focused on comparing the effectiveness of bariatric surgery and CPAP therapy for severe obstructive sleep apnea (OSA) and morbid obesity patients.[22]
- The Cleveland Family Study (CFS) investigated sleep apnea within a large family-based context to explore its genetic and environmental determinants.[23]
- The Home Positive Airway Pressure (HPAP) study examined the efficacy of home-based portable monitoring for diagnosing and treating OSA compared to laboratory-based PSG.[24]
- The Multi-Ethnic Study of Atherosclerosis (MESA) aimed to understand factors associated with cardiovascular disease progression by analyzing sleep patterns across different ethnic groups.[25]
- The MrOS Sleep Study explored the relationship between sleep patterns, sleep-disordered breathing, and cognition in elderly men.[26]
- The NCH Sleep DataBank (NCHSDB) provided insights into pediatric sleep patterns in children in the United States.[27]
- The Sleep Heart Health Study (SHHS) examined sleep-disordered breathing as a risk factor for cardiovascular disease development in a large prospective cohort.[28]
- The Wisconsin Sleep Cohort (WSC) is a still ongoing longitudinal study focusing on various sleep disorders, collected in the United States.[29,30]

These cohorts collectively comprised 18,767 PSG records from 13,616 subjects, with each PSG including EEG recordings from at least two central channels. ABC, MROS, and SHHS, in particular, contribute multiple datasets due to the inclusion of multiple visits from the same subjects. HPAP serves as an independent test set held-out from neural network training. These cohorts can be approved by the National Sleep Research Resource.[22,31] There are a total of 12 datasets included in these 8 cohorts covering a wide range of age groups, and the Figure 1 visually depicts the distribution of the CAs for each dataset. The demographics and profiles of the included subjects are shown in Table 1.

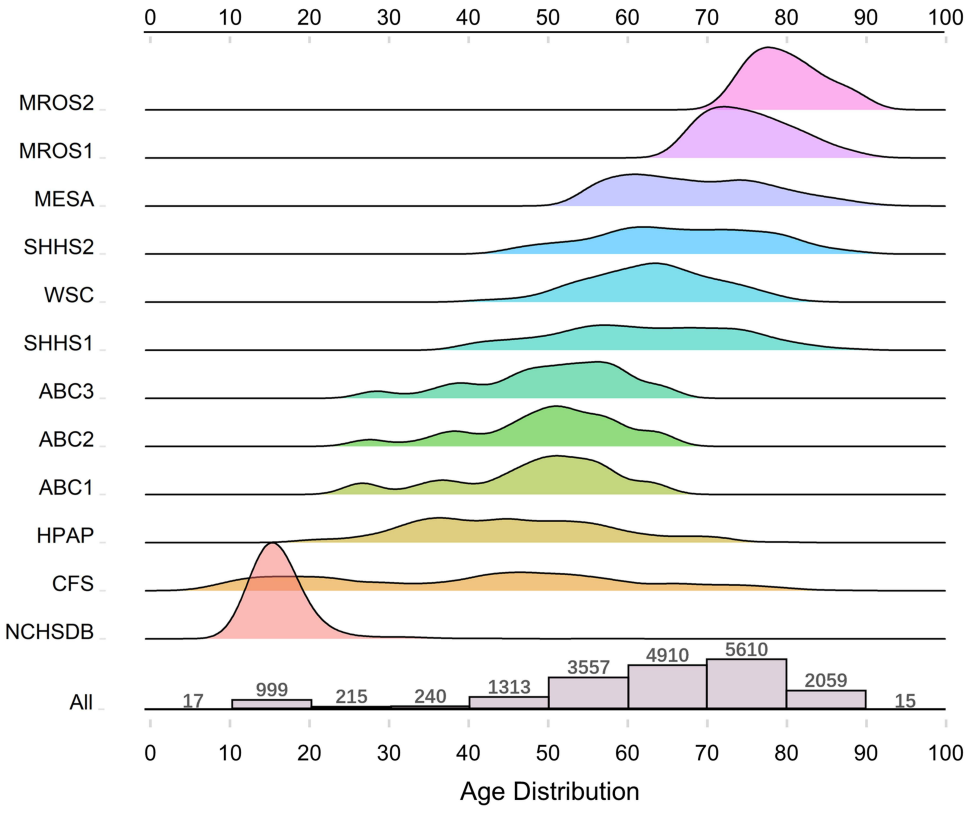

Figure 1 Distribution of chronological age in the adopted dataset. The "All" row displays the number of individuals aged 1–100 years, grouped in 10-year intervals.

Table 1 Demographics and Profiles of Subjects Included from Each Cohort

Purpose	Cohort Name	Subject Num (Train/Valid/Test)	Subgroup or Visit	PSG Num	EEG Length (Minute)	CA (year)	Sex % (M/F)	BMI	AHI (Events/ Hour)
Train and evaluation	ABC	49 (37/5/7)	Baseline	49	508.5±20.9	48.8 ± 9.8	57/43	38.9±2.9	40.4±27.8
			Month9	42	500.2±13.9	50.5±9.3	58/42	36.3±3.6	28.8±28.0
			Month18	40	502.7±21.7	51.0±9.3	62/38	36.2±4.1	27.6±21.7
	CFS	722 (572/50/100)	Visit5	722	593.1±53.6	41.4±19.4	45/55	32.3±9.4	12.5±17.0
	MESA	2034 (1884/50/100)	–	2034	634.7±95.5	69.4±9.1	46/54	–	24.1±19.5
	MROS	2885 (2735/50/100)	Visit1	2885	647.9±102.2	76.4±5.5	100/0	27.2±3.8	–
			Visit2	1001	783.8±174.9	81.0±4.4	100/0	26.9±3.8	–
	NCHSDB	907 (757/50/100)	Multiple	982	454.1±45.1	16.8±3.8	52/48	32.4±12.5	–
	SHHS	5713 (5563/50/100)	Visit1	5667	506.0±37.5	63.1±11.2	48/52	28.2±5.1	18.0±16.1
			Visit2	2621	602.0±68.2	67.6±10.4	81/19	28.3±5.0	18.4±16.3
	WSC	1116 (966/50/100)	Multiple	2534	459.3±46.3	62.3±8.5	54/46	31.7±7.3	11.8±16.4
Hold-out evaluation	HPAP	190 (0/0/190)	Full-night	190	465.2±65.6	46.0±12.0	52/48	36.9±9.0	13.8±18.3

Notes: ABC, MROS, and SHHS are reported separately with multiple visits.
Abbreviations: PSG, Polysomnography; CA, chronological age; BMI, Body Mass Index; AHI, Apnea–Hypopnea Index. ABC, the Apnea, Bariatric surgery, and CPAP; CFS, the Cleveland Family Study; MESA, the Multi-Ethnic Study of Atherosclerosis; MROS, the MrOS Sleep Study; NCHSDB, the NCH Sleep DataBank; SHHS, the Sleep Heart Health Study; WSC, the Wisconsin Sleep Cohort; HPAP, the Home Positive Airway Pressure.

Data Pre-Processing and Preparation

The raw time series data is composed of six EEG channels (F3, F4, C3, C4, O1, O2, if available) that were filtered (bandpass 0.3–35 Hz), clipped (−1000 to 1000 μV), and resampled (128 Hz). Three channels of EEG were used for model input: frontal, central and occipital regions of the right hemisphere. (The left hemisphere was used as an alternative if signal quality was poor) Missing channels were manually set to zero during training and testing. During training, each EEG inputs underwent various data enhancement techniques: 1) Channel missing, with a 10% probability of randomly masking certain EEG channels with zeros. 2) Channel reversal, with a 10% probability of flipping the polarity of each EEG channel. 3) Gaussian noise, with a 10% probability of being added to the EEG signals as random noise to introduce variability and mimic real-world conditions (noise amplitude = 0–30% of the raw signal).

Except for HPAP, the other data were split by subject according to Ref,[32] into training set, validation set (10%, up to 50, per cohort), and test set (15%, up to 100, per cohort). The split data were mixed among cohorts for model training and evaluation, and HPAP was used as an independent test set solely to evaluate the generalization of the predictive model.

Two-Flow Prediction: Brain Age Estimation by 5-Minute Epoch

The multi-flow sequence learning network proposed in this study uses two scales of modal sequences as inputs: a 5-minute EEG epoch and the hypnodensity throughout the night. In this model, EEG deep features are crucial in estimating brain age, while sleep structural features from hypnodensity enhance the EEG deep features. We intercepted the overnight EEG into 5-minute epochs with 50% overlap as separate samples. The method described in Ref[33] was utilized for sleep staging and to obtain hypnodensity. The hypnodensity demonstrates the possibilities of sleep stages for each epoch, thereby capturing the dynamic shifts and variability of the sleep structure. By employing a deep feature extraction of hypnodensity, it is possible to generalize beyond manually extracted sleep structural information such as total sleep time (TST), sleep efficiency (SE), or wake after sleep onset (WASO).[34] Figure 2 presents the overall framework and details of the proposed network model.

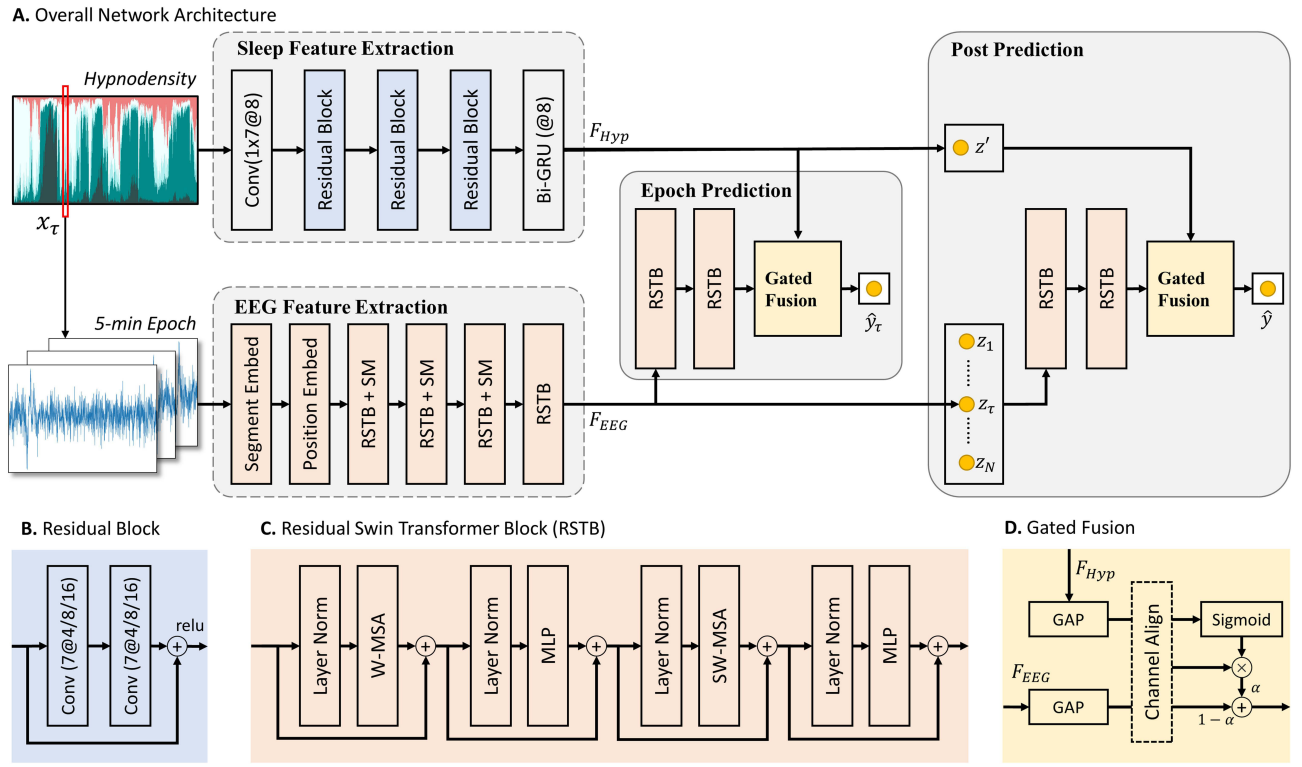

Figure 2 Neural network structures. (**A**) An overview architecture of the proposed brain age estimation model containing epoch-wise prediction and post-prediction. (**B**) Design of the CNN residual module for extracting sleep features. (**C**) Design of the residual Swin Transformer block for extracting EEG features. (**D**) Design of the attentional mechanism for gated feature fusion from two flows.

A one-dimensional Swin Transformer neural network is designed for sleep EEG feature extraction. The multi-layered feature pyramid structure inherent in the network model proved helpful in capturing information embedded in EEG features at various scales. For EEG input, the segment embedding is proposed, which involves partitioning the sequence into non-overlapping segments through spaced sampling and projecting them into a high-dimensional space via linear embedding. A trainable vector is incorporated into the input's channel dimension for position embedding. Four pairs of Swin Transformer layers are employed for feature extraction, encompassing diverse temporal scales. Each pair comprises a window (W-) and a shifted window (SW-) based multi-head self-attention (MSA) module. These modules perform self-attention mapping along the time dimension within the window and feature interactions through the shift window mechanism. Furthermore, we introduce segment merging (SM) to reduce feature dimensionality and preserve sequence information, which is a one-dimensional implementation of patch merging in Swin Transformer.

Another one-dimensional neural network is also introduced to extract sleep structure features from hypnodensity, which are used to attentional enhance the abovementioned EEG features. The network is composed of three convolutional residual blocks connected in series and a gated recurrent neural network. It is implemented in parallel with the aforementioned EEG feature extractor, forming two feature extraction flows. The two flows' outputs utilize a global average pooling layer to remove sequence dependence and maintain channel information. The extracted sleep features z' are subsequently used to attention-enhance EEG features.

To consider the contribution of the sleep structure to the brain age estimation, a gated multimodal fusion method is introduced.[35] In this method, the interactive attention map of the two modalities is implemented through a bilinear-gating approach, which enables hypnodensity enhancement on the EEG features, the fusion output o(F_H, F_E) is calculated as:

$$Attn(F_H, F_E) = UF_H \otimes VF_E,$$

$$o(F_H, F_E) = W(\lambda \cdot Attn(F_H, F_E) + VF_E),$$

where *Attn* represents the attention map, F_H is normalized hypnodensity feature, F_E is EEG epoch feature, *W, U, V* are weight matrices, λ is a harmonic weight, and the symbol ⊗ represents matrix-vector multiplication. The fusion utilizes information on long-distance sleep dynamics to enhance the constraints on high-dimensional EEG features, enabling more robust brain age estimation. The model takes in the τ-th EEG epoch and the overnight hypnodensity as inputs and produces an estimated brain age value $\hat{y}_\tau$, for that epoch. The model is trained individually to minimize the difference between $\hat{y}_\tau$ and the chronological age *y*.

Post-Prediction: Ensemble Overnight Brain Age

After fine-tuning the two-flow network for epoch prediction, we introduce a new network model to estimate the brain age over the entire night. The potential features *z* are defined as the deep features in the last layer of the two-flow network. These potential features are averaged over the temporal dimensions, preserving channel information. The post-prediction network takes the ensemble of *z* from multiple epochs during the overnight period as input.

To capture extended dependencies within the sleep structure, we introduce a two-layer Swin Transformer with window sizes resembling the sleep cycle (90 minutes). For the final step, we apply the gated fusion mentioned above, using *z'* to attentively enhance the sequence feature *z* and generate category predictions with a fully connected network.

Loss Function

The age distribution of the training samples is inhomogeneous, which presents a significant challenge for the regression problem. Given that age is a continuous variable, we proposed the DecadeCE Loss, which transforms the regression problem into an ordered multiclassification problem, with each class spanning 10 years. The DecadeCE Loss samples numerical labels from a Gaussian distribution for each range to obtain class vectors, preventing precision lost. The soft cross-entropy between these vectors and the model output is then calculated. For a given age y, it is first discretized into a distribution that follows the N(*y*, σ), calculated as:

$$p(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mathrm{t}-\mathrm{y})^2}{2\sigma^2}\right)$$

where t=5,15,…,95, representing the midpoints of each decade. The numerical *y* value is converted to vector *Y* containing 10 levels of probability as

$$Y^{(t)} = p(t)/max_t(p(t)).$$

Then, the loss between a target *Y* and a predicted ŷ is computed by weighted cross-entropy, expressed as:

$$Loss = -\frac{1}{10}\sum_t \left[Y^{(t)} log\left(\hat{Y}^{(t)}\right) + \left(1 - Y^{(t)}\right)\log\left(1 - \hat{Y}^{(t)}\right)\right]$$

where $\hat{y}^{(t)}$ denotes the predicted vector of age range (t-5:t+5).

During model inference, the 10 class outputs of the neural network are passed through the softmax function to obtain their corresponding predictive probabilities, calculated as:

$$\hat{P} = softmax\left(\hat{Y}\right)$$

Therefore, $P = \{\hat{P}^{(t)} | t = 5, 15, \ldots, 95\}$ is composed of probabilities from 10 classes. We define the ordered median predictive probability as the output BA value of the model to enhance the stability of converting continuous classification results into regression values. It is calculated as follows:

$$P(x) = \frac{1}{10}\sum_{i=1}^{x}\sum_t \hat{P}^{(t)} \cdot 1_{[t-5<i\leq t+5]}$$

$$BA = \min(x \mid P(x) \geq 0.5)$$

where i ranges from 1 to 100 in one-year increments, while the indicator function **1** equals one when i is within the decadal range of t and zero otherwise. Briefly, in steps of 1 year, the age of the left and right area equivalents of the 10 class probability densities was determined as the predicted BA.

Figure 3 illustrates the DecadeLoss and the inference process. It demonstrates how the regression problem is converted into a continuous categorical classification problem and how predicted multi-class probabilities are translated into a numerical prediction.

Training and Evaluation

For the epoch prediction neural network training, we used a batch size of 64 EEG and hypnodensity combinations as inputs for one iteration of training, totaling 320,000 iterations. An AdamW[36] optimizer with a learning rate of 0.001 and cosine decayed to 1e-6 is used. For the post prediction network training, we utilized identical optimizer parameters, set the batch size to 128, and completed 50 iterations of training data.

The model's best MAE on the validation set was used for test evaluation in each training session. The training process consisted of two steps, which took approximately 30 hours and 0.2 hours, respectively, on a workstation with an Intel Core (TM) i9-10900X CPU and NVIDIA RTX 3080 graphics card. It takes about two seconds to estimate overnight brain age.

Evaluation Metrics and Statistical Analysis

The MAE and Pearson correlation (R) are used to evaluate the predictive performance of the model across varied age ranges.[9] They are denoted as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i|$$

$$R = \frac{\sum(y-\bar{y})(\hat{y}-\bar{\hat{y}})}{\sqrt{\sum(y-\bar{y})^2\sum(\hat{y}-\bar{\hat{y}})^2}},$$

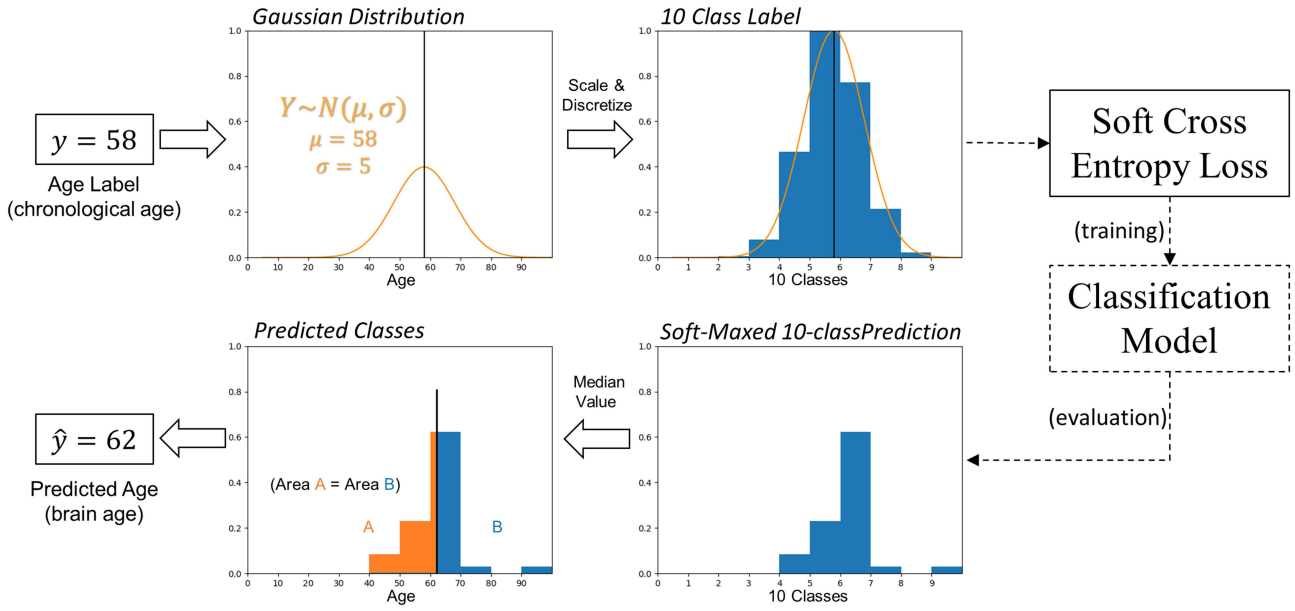

Figure 3 Schematic representation of the proposed DecadeCE Loss and the inferencing results. The regression problem for *y* is initially transformed into a 10-consecutive-category classification problem for the training and inferring of the neural network model. The 10 output categories are converted into the predicted output ŷ by the median probability method.

where $\hat{y}$ and y are predicted brain age and chronological age labels, respectively. The BAI is defined as BA-CA and quantifies the extent to which predicted brain age exceeds chronological age.

Unless otherwise specified, the results presented in this paper are subject-wise, meaning that they aggregate each individual's BA assessment with the difference from their CA. Paired t-tests compare model performance on the same data, while independent t-tests and ANOVA assess variations between different datasets or populations.

Experimental Setup

This section outlines the experimental procedures conducted to assess the performance, robustness, and reliability of our model, as well as to investigate the BA and BAI.

Performance Evaluation

To evaluate the performance of the model, three evaluation experiments were conducted. The mixed-cohort evaluation involved training and testing the multiple cohort data, except HPAP, after the aforementioned data splitting. The optimized model was then used for subsequent experimental evaluations. The hold-out evaluation entailed using the optimized model to test it directly on the HPAP data. Finally, the comparative evaluation compared EEG feature extraction performance, introducing four typical deep learning methods for EEG analysis.[21,33,37,38] These methods represent the most recent mainstream deep frameworks: CNN, CNN+GRU, CNN+LSTM, and Transformer.

Ablation Experiments

To understand the specific contributions of certain components in our proposed models, we conducted three ablative experiments by removing or replacing them with other approaches.

We first take apart proposed multi-flow network taking hypnodensity and EEG epochs separately for brain age prediction. The structure of the networks remains unchanged except that the merging module is replaced by a simple classification header. To simplify the aggregation of overnight estimates from different approaches, we replace the post-prediction model with the estimated mean value of the overnight epoch.

Next, to assess the effectiveness of proposed DecadeCE Loss in model training, we compared it to two commonly used regression losses: MSE Loss and Huber Loss. They are calculated as follows:

$$MSE\,Loss(y,\hat{y}) = (y_i - \hat{y}_i)^2$$

$$Huber\,Loss(y_i,\hat{y}_i) = \begin{cases} 0.5(y_i - \hat{y}_i)^2, & if\,|y_i - \hat{y}_i| \le 5 \\ 5(|y_i - \hat{y}_i| - 2.5), & if\,|y_i - \hat{y}_i| \le 5 \end{cases}$$

where y and $\hat{y}$ are the CA and predicted values, respectively, and the two comparison losses use the optimizer described in Ref.[21] To simplify the aggregation of overnight estimates from different approaches, we replace the post-prediction model with the estimated mean value of the overnight epoch.

Finally, we investigated the effect of the overnight post-prediction model on brain age prediction, in its place: no processing and overnight averaging.

Longitudinal Reliability Evaluation

We investigated the longitudinal reliability of proposed brain age prediction model across multiple follow-up visits in datasets SHHS, MROS, and ABC. By analyzing the incremental changes in predicted BA and CA at each visit, we assessed how this model capturing the natural progression of brain aging over time.

Investigation of Brain Age Index

We explored the relationship between sleep stages and the predictive performance of BA. Here, we defined a 5-minute epoch as a "stage domain", representing the predominant or most likely sleep stage within its duration, identified by the sleep stage occupying the largest area on the hypnodensity. The ability of each stage domain to estimate brain age was quantified individually without post-prediction.

We also investigated the association between specific populations or diseases and the calculated BAI, that is, OSA and psychiatric or neurological disorders on the test set. We used a threshold of 15 on the Apnea Hypopnea Index (AHI) to divide subjects into two groups: no to mild obstructive sleep apnea (NM-OSA), and moderate to severe obstructive sleep apnea (MS-OSA). Furthermore, we divided a portion of the test subjects into two groups: the psychiatric or neurological disorders (PSY) group and the healthy control (HC) group, based on publicly available medical records. This grouping was exclusively performed within the SHHS, CFS, MROS, and WSC cohorts, as only these cohorts provided individual medical records pertaining to neurological or psychiatric conditions. The PSY group consisted subjects with diagnosed Alzheimer's disease, Parkinson's disease, cerebrovascular disease, stroke, dementia, attention-deficit disorder, seizure-sleeping disorder, cerebrovascular disease, depression, anxiety, epilepsy, head trauma, or other psychiatric or neurologic disorders. The HC group consisted of subjects who did not report any neurological or psychiatric disorders.

Results

Prediction Performance

The proposed model was optimized using 17,294 PSG records from 12,514 subjects and validated with 435 PSG records from 305 subjects. The mixed-cohort test set comprises 848 PSG records from 607 subjects, while the hold-out test set includes PSG records from 190 subjects.

In the mixed-cohort evaluation, the mean CA in training, validation, and test sets were 64.62 ± 16.12, 57.02 ± 21.47, and 57.77 ± 21.20 years, respectively. ANOVA revealed a significant difference in CA distribution among these sets ($p < 0.01$), while pairwise t-tests between validation and test sets showed no significant difference ($p = 0.55$). The model attained a predictive performance of MAE = 4.19 years and R = 0.97 on the test set. Figure 4A shows the scatterplot of BA vs CA for the test sample on each cohort, along with the correlation within each cohort. According to the test results of a single cohort, the best MAE is for NCHSDB (2.25 years), the worst is for WSC (5.07 years), the best R is for CFS (0.95), and the worst is for NCHSDB (0.42).

In the hold-out test set, the mean CA was 46.04 ± 12.03, significantly different from that in the mixed-cohort test set ($p < 0.01$, *t*-test). The model attained a predictive performance of MAE = 6.18 years and R = 0.78 on this test set. Figure 4B shows the scatterplot of BA vs CA for the test sample on each cohort, along with their correlation.

Comparison with Existing Approaches

We conduct a comparison by evaluate the performance of our proposed model against recent sleep EEG analyzing approaches in the literature. To maintain fairness, we employ identical EEG channel configurations and input lengths

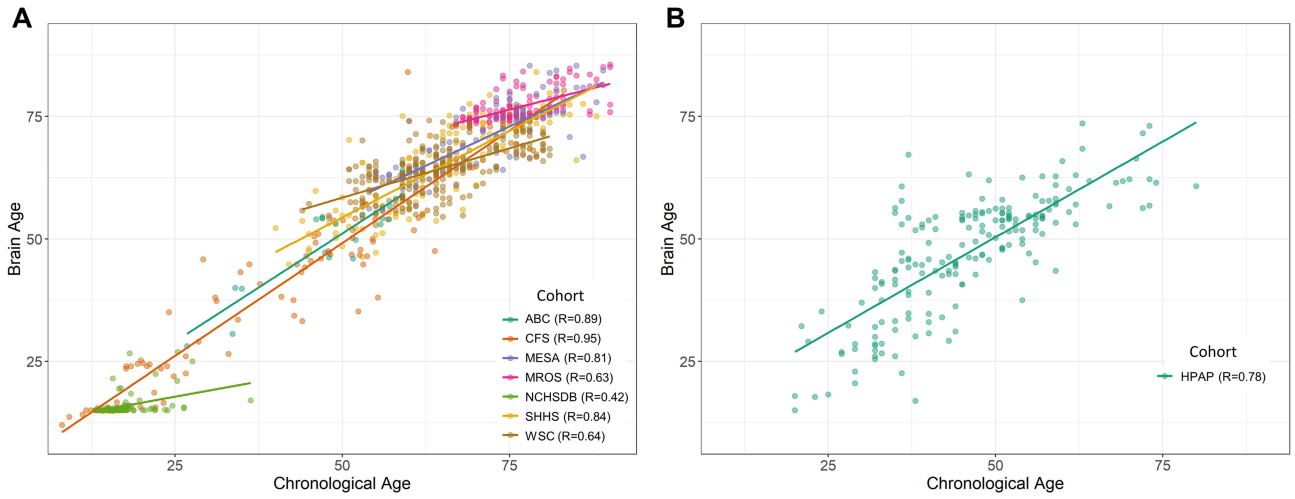

Figure 4 Scatterplots of the predicted brain age (BA) plotted against the chronological age (CA) for individual subjects. (**A**) is on a mixed-cohort test set, separated by each cohort. (**B**) is on a hold-out test cohort.

across all models. For non-native brain age prediction frameworks, we utilize the Huber loss, following the methodology outlined in Ref[21] for model training. Optimization parameters are prioritized to those used in the public projects associated with these models, otherwise they are reconciled according to the hyperparameters used in the models presented in this study. To simplify the aggregation of overnight estimates from different approaches, we replace the post-prediction model with the estimated mean value of the overnight epoch.

As shown in Table 2, the proposed method achieves a closer fit to the chronological age than other EEG analysis methods using the same inputs. It achieves the lowest MAE value and a high level of correlations between BA and CA.

Ablation Results

Table 3 illustrates the ablation flow analysis results. It demonstrates that the unablated two-flow network performs the best. However, when relying solely on hypnodensity, there is a significant reduction in performance. Using only EEG epoch predictions results in only a slight decrease.

Table 4 compares the loss functions. It shows that the regression results optimized using the proposed DecadeCE loss are stronger than those using the other two losses.

Table 5 presents the results of the post-prediction investigation. Each night comprises approximately 200 samples of 5-minute EEG epochs. Predictions for these samples are made through overnight averaging, resulting in a 1–2 year MAE improvement in fitting performance. Additionally, using neural network-based ensembles instead of direct arithmetic averages significantly improves predictive results.

Longitudinal Reliability of Brain Age Predicting

To assess the longitudinal reliability in estimated brain age, we computed the differences in BA and CA (ΔBA, ΔCA) across the two visits for each subject. Subsequently, we introduced brain age acceleration (BAA=ΔBA-ΔCA), a metric that quantifies the increase in brain age relative to chronological age. The longitudinal reliability results are shown in Figure 5.

For SHHS, 69 subjects on the validation and test sets have multiple visits. The mean chronological age difference ΔCA=CA2-CA1 is 5.57 years (SD=0.71), and the mean brain age difference ΔBA=BA2-BA1 is 5.20 years (SD=3.73).

Table 2 Comparison of Brain Age Estimates Between EEG Analysis Approaches

Network Model	Backbone	Mixed-Cohort test		Hold-Out test	
		MAE	R	MAE	R
Zhang[33]	ResAttNet	4.88±4.14	0.957	8.73±6.90	0.601
TinySleepNet[37]	CNN+GRU	5.02±4.41	0.949	9.17±7.35	0.535
Brink-Kjaer[21]	MobileNetV2+LSTM	4.82±3.92	**0.958**	8.82±7.17	**0.611**
Pradeepkumar[38]	Transformer	5.90±5.61	0.926	12.16±8.54	0.443
Proposed	Swin Transformer	**4.76±3.94**	**0.958**	**8.51±6.72**	0.608

Note: Bolded values indicate the closest estimate.
Abbreviation: MAE, mean absolute error.

Table 3 Comparison Results of Ablated Flow

Network Flow	Mixed-Cohort test		Hold-Out test	
	MAE	R	MAE	R
EEG epoch only	4.76±3.94	0.958	8.51±6.72	0.608
Hypnodensity only	7.97±8.49	0.842	15.59±9.54	0.451
EEG epoch + Hypnodensity	**4.64±3.84**	**0.960**	**7.94±6.54**	**0.641**

Notes: Bolded values indicate the closest estimate. *Italics* indicate proposed method in this study.
Abbreviation: MAE, mean absolute error.

Table 4 Comparison Results for Different Loss Functions

Loss Function	Mixed-Cohort test		Hold-Out test	
	MAE	R	MAE	R
MSE Loss	5.13±4.27	0.930	8.87±6.95	0.601
Huber Loss	5.26±4.44	0.935	9.18±7.69	0.598
DecadeCE Loss	**4.64±3.84**	**0.960**	**7.94±6.54**	**0.641**

Notes: Bolded values indicate the closest estimate. *Italics* indicate proposed method in this study.
Abbreviation: MAE, mean absolute error.

Table 5 Comparison Results for Different Post-Processing

Post-Processing	Mixed-Cohort test			Hold-Out test		
	N_{sample}	MAE	R	N_{sample}	MAE	R
None (by-epoch)	183,086	5.50±5.21	0.933	35,050	9.94±8.60	0.499
Overnight average	848	4.64±3.84	0.960	190	7.94±6.54	0.641
Overnight NN	848	**4.19±3.57**	**0.966**	190	**6.18±5.32**	**0.777**

Notes: Bolded values indicate the closest estimate. Italics indicate proposed method in this study.
Abbreviations: NN, neural network; MAE, mean absolute error.

The average BAA is 0.37 years (SD=3.82), and the paired *t*-test shows the increments of BA and CA is not statistical significance (p=0.43).

For MROS, 56 subjects on the validation and test sets have multiple visits. The mean chronological age difference ΔCA=CA2-CA1 is 6.48 years (SD=0.86), and the mean brain age difference ΔBA=BA2-BA1 is 2.53 years (SD=2.38). The average BAA is −3.95 years (SD=2.70), indicating the increments of BA and CA have a statistical significance (p=0.43, paired *t*-test).

For ABC, 10 subjects on the validation and test sets have multiple visits. The ΔCA intervals between the two visits are both 0.8 years, with a brain age difference of ΔBA=BA2-BA1 of 3.47 years (SD=3.65) for the first interval, and a brain age difference of ΔBA=BA3-BA2 of 0.86 years (SD=3.05) for the second interval. The mean BAA for the two intervals of visits is 2.66 years (SD=3.82, p<0.01 in paired *t*-test) and 0.08 years (SD=3.01, p=0.38 in paired *t*-test), respectively, demonstrating a significant difference between the increments of BA and CA for the first interval, but not for the second interval.

Correlation Between Sleep Stage and Brain Age Index

We investigated the relationship between BA (BAI) and sleep stage domains. As shown in Figure 6A, compared to the overnight averaging, brain age estimation using only the N2 domain is closer to chronological age, whereas on the Wake and REM domains, the model exhibits a decrease in fitting ability.

Additionally, we conducted statistical analyses to investigate changes in BAI estimated by various sleep stages relative to baseline. As shown in Figure 6B, the Wake and light sleep domains provide accelerated aging estimates, and the deep sleep and REM domains provide decelerated aging estimates.

Correlation Between Disease and Brain Age Index

We investigated the statistical differences between BAI in subgroups: NM-OSA (n = 401) vs MS-OSA (n = 201), and PSY (n = 246) vs HC (n = 370), where n represents the number of subjects in each subgroup. The results are presented in Figure 7.

The mean BAI of NM-OSA is 0.5 years higher than that of MS-OSA, no significant difference (P>0.05, *t*-test). Furthermore, the model predicted a MAE of 4.41 years and a BA and CA correlation of 0.92 in the NM-OSA group, whereas in the MS-OSA group, the MAE was larger at 4.84 years and the correlation was smaller at 0.90.

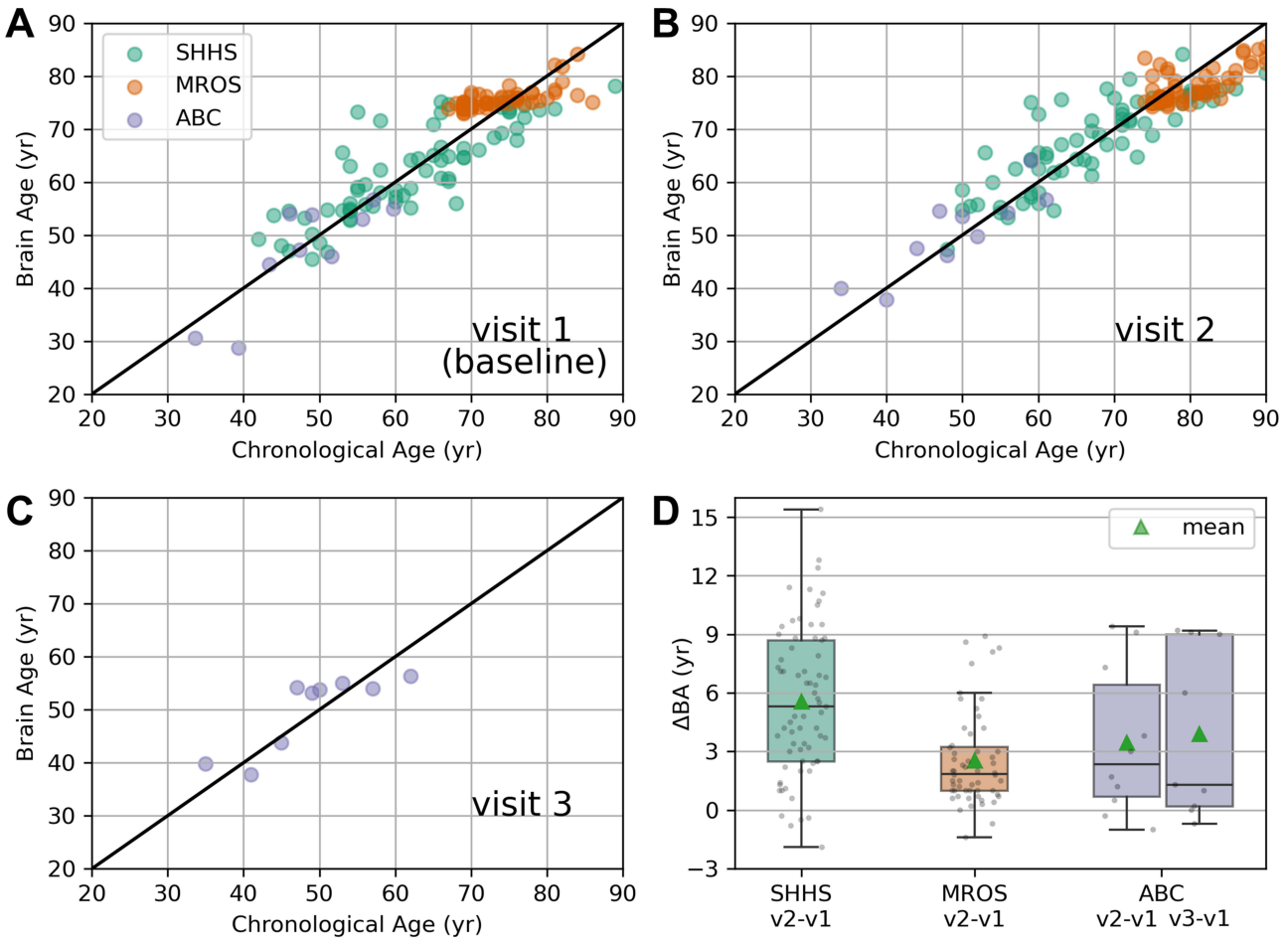

Figure 5 Longitudinal brain age estimates for three cohorts. (**A-C**) BA and CA comparisons at baseline and at two visits for the same subject. (**D**) Distribution of the increase in brain age from baseline at each visit.

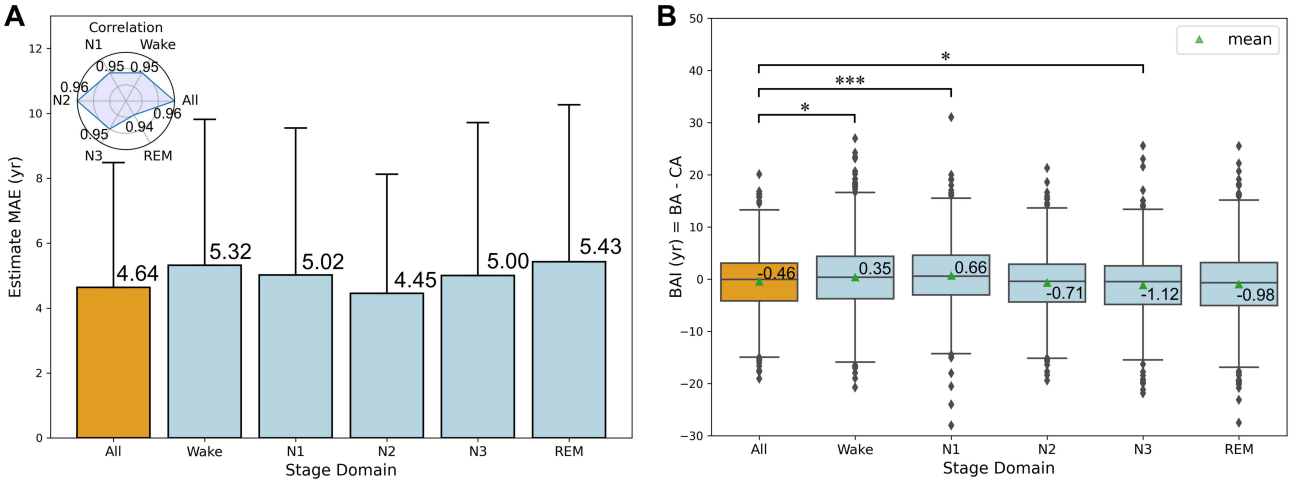

Figure 6 Correlation of sleep stage domains with brain age estimates. The average overnight results for BA prediction by each sleep stage domain are summarized by subject. (**A**) A histogram and radar plot illustrate the performance of BA estimation using each sleep stage domain, indicating the MAE and correlation R values for BA vs CA. (**B**) A box plot illustrates the differences in predicted BAI for each sleep stage domain. (*: p<0.05, ***: p<0.001).

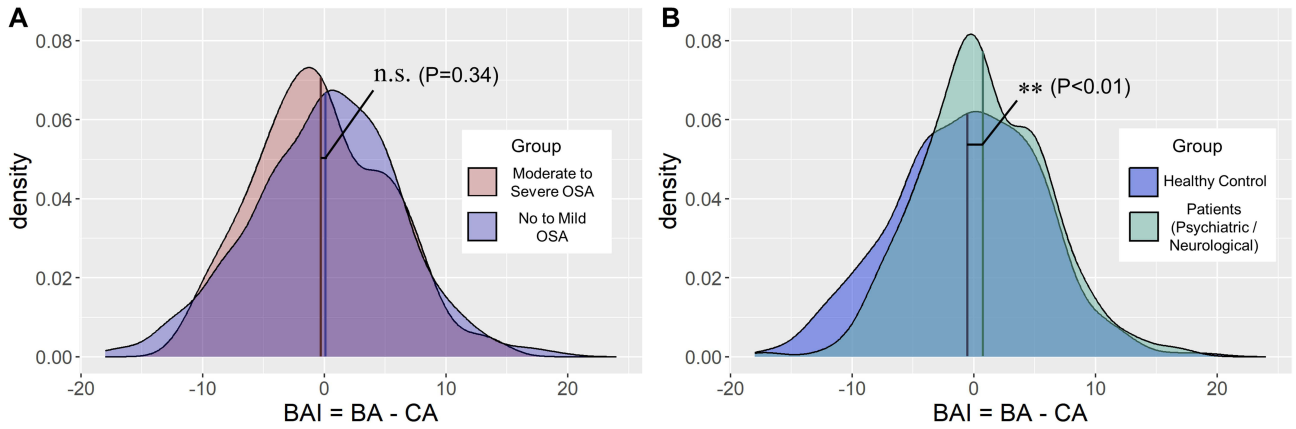

Figure 7 Density distributions of Brain Age Index (BAI) among different groups. (**A**) Comparison of individuals with no to mild obstructive sleep apnea (NM-OSA) and moderate to severe obstructive sleep apnea (MS-OSA), showing a non-significant trend in BAI distribution. (**B**) Comparison of individuals with psychiatric or neurological disorders (PSY) compared to healthy controls (HC) highlights significantly higher BAI in PSY subgroup.

The PSY group demonstrated the mean BAI that exceeded the HC group by 1.27 years (P<0.001, *t*-test). Furthermore, Furthermore, the model predicted a MAE of 4.05 years and the correlation of 0.927 in the HC group, whereas in the PSY group, the MAE was larger at 4.87 years and the correlation was comparable at 0.928.

Discussion

Brain Age from Overnight EEG

This study presents an innovative deep learning-based method for estimating brain age using overnight EEG recordings, capturing changes in brain activity patterns during sleep. Extracting features from sleep EEG is a non-invasive, participant-friendly approach that offers several advantages over both neuroimaging-based and waking EEG-based methods.

Compared to brain age estimation methods developed with structural MRI data, EEG-based methods lack highly intuitive features such as cortical thickness and brain volume.[39] However, EEG-based brain age estimation has several unique advantages. Previous research, such as that by Muehlroth and Werkle-Bergner, has highlighted significant correlations between sleep characteristics (like slow wave power and spindle wave frequency) and age.[40] EEG-based approach has higher sensitivity compared to neuroimaging-based brain age measures and can provide timely feedback for interventions. Moreover, EEG is a more dynamic and cost-effective method that can be applied in various clinical and home monitoring settings.

Compared to using waking resting-state or task-state EEG to analyze brain age, sleep EEG has several distinct advantages: 1) Rich Information: Sleep and aging are closely related, and sleep EEG contains richer potential information revealing brain aging.[41] 2) Natural and Stable Reflection: Sleep EEG signals are acquired in the natural sleep state, providing a more realistic and stable reflection of brain activity.[42] 3) Extended Signal Duration: Sleep EEG typically provides a longer time range of signals, with a large number of EEG epochs helping to train more robust models.

In addition to sleep EEG, sleep structure information also contains indicators of accelerated brain aging. A recent study has accurately modeled sleep dynamics using 3-second mini-epochs and predicted brain aging using overnight sleep features.[43] Another study demonstrated the correlation between sleep structure and mortality.[41]

In this work, we have innovatively utilized both sleep structure information and EEG segments for modeling. By leveraging the powerful representational capabilities of deep neural networks across multiple time scales and the enhanced effects of our multi-flow fusion algorithm, we have achieved a more accurate fit in brain age estimation.

Performance of Proposed Model

The machine learning model proposed in this study is a deep artificial neural network using multi-flow learning, consisting of three key modules: sleep feature extraction module, EEG feature extraction module and fusion prediction module.

For the two feature extractors, which are the two flows of the model, they input data from two complementary modalities: 5-minute epoch information and sleep structure information. A previous study revealed a significantly greater contribution from EEG epoch information compared to sleep structure information (MAE=7.8 years vs MAE=23.3 years).[19] Our ablation experiments corroborated this observation, also validating the contribution of multimodality compared to unimodality to the performance of model assessment. The EEG features were extracted from 5-minute segments. This duration strikes a balance between capturing sufficient information and managing computational demands. It effectively captures both stable and transitional sleep states and is a commonly used configuration in deep learning-based EEG studies.[21,44,45]

For the fusion module, we introduced the gated fusion method to allow sleep structure information to supervise and augment the EEG. We also implemented a two-phase optimization approach inspired by Ref.[21] This approach tackles the computational challenges arising from extensive parameterization in the end-to-end learning process of overnight recordings. The two phases involve epoch prediction, which estimates BA from 5-minute epochs, and post-prediction, which estimates BA from overnight deep features. An ablation experiment confirmed the enhancement of overnight-based post-prediction for brain age estimation.

The publicly available databases exhibit an uneven age distribution, with a noticeable vacancy in the 20–40 age range. This could introduce potential pitfalls in pattern recognition tasks, such as overfitting to oversampled age groups in regression or neglecting the vacancy range in classification. To address these challenges, we introduced the DecadeCE loss, which effectively mitigates these issues. For age vacancy, we applied a Gaussian-based soft cross-entropy loss. This approach enhances the influence of labels within the gap range and avoids the imbalance problem that can arise due to age vacancy undersampling. When converting classification probabilities into continuous brain age (BA) values, we utilized the median probability method. Compared to traditional methods based on maximum probability or weighted average probability, this approach maintains regression accuracy and is more robust in boundary cases. For instance, when there is a vacancy in the predicted ten classes, this method can output BA within the vacancy with a lower probability. In general, the proposed DecadeCE loss demonstrated superior performance compared to the MSE and Huber loss functions, as evidenced by the results of our ablation experiments.

Furthermore, our proposed method achieved compelling results across diverse cohorts. On a mixed cohort dataset, our model attained an MAE of 4.19 years and an R of 0.97, outperforming existing multi-cohort studies.[19,21] Notably, on the unseen HPAP cohort, our method achieved an MAE of 6.18 years and an R of 0.78, showcasing competitive performance compared to benchmarks (MAE=8.16, R=0.78) integrating additional physiological signals like electromyogram and electrocardiogram.[21] These outcomes underscore the adaptability and effectiveness of our approach in varied clinical settings.

Estimated Brain Age

Our proposed model achieved the best predictive performance on mixed-cohorts, with an MAE of 4.19 years and a correlation coefficient of 0.97 for brain age and chronological age, as demonstrated through multiple model comparisons. Our brain age estimation outperformed other EEG-based studies, reducing errors by more than 1 year and reaching the level of neuroimaging accuracy.[10,46,47]

In the longitudinal reliability assessment, our proposed model exhibited distinct trends across three cohorts. In the SHHS cohort, the mean ΔBA from two visits showed no statistically significant difference from ΔCA, indicating a comparable rate of brain aging to natural aging. This is consistent with similar findings from a previous study on brain age estimation using the SHHS with internal training.[19] Conversely, in the MROS cohort, although BA at the second visit showed a significant increase from baseline, ΔBA was significantly smaller than ΔCA. Given that MROS represents the oldest cohort with a concentrated age range, possible explanations include: 1) age boundary effects rendering the model insensitive to brain aging, 2) health status bias due to survivorship bias, or 3) treatment or other physiological factors. The ABC cohort was studied in a smaller sample with low statistical power, and its individuals had high correlations between BA and CA at three visits and excessive ΔBA at the first two visits.

Examining BA estimates across individual sleep stages, we observed that prediction accuracy solely based on N2 sleep outperformed the accuracy obtained from averaging overnight epochs. This discrepancy underscores the

significance of the post-prediction based on neural network models. Stage N2, being a most prevalent sleep stage, extracts a wealth of information in brain age estimation from sleep electroencephalography. An interesting observation regarding the BAI is the association of deep sleep and REM sleep with decelerated aging, while Wake and light sleep correlate with accelerated aging. This aligns intuitively with sleep dynamics, as aging sleep exhibits fewer N3 and REM period EEG connections, while various diseases manifest in sleep fragmentation.[48,49]

We also examined possible associations between sleep disorders and BAI, and between psychiatric or neurological disorders and BAI. As reported in a single cohort study, a mean BAI difference of +2.9 years was found in people with OSA compared to healthy sleepers.[20] However, this association was not found in our work and in another study on multiple cohorts, suggesting an association between AHI and CA and BA, but not BAI.[19]

In accordance with previous reports, our study has demonstrated a positive correlation between psychiatric or neurological disorders and BAI.[2] Our results indicate that individuals in the PSY group exhibited an average of 1.27 years higher BAI than the HC group, a difference that reached statistical significance. Although this discrepancy was less than that observed in a comparable study (+4 years),[19] we attribute this variance to differences in the PSY inclusion criteria and the use of mixed-cohort datasets, which may have refined the model but mitigated this effect to some extent.

Limitations and Future Work

Although the proposed model achieved better BA predictive performance compared to similar studies in independent cohort evaluation, however, the variation between cohorts has a non-negligible impact on accurate BA estimation. To increase the robustness of the model, it is necessary to adopt more diverse PSG recordings from different populations, devices, and acquisition conditions. Furthermore, due to the constraints of utilizing public cohorts, this study lacked sufficient medical information to support a more nuanced brain disease-related assessment.

Conclusion

This study introduces a novel artificial neural network model for predicting brain age using overnight sleep EEG. The model utilizes the one-dimensional Swin Transformer to enable efficient extraction of complex patterns at multiple time scales from sleep EEG signals. This multi-flow network can fuse hypnogram and EEG information to estimate brain age from multiple perspectives of sleep dynamics. To address the challenges posed by the inhomogeneous age distribution in the dataset, we propose the DecadeCE Loss function, which enhances the model's ability to accurately predict brain age across a range of ages.

Trained and evaluated on more than 18,000 overnight PSG recordings, the proposed model demonstrates a robust measurement of brain age, achieving age prediction at the neuroimaging level. The study validates the model's sensitivity to aging through a longitudinal investigation. The estimated BAI showed significant differences between psychiatric or neurological disorders and healthy individuals, demonstrating that diseased individuals have stronger brain aging effects. These achievements demonstrate the potential of EEG in predicting brain age, offering a non-invasive and accessible method for evaluating brain aging. The code for this work has been uploaded on https://github.com/dizhang-xdu/brain-age and is publicly available.

Acknowledgments

We would like to thank the National Sleep Research Resources (https://www.sleepdata.org) team for their work in collecting, organizing the PSG data. The National Sleep Research Resource was supported by the National Heart, Lung, and Blood Institute (R24 HL114473, 75N92019R002).

Regarding the data provider, the Apnea, Bariatric surgery, and CPAP study (ABC Study) was supported by National Institutes of Health grants R01HL106410 and K24HL127307. Philips Respironics donated the CPAP machines and supplies used in the perioperative period for patients undergoing bariatric surgery. The Cleveland Family Study (CFS) was supported by grants from the National Institutes of Health (HL46380, M01 RR00080-39, T32- HL07567, RO1-46380). The Home Positive Airway Pressure study (HomePAP) was supported by the American Sleep Medicine Foundation 38- PM-07 Grant: Portable Monitoring for the Diagnosis and Management of OSA. The Multi-Ethnic Study of Atherosclerosis (MESA) Sleep Ancillary study was funded by NIH-NHLBI Association of Sleep Disorders

with Cardiovascular Health Across Ethnic Groups (RO1 HL098433). MESA is supported by NHLBI funded contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute, and by cooperative agreements UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420 funded by NCATS. The National Heart, Lung, and Blood Institute provided funding for the ancillary MrOS Sleep Study, "Outcomes of Sleep Disorders in Older Men", under the following grant numbers: R01 HL071194, R01 HL070848, R01 HL070847, R01 HL070842, R01 HL070841, R01 HL070837, R01 HL070838, and R01 HL070839. NCH Sleep DataBank was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number R01EB025018. The Sleep Heart Health Study (SHHS) was supported by National Heart, Lung, and Blood Institute cooperative agreements U01HL53916 (University of California, Davis), U01HL53931 (New York University), U01HL53934 (University of Minnesota), U01HL53937 and U01HL64360 (Johns Hopkins University), U01HL53938 (University of Arizona), U01HL53940 (University of Washington), U01HL53941 (Boston University), and U01HL63463 (Case Western Reserve University). This Wisconsin Sleep Cohort Study was supported by the US National Institutes of Health, National Heart, Lung, and Blood Institute (R01HL62252), National Institute on Aging (R01AG036838, R01AG058680), and the National Center for Research Resources (1UL1RR025011).

In accordance with Article 32, Paragraph 1 of the "Ethical Review Measures for Life Sciences and Medical Research Involving Human Beings", as promulgated by the National Health Commission of the People's Republic of China, research using lawfully obtained publicly available data may be exempt from ethical review. All data used in this study were obtained from publicly available data approved from the NSRR, thus exempting it from approval. Please refer to https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm for more information.

Funding

This work was funded by the Innovation Team and Talents Cultivation Program of National Administration of Traditional Chinese Medicine (Grant No. ZYYCXTD-C-202004), the National Key R&D Program of China (Grant No. 2021YFF0306500), the Key Research and Development Program of Shaanxi (Grant No. 2020ZDLSF04-05), Natural Science Basic Research Program of Shaanxi (Grant No. 2022JQ-649), Natural Science Basic Research Program of Shaanxi (2023-JC-YB682).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Cole JH, Leech R, Sharp DJ. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann Neurol*. 2015;77(4):571–581. doi:10.1002/ana.24367
2. Shahab S, Mulsant BH, Levesque ML, et al. Brain structure, cognition, and brain age in schizophrenia, bipolar disorder, and healthy controls. *Neuropsychopharmacology*. 2019;44(5):898–906. doi:10.1038/s41386-018-0298-z
3. Tønnesen S, Kaufmann T, de Lange AMG, et al. Brain Age Prediction Reveals Aberrant Brain White Matter in Schizophrenia and Bipolar Disorder: a Multisample Diffusion Tensor Imaging Study. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2020;5(12):1095–1103. doi:10.1016/j.bpsc.2020.06.014
4. Pardoe HR, Cole JH, Blackmon K, Thesen T, Kuzniecky R. Structural brain changes in medically refractory focal epilepsy resemble premature brain aging. *Epilepsy Res*. 2017;133:28–32. doi:10.1016/j.eplepsyres.2017.03.007
5. Vanni S, Colini Baldeschi A, Zattoni M, Legname G. Brain aging: a Ianus -faced player between health and neurodegeneration. *J Neurosci Res*. 2020;98(2):299–311. doi:10.1002/jnr.24379
6. Han LKM, Dinga R, Hahn T, et al. Brain aging in major depressive disorder: results from the ENIGMA major depressive disorder working group. *Mol Psychiatry*. 2020. doi:10.1038/s41380-020-0754-0
7. Wu W, Zhang Y, Jiang J, et al. An electroencephalographic signature predicts antidepressant response in major depression. *Nat Biotechnol*. 2020;38(4):439–447. doi:10.1038/s41587-019-0397-3
8. Ye E, Sun H, Leone MJ, et al. Association of Sleep Electroencephalography-Based Brain Age Index With Dementia. *JAMA Network Open*. 2020;3(9):e2017357. doi:10.1001/jamanetworkopen.2020.17357
9. de Lange AMG, Anatürk M, Rokicki J, et al. Mind the gap: performance metric evaluation in brain-age prediction. *Hum Brain Mapp*. 2022;43(10):3113–3129. doi:10.1002/hbm.25837

10. Cole JH, Franke K. Predicting Age Using Neuroimaging: innovative Brain Ageing Biomarkers. *Trends Neurosci*. 2017;40(12):681–690. doi:10.1016/j.tins.2017.10.001
11. Leone MJ, Sun H, Boutros CL, et al. HIV increases sleep-based brain age despite antiretroviral therapy. *Sleep*. 2021;44(8):1–32. doi:10.1093/sleep/zsab058
12. Franke K, Ziegler G, Klöppel S, Gaser C. Alzheimer's Disease Neuroimaging Initiative. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *Neuroimage*. 2010;50(3):883–892. doi:10.1016/j.neuroimage.2010.01.005
13. Wang J, Li W, Miao W, Dai D, Hua J, He H. Age estimation using cortical surface pattern combining thickness with curvatures. *Med Biol Eng Comput*. 2014;52(4):331–341. doi:10.1007/s11517-013-1131-9
14. Kondo C, Ito K, Wu K, et al. An age estimation method using brain local features for T1-weighted images. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE; 2015:666–669. doi:10.1109/EMBC.2015.7318450.
15. Jonsson BA, Bjornsdottir G, Thorgeirsson TE, et al. Brain age prediction using deep learning uncovers associated sequence variants. *Nat Commun*. 2019;10(1):1–10. doi:10.1038/s41467-019-13163-9
16. Bashyam VM, Erus G, Doshi J, et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain*. 2020;143(7):2312–2324. doi:10.1093/brain/awaa160
17. Ning K, Duffy BA, Franklin M, et al. Improving brain age estimates with deep learning leads to identification of novel genetic factors associated with brain aging. *Neurobiol Aging*. 2021;105:199–204. doi:10.1016/j.neurobiolaging.2021.03.014
18. Zoubi Al O, Wong CK, Kuplicki RT, et al. Predicting age from brain EEG signals-a machine learning approach. *Front Aging Neurosci*. 2018;10(JUL):1–12. doi:10.3389/fnagi.2018.00184
19. Sun H, Paixao L, Oliva JT, et al. Brain age from the electroencephalogram of sleep. *Neurobiol Aging*. 2019;74:112–120. doi:10.1016/j.neurobiolaging.2018.10.016
20. Yook S, Miao Y, Park C, et al. Predicting brain age based on sleep EEG and DenseNet. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2021:245–248. doi:10.1109/EMBC46164.2021.9631064.
21. Brink-Kjaer A, Leary EB, Sun H, et al. Age estimation from sleep studies using deep learning predicts life expectancy. *Npj Digit Med*. 2022;5(1). doi:10.1038/s41746-022-00630-9
22. Bakker JP, Tavakkoli A, Rueschman M, et al. Gastric Banding Surgery versus Continuous Positive Airway Pressure for Obstructive Sleep Apnea: a Randomized Controlled Trial. *Am J Respir Crit Care Med*. 2018;197(8):1080–1083. doi:10.1164/rccm.201708-1637LE
23. Redline S, Tishler PV, Tosteson TD, et al. The Familial Aggregation of Obstructive Sleep Apnea. *Am J Respir Crit Care Med*. 1995;151(3_pt_1):682–687. doi:10.1164/ajrccm/151.3_pt_1.682
24. Rosen CL, Auckley D, Benca R, et al. A Multisite Randomized Trial of Portable Sleep Studies and Positive Airway Pressure Autotitration Versus Laboratory-Based Polysomnography for the Diagnosis and Treatment of Obstructive Sleep Apnea: the HomePAP Study. *Sleep*. 2012;35(6):757–767. doi:10.5665/sleep.1870
25. Chen X, Wang R, Zee P, et al. Racial/Ethnic Differences in Sleep Disturbances: the Multi-Ethnic Study of Atherosclerosis (Mesa). *Sleep*. 2015. doi:10.5665/sleep.4732
26. Song Y, Blackwell T, Yaffe K, Ancoli-Israel S, Redline S, Stone KL. Relationships Between Sleep Stages and Changes in Cognitive Function in Older Men: the MrOS Sleep Study. *Sleep*. 2015;38(3):411–421. doi:10.5665/sleep.4500
27. Lee H, Li B, DeForte S, et al. A large collection of real-world pediatric sleep studies. *Sci Data*. 2022;9(1):421. doi:10.1038/s41597-022-01545-6
28. Quan SF, Howard BV, Iber C, et al. The Sleep Heart Health Study: design, rationale, and methods. *Sleep*. 1997;20(12):1077–1085.
29. Moore H, Leary E, Lee SY, et al. Design and Validation of a Periodic Leg Movement Detector. *PLoS One*. 2014;9(12):e114565. doi:10.1371/journal.pone.0114565
30. Young T, Palta M, Dempsey J, Peppard PE, Nieto FJ, Hla KM. Burden of sleep apnea: rationale, design, and major findings of the Wisconsin Sleep Cohort study. *WMJ*. 2009;108(5):246–249.
31. Zhang GQ, Cui L, Mueller R, et al. The National Sleep Research Resource: towards a sleep data commons. *J Am Med Inf Assoc*. 2018;25(10):1351–1358. doi:10.1093/jamia/ocy064
32. Perslev M, Darkner S, Kempfner L, Nikolic M, Jennum PJ, Igel C. U-Sleep: resilient high-frequency sleep staging. *Npj Digit Med*. 2021;4(1):72. doi:10.1038/s41746-021-00440-5
33. Zhang D, Sun J, She Y, et al. A two-branch trade-off neural network for balanced scoring sleep stages on multiple cohorts. *Front Neurosci*. 2023:17. doi:10.3389/fnins.2023.1176551
34. Stephansen JB, Olesen AN, Olsen M, et al. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat Commun*. 2018;9(1):1–15. doi:10.1038/s41467-018-07229-3
35. Kiela D, Grave E, Joulin A, Mikolov T. Efficient Large-Scale Multi-Modal Classification. *Proc AAAI Conf Artif Intell*. 2018;32(1):5198–5204. doi:10.1609/aaai.v32i1.11945
36. Loshchilov I, Hutter F Decoupled Weight Decay Regularization; 2017. Available from: http://arxiv.org/abs/1711.05101. Accessed June 26, 2024.
37. Supratak A, Guo Y. TinySleepNet: an Efficient Deep Learning Model for Sleep Stage Scoring based on Raw Single-Channel EEG. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2020:641–644. doi:10.1109/EMBC44109.2020.9176741.
38. Pradeepkumar J, Anandakumar M, Kugathasan V, et al. Towards Interpretable Sleep Stage Classification Using Cross-Modal Transformers. *arXiv preprint arXiv*. 2022;1–12.
39. Cole JH, Franke K, Cherbuin N. Quantification of the Biological Age of the Brain Using Neuroimaging. *Biomarkers Human Aging*. 2019:293–328. doi:10.1007/978-3-030-24970-0_19
40. Muehlroth BE, Werkle-Bergner M. Understanding the interplay of sleep and aging: methodological challenges. *Psychophysiology*. 2020;57(3). doi:10.1111/psyp.13523
41. Shi H, Huang T, Ma Y, Eliassen AH, Sun Q, Wang M. Sleep duration and snoring at midlife in relation to healthy aging in women 70 years of age or older. *Nat Sci Sleep*. 2021;13:411–422. doi:10.2147/NSS.S302452
42. Adra N, Dümmer LW, Paixao L, et al. Decoding information about cognitive health from the brainwaves of sleep. *Sci Rep*. 2023;13(1):11448. doi:10.1038/s41598-023-37128-7

43. Cesari M, Stefani A, Mitterling T, Frauscher B, Schönwald SV, Högl B. Sleep modelled as a continuous and dynamic process predicts healthy ageing better than traditional sleep scoring. *Sleep Med*. 2021;77:136–146. doi:10.1016/j.sleep.2020.11.033
44. Seo H, Back S, Lee S, Park D, Kim T, Lee K. Intra- and inter-epoch temporal context network (IITNet) using sub-epoch features for automatic sleep scoring on raw single-channel EEG. *Biomed Signal Process Control*. 2020;61:102037. doi:10.1016/j.bspc.2020.102037
45. Olesen AN, Mignot E, Bjarup H, Sorensen D, Alto P, Olesen AN. Automatic sleep stage classification with deep residual networks in a mixed-cohort setting. *Sleep*. 2020.
46. Marx GA, Kauffman J, McKenzie AT, et al. Histopathologic brain age estimation via multiple instance learning. *Acta Neuropathol*. 2023;146 (6):785–802. doi:10.1007/s00401-023-02636-3
47. Jusseaume K, Valova I. Brain Age Prediction/Classification through Recurrent Deep Learning with Electroencephalogram Recordings of Seizure Subjects. *Sensors*. 2022;22(21). doi:10.3390/s22218112
48. Suzuki Y, Abe T, Kawana F. Instability of scoring stage N1 is a factor that reduces repeatability of human sleep staging. *Sleep Med*. 2019;64:S369. doi:10.1016/j.sleep.2019.11.1029
49. Bouchard M, Lina JM, Gaudreault PO, Dubé J, Gosselin N, Carrier J. EEG connectivity across sleep cycles and age. *Sleep*. 2019. doi:10.1093/sleep/zsz236