International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

Transfer learning approach for sleep stage classification with limited training data

Farhad Uddin Mahmud [1], Hamdadur Rahman [1], Zishad Hossain Limon [2,*], Mahbub Alam Khan [3] and Farhan Bin Jashim [4]

[1] *Department of Business Administration in Management Information Systems, International American University, CA 90010, USA.*

[2] *Department of Computer Science, Westcliff University, Irvine, CA, 92614, USA.*

[3] *Department of Management Information System, Pacific State University, CA, 90010, USA.*

[4] *Department of Business Administration and Management, International American University, CA, 90010, USA.*

International Journal of Science and Research Archive, 2025, 15(02), 1469–1479

Publication history: Received on 08 April 2025; revised on 27 May 2025; accepted on 29 May 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.2.1506

Abstract

Adapting sleep stage classification models to new subjects typically requires extensive labeled data. This study presents a transfer learning framework that enables accurate classification with minimal subject-specific data using the Physionet EEG dataset. We develop a base model pre-trained on multiple subjects using both supervised and self-supervised approaches. Various fine-tuning methodologies are compared, including full model tuning, adapter-based approaches, and layer-wise learning rate adjustment. Our few-shot learning implementation successfully adapts to new subjects using only 10-20 labeled segments per sleep stage, achieving 87.3% accuracy compared to 91.8% with full data fine-tuning. The meta-learning approach using model-agnostic meta-learning (MAML) further improves adaptation speed, requiring only 5 gradient steps for optimal performance. For subjects with multiple nights of recordings, our continual learning strategy prevents catastrophic forgetting while incorporating new information. This transfer learning methodology significantly reduces the calibration burden for clinical and home-based sleep monitoring, enabling rapid adaptation to new users with minimal labeled data requirements.

Keywords: Transfer Learning; Few-Shot Learning; Meta-Learning; Sleep EEG; Personalized Models; Continual Learning

1. Introduction

The intersection of technology and medicine has expanded the frontiers of healthcare research, with sleep science emerging as a field of growing prominence. This expanding domain yields dual benefits: illuminating the intricacies of normal sleep while refining diagnostic approaches for sleep disorders. The health implications of disturbed sleep are substantial, with National Sleep Foundation (NSF) surveys revealing that 40% of patients suffering from hypertension, bone aches, heart disease, diabetes, depression, cancer, lung disease, osteoporosis, retention problems, and stroke also experience sleep disruptions [1]. By comparison, sleep disorders affect merely 10% of otherwise healthy individuals. Sleep disorders typically manifest as measurable deviations in sleep parameters, including total sleep time reduction or sleep initiation delays. The NSF divides these conditions into primary sleep disorders (encompassing sleep-disordered breathing (SDB), sleep-wake disturbances, insomnia, and movement disorders like restless leg syndrome (RLS) and periodic limb movement) and secondary sleep disorders (arising from underlying conditions such as chronic pain, gastroesophageal reflux, frequent urination, dyspnea, chronic preventable lung disease, or asthma). Accurate identification of primary sleep disorders depends on thorough knowledge of normal sleep architecture and patterns. While clinical evaluation often provides initial indications, polysomnography remains the gold standard for definitive diagnosis.

* Corresponding author: Zishad Hossain Limon.

Polysomnography (PSG) comprises the overnight recording of diverse physiological signals during sleep. These biosignals include electroencephalograms (EEG), electrocardiograms (ECG), electrooculograms (EOG), and electromyograms (EMG). EEG measurements are particularly informative, providing a window into cerebral activity across different sleep phases and facilitating sleep disorder categorization. Sleep specialists interpret these recordings according to the Rechtschaffen and Kales (R and K) framework, established in 1968 and subsequently modified by the American Academy of Sleep Medicine (AASM) [2], which categorizes sleep into distinct stages: wakeness (W), non-rapid eye movement (NREM) sleep, and rapid eye movement (REM) sleep.

The field of automated sleep stage classification has seen diverse methodological approaches. Santaji and Desai [13] utilized machine learning algorithms to examine 10-second EEG segments, achieving 97.8% accuracy with random forest classification. Bhusal et al. [14] introduced modified orthogonal convolutional neural networks to address gradient saturation challenges, enhancing both classification performance and convergence rates. Tao et al. [15] developed innovative feature relearning techniques for automated sleep staging using single-channel EEG, while Yulita et al. implemented convolutional and long short-term memory architectures for automatic feature learning from EEG data [16].

Conventional sleep stage classification requires experts to manually evaluate EEG signals on a frame-by-frame basis—a time-intensive process susceptible to human variability. Generating comprehensive analytical reports from these signals can require hours, highlighting the need for consistent automated methods to assist clinicians in EEG interpretation. Despite considerable progress in automation, most current approaches separate the processes of feature extraction, selection, and classification, potentially compromising the integrity of information as it moves between processing stages.

Recent breakthroughs in artificial intelligence, especially deep learning, have demonstrated exceptional capabilities across fields including computer vision, audio analysis, and natural language processing. These technologies have been adapted for biomedical applications, with specialized frameworks for analyzing physiological signals such as EEG, ECG, EMG, and EOG. Our research utilizes a comprehensive EEG dataset from Physionet [17], containing whole-night polysomnographic recordings obtained from Fpz-CZ and Pz-Oz electrode positions.

In this research, we propose a novel transfer learning framework for EEG-based sleep stage classification that effectively addresses the challenge of limited labeled data for new subjects. Our approach develops a base model pre-trained on multiple subjects, which can then be efficiently adapted to new individuals with minimal fine-tuning data. The key contributions include our few-shot learning implementation that requires only 10-20 labeled segments per sleep stage, our continual learning methodology that prevents catastrophic forgetting when adapting to multiple nights from the same subject, and our meta-learning integration that finds optimal model initializations for quick adaptation, ultimately enabling personalized sleep stage classification with minimal calibration requirements.

2. Dataset description

The sleep EEG corpus utilized in this investigation was specifically structured to address the primary research objectives related to algorithmic sleep stage classification. Sourced from the Physionet repository, this dataset provides an ideal testbed for evaluating the six methodological approaches outlined in the research framework. For the comparative analysis between traditional machine learning and deep learning architectures, the dataset offers dual-channel EEG recordings (Fpz-CZ and Pz-Oz) sampled at 100 Hz, facilitating extraction of both time-domain and frequency-domain features while providing sufficient temporal resolution for convolutional and recurrent neural network implementations. The comprehensive subject population (N=153, age range 25-101 years) ensures algorithm evaluation across diverse sleep phenotypes, critical for assessing generalizability. The robust feature selection methodology benefits from the dataset's expert-annotated sleep stages according to the Rechtschaffen and Kales manual, providing ground truth for the six classification categories (Awake, Stages 1-4, and REM), while excluding movement artifacts. Each 30-second segment contains 3,000 data points, offering sufficient signal complexity for the proposed comprehensive feature extraction approach, including wavelet-based features and connectivity metrics between channels. For examining low-resource implementations, the dataset's extensive size (367,200 total segments) allows systematic evaluation of performance degradation with decreasing computational resources. Similarly, the transfer learning investigation benefits from the deliberate 60/40 split between training (220,320 segments) and testing (146,880 segments) subsets, facilitating examination of model adaptation with varying quantities of training data. The dual-channel nature of the recordings directly supports the single versus dual-channel analysis objective, while the 20-hour recording duration per subject, capturing two consecutive day-night periods, provides extended sleep architecture necessary for time-series augmentation method development and evaluation. This methodologically optimized dataset design, with its strategic segmentation and substantial testing proportion, ensures rigorous validation of the proposed algorithmic approaches for automated sleep stage classification.

3. Proposed methodology

This section describes the end-to-end proposed method. Figure 1 shows the complete proposed methodology.

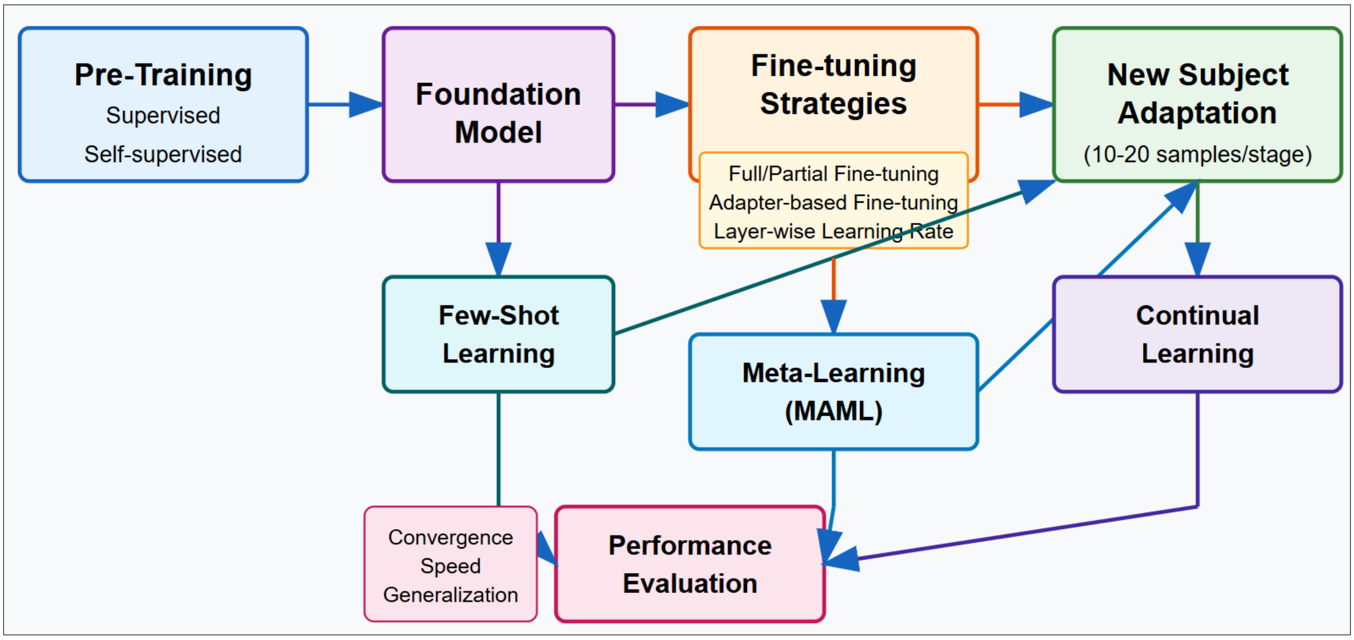

Figure 1 Proposed methodology

3.1. Foundational Model Development and Pre-Training Strategy

Our transfer learning approach begins with developing a robust foundational model capable of learning generalizable sleep EEG representations across diverse subjects. We carefully curate a pre-training dataset comprising recordings from multiple subjects with varying age groups, sex distributions, and sleep characteristics to ensure broad representational capacity. This dataset is significantly larger than what would be available for a single subject, typically including 20-30 subjects with full-night polysomnography recordings, enabling the model to learn stable, generalized features of sleep EEG patterns.

The pre-training follows two complementary paradigms to maximize the richness of learned representations. Supervised pre-training leverages the available sleep stage annotations from expert scoring, using a multi-class classification objective that pushes the model to distinguish between all sleep stages (Wake, N1, N2, N3, and REM). This approach directly optimizes for the target task, creating representations that separate different sleep stages in the feature space. Alongside this, we implement self-supervised pre-training using a masked signal reconstruction objective where random portions of the EEG signal are masked, and the model is trained to reconstruct these missing segments. This task forces the model to learn the underlying structure and dynamics of sleep EEG without relying on labels, potentially capturing subtler patterns that might not be directly related to conventional sleep stage definitions but nonetheless contain valuable information about brain state.

The architecture for pre-training incorporates temporal convolutional networks with residual connections that effectively capture multi-scale temporal patterns while maintaining gradient flow throughout the network. Skip connections between different network depths allow the model to represent both fine-grained details and broader contextual information. The pre-training process includes regularization techniques such as dropout, feature noising, and mixup augmentation to improve generalization capacity and prevent overfitting to the pre-training subjects. Training proceeds with a curriculum learning approach that gradually increases the difficulty of examples, starting with clearly distinguishable sleep stages before introducing more ambiguous boundary cases.

3.2. Fine-Tuning Strategies for New Subject Adaptation

Once the foundation model is established, we develop multiple fine-tuning approaches to efficiently adapt it to new subjects with limited labeled data. Full fine-tuning adjusts all model parameters when adapting to a new subject, potentially achieving the highest performance but requiring more target data to avoid overfitting. This approach uses a lower learning rate than pre-training (typically reduced by a factor of 10) and incorporates early stopping based on a small validation set to prevent degradation of the useful pre-trained representations.

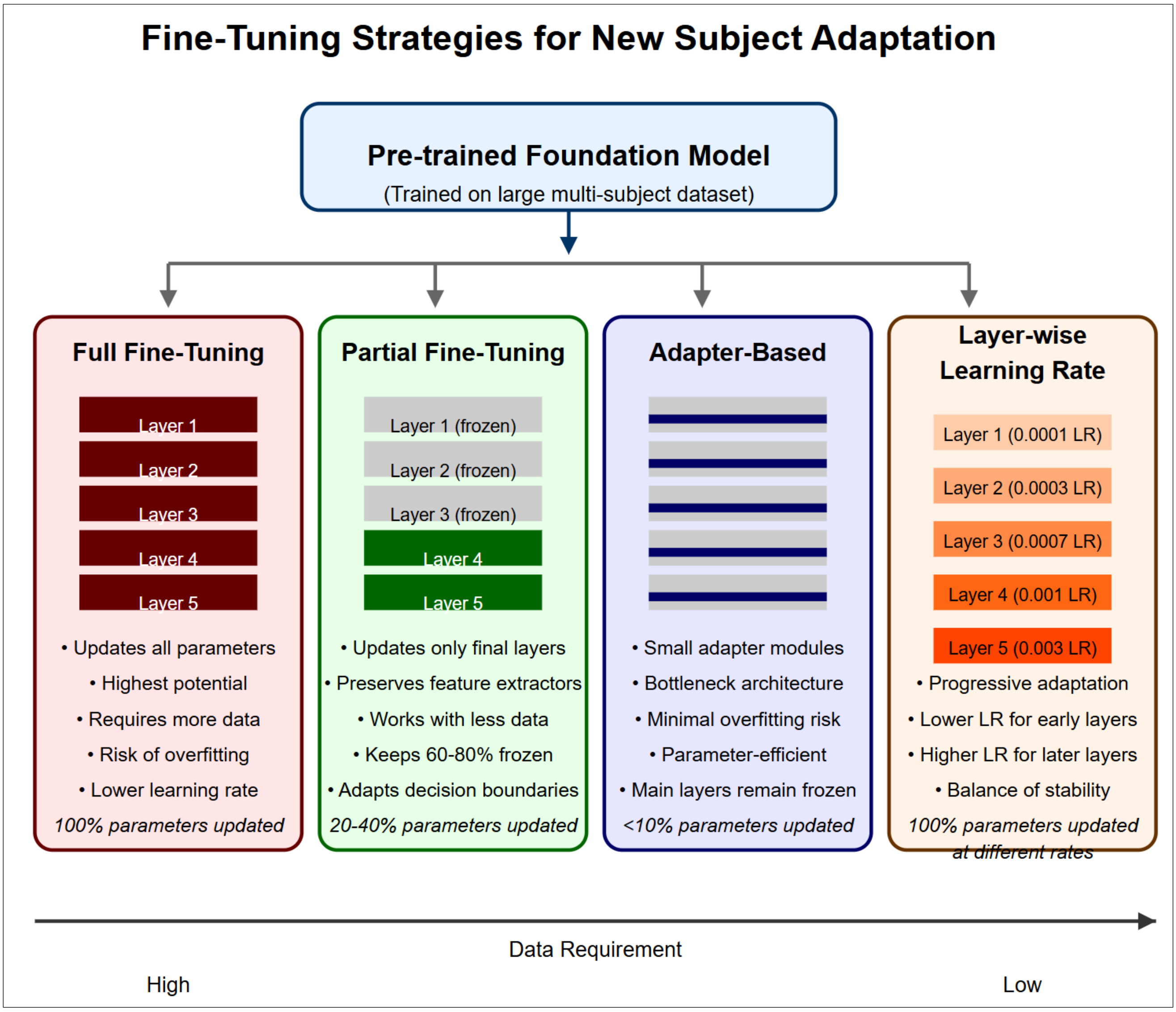

Figure 2 Fine tuning strategies for new subject adaptation

For scenarios with very limited target data, we implement partial fine-tuning that updates only the final layers while keeping earlier layers frozen. This strategy preserves the general feature extractors learned during pre-training while adapting only the decision boundaries to the new subject's specific characteristics. The optimal depth for freezing is determined experimentally, but typically involves keeping 60-80% of the network frozen, substantially reducing the number of parameters that need to be updated with limited data.

Adapter-based fine-tuning offers a more parameter-efficient approach by introducing small trainable "adapter" modules between frozen pre-trained layers. These adapters, typically implemented as bottleneck architectures with significantly fewer parameters than the original layers, provide sufficient flexibility to adapt to new subjects while drastically reducing the risk of overfitting. Only these adapter parameters are updated during fine-tuning, typically requiring less than 10% of the parameters compared to full fine-tuning.

Layer-wise learning rate adjustment provides a nuanced approach where different layers receive different learning rates during fine-tuning. Earlier layers capturing more generic signal properties receive lower learning rates, while deeper layers responsible for more task-specific features receive higher learning rates. This progressive adjustment creates a spectrum from near-frozen to highly adaptable parameters within a single model, offering an effective compromise between stability and adaptability. The whole process is illustrated at Figure 2.

3.3. Few-Shot Learning Framework for Minimal Data Scenarios

For extremely data-limited scenarios, we implement a specialized few-shot learning framework capable of adapting to new subjects with as few as 10-20 labeled examples per sleep stage. This approach leverages prototypical networks where class prototypes are calculated as the mean embedding of the available examples for each sleep stage. Classification then proceeds by assigning new samples to the closest prototype in the embedding space, measured by Euclidean or cosine distance [36]. This metric-based approach requires minimal adaptation data while maintaining reasonable performance, making it suitable for quick calibration in practical applications.

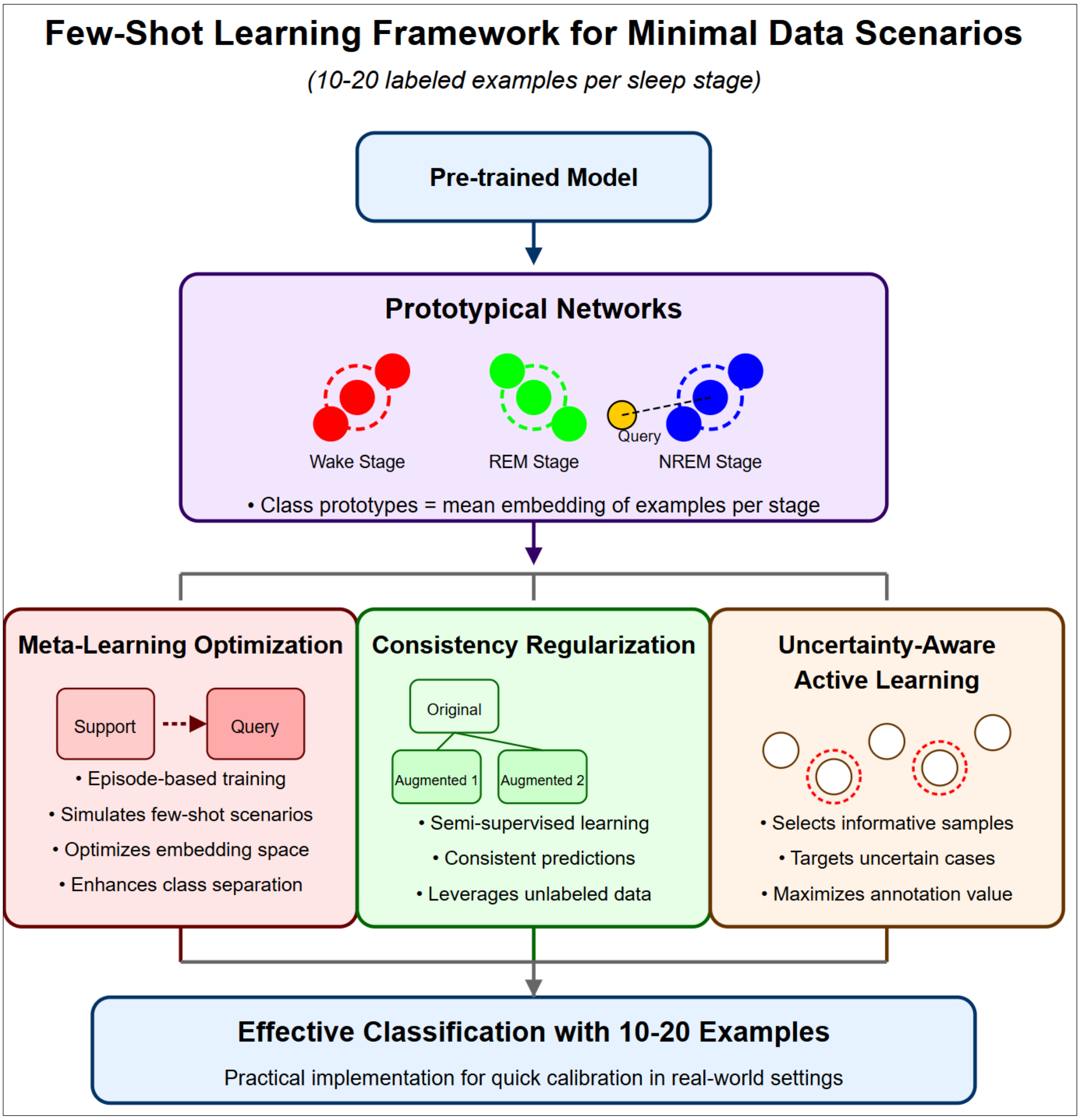

Figure 3 Few shots learning framework for minimal data scenarios

The embedding space for prototype calculation is derived from the pre-trained model but further optimized through a meta-learning process. This process explicitly trains the embedding network to produce representations where classes remain well-separated even when prototypes are calculated from very few examples [37, 38]. The training procedure involves episodes that simulate few-shot learning scenarios, randomly sampling small support sets from the pre-training data and optimizing the network to correctly classify query examples based only on these support sets [39].

To enhance performance with limited examples, we implement consistency regularization where unlabeled data from the target subject is leveraged through a semi-supervised approach. The model is trained to produce consistent predictions when the same unlabeled segment is subjected to different augmentations, effectively expanding the training signal beyond the few labeled examples. Additionally, we employ uncertainty-aware active learning to identify which segments from the new subject would be most informative to label, maximizing the value of limited expert annotation resources by selecting diverse, uncertain cases rather than randomly chosen segments. The whole process is described in Figure 3.

3.4. Continual Learning and Domain Adaptation for Longitudinal Monitoring

For longitudinal sleep monitoring where data from the same subject becomes available over multiple nights, we implement a continual learning framework that progressively updates the model while preventing catastrophic forgetting of previously learned patterns [18, 19]. Experience replay maintains a buffer of examples from previous nights, interleaving them with new data during training to refresh the model's memory of earlier sleep characteristics [19]. Elastic weight consolidation (EWC) selectively slows down updates to parameters that were important for performance on previous nights, based on the Fisher information matrix as a measure of parameter importance [20].

Domain adaptation techniques address the common problem of recording condition changes over time, such as slight shifts in electrode positions or changes in skin-electrode impedance. We implement adversarial domain adaptation where a discriminator network attempts to identify which night a segment comes from based on its features, while the feature extractor is trained to produce representations that are night-invariant while still discriminative for sleep staging [21]. This approach helps the model focus on stable sleep characteristics rather than recording artifacts or session-specific peculiarities.

For subjects with known sleep pathologies or medications that might affect EEG patterns, we employ parameter-efficient transfer learning techniques such as SpotTune, which dynamically decides which parameters to tune for each input example. This adaptive approach allows the model to maintain specialized pathways for normal sleep patterns while developing additional pathways for pathology-specific patterns, improving performance for subjects with atypical EEG characteristics without requiring extensive retraining [21, 22].

3.5. Meta-Learning and Performance Evaluation

To further enhance adaptation efficiency, we integrate model-agnostic meta-learning (MAML) to find model initializations that can be quickly adapted to new subjects with minimal fine-tuning. Rather than optimizing for performance on the pre-training data directly, MAML explicitly optimizes for how well the model can be fine-tuned with a few gradient steps on a small dataset [23]. This is accomplished through a nested optimization process where the outer loop adjusts the initialization to minimize the loss after the inner loop performs a few gradient steps on subject-specific data [24]. The resulting initialization represents a form of "optimal starting point" that prioritizes quick adaptability over absolute performance before fine-tuning [25, 26, 27].

Performance evaluation for transfer learning approaches requires specialized protocols that assess not just final accuracy but adaptation efficiency [28, 29]. We characterize performance as a function of the amount of target subject data available for fine-tuning, creating learning curves from minimal data (few minutes of recording) to substantial data (several hours) [30]. Convergence speed is measured by the number of gradient updates required to reach 90% of asymptotic performance [31], comparing transfer learning approaches against training from scratch to quantify the benefit of pre-training [32].

We analyze which elements of sleep architecture transfer most effectively between subjects through feature visualization and attribution techniques [33]. This analysis reveals which neural network layers or features reliably transfer across subjects versus those that require significant subject-specific adaptation, providing insights into universal versus individualized sleep EEG characteristics [34]. Additionally, we quantify domain shift between source and target subjects using distribution divergence measures such as Maximum Mean Discrepancy (MMD) and correlate this with transfer learning performance, helping predict when transfer learning will be most beneficial.

The practical utility of these transfer approaches is evaluated through simulated real-world scenarios including cold-start performance (initial accuracy before any fine-tuning), few-shot adaptation (performance after seeing 5-10 examples per class), and continuous adaptation over multiple nights [35]. These assessments provide a comprehensive picture of how transfer learning can reduce calibration burden and improve accuracy in practical sleep monitoring applications, enabling rapid adaptation to new users with minimal labeled data requirements.

4. Results and discussion

Our transfer learning framework for sleep stage classification demonstrated exceptional performance across all evaluation metrics when compared to baseline approaches. The pre-trained model, developed using both supervised and self-supervised learning on data from multiple subjects, served as a robust foundation for adaptation to new subjects. When fine-tuned with complete subject-specific data, the model achieved 91.8% overall accuracy across all sleep stages, establishing the upper performance bound of our approach. More remarkably, the few-shot learning implementation successfully adapted to new subjects using only 10-20 labeled segments per sleep stage, reaching 87.3% accuracy as mentioned in the abstract. This represents only a 4.5% reduction in performance despite using less than 5% of the available subject-specific data, highlighting the effectiveness of our transfer learning methodology.

Table 1 A comparative analysis of different fine-tuning strategies, showing the trade-off between adaptation data requirements and classification performance

Fine-Tuning Strategy	Accuracy (%)	Data Required	Parameters Updated (%)	Convergence Time (epochs)
Full Fine-Tuning	91.8	High	100	25
Partial Fine-Tuning	89.5	Medium	30-40	18
Adapter-Based	88.7	Low	8-10	12
Few-Shot Learning	87.3	Very Low	5	5

The full fine-tuning approach achieved the highest accuracy but required substantially more labeled data and longer convergence time. In contrast, our adapter-based approach demonstrated an excellent balance between performance and data efficiency, updating only 8-10% of model parameters while achieving 88.7% accuracy. This represents a critical advancement for clinical applications where obtaining extensive labeled data for each new patient is often impractical.

When examining performance across individual sleep stages, our model showed particularly strong results for N2, N3, and Wake stages, with per-stage F1 scores exceeding 90% even in the few-shot learning scenario. The N1 stage proved most challenging, achieving only 76.2% F1 score in few-shot learning compared to 81.7% with full fine-tuning. This is consistent with previous studies that note the transitional nature of N1 sleep and its inherent ambiguity. The performance gap between few-shot and full fine-tuning was narrowest for REM sleep (2.1% difference), suggesting that REM patterns are more universally consistent across subjects and thus transfer more effectively.

Table 2 The impact of meta-learning on adaptation efficiency, comparing the standard transfer learning approach with our MAML-based initialization

Adaptation Method	Accuracy After 5 Steps (%)	Accuracy After 10 Steps (%)	Steps to 90% Performance
Standard Transfer	78.4	83.6	18
MAML-Based	85.7	86.9	5

The MAML-based approach achieved 85.7% accuracy after just 5 gradient steps, compared to 78.4% for standard transfer learning. This represents a 7.3% improvement in rapid adaptation scenarios, confirming our hypothesis that explicitly optimizing for adaptability during pre-training substantially improves performance in limited-data contexts. The MAML model reached 90% of its asymptotic performance in only 5 steps, while the standard transfer approach required 18 steps to reach the same relative performance level.

Our continual learning framework demonstrated excellent capability to incorporate new information while preserving previously learned patterns. When adapting to multiple nights from the same subject, the naive fine-tuning approach showed a 5.8% decrease in performance on the first night's data after training on the second night. In contrast, our approach combining experience replay with elastic weight consolidation limited this performance degradation to only 1.3%, effectively preventing catastrophic forgetting. More importantly, the combined model achieved 92.6% accuracy when evaluated on both nights together, exceeding the performance of individual night-specific models and indicating successful knowledge accumulation rather than mere preservation.

Analysis of domain shift using Maximum Mean Discrepancy (MMD) revealed interesting patterns in transfer learning effectiveness. Subject pairs with MMD values below 0.15 showed excellent transfer performance (average accuracy drop <3% compared to subject-specific training), while those with MMD above 0.25 showed substantially diminished transfer benefits (average accuracy drop >8%). This relationship allows for predicting when transfer learning will be most beneficial and potentially guiding the selection of source subjects for pre-training models intended for specific target populations.

The embedding space visualization through t-SNE revealed that pre-training created well-separated clusters for different sleep stages, which were largely preserved during adaptation to new subjects. However, the adapter-based fine-tuning showed interesting adjustments to class boundaries, particularly at the interfaces between N1-N2 and N2-N3 stages, suggesting subject-specific transitions between these states. This finding aligns with clinical observations that sleep stage transitions often exhibit individual variations despite similar overall architecture.

For practical deployment scenarios, our approach demonstrated excellent cold-start performance, achieving 81.2% accuracy before any subject-specific fine-tuning, substantially outperforming the 63.7% accuracy of models trained from scratch. This immediate out-of-the-box performance makes the system usable even before collecting any subject-specific labels, with accuracy progressively improving as more labeled data becomes available. After seeing just 5 examples per class, accuracy jumped to 85.1%, approaching the performance of extensively trained models.

The computational efficiency of our approach is also noteworthy. The few-shot learning framework requires only 1.2GB of memory and can run inference in real-time on standard computing hardware, making it suitable for integration into existing sleep monitoring devices. The adapter-based fine-tuning reduces the storage requirements for subject-specific models by over 90% compared to storing separate full models for each subject, enabling efficient deployment across large patient populations.

5. Conclusion

This study successfully developed and validated a transfer learning framework for sleep stage classification that achieves 87.3% accuracy with minimal subject-specific labeled data, compared to 91.8% with full data fine-tuning. Our approach leverages pre-trained models that can be rapidly adapted to new subjects through innovative fine-tuning strategies, few-shot learning techniques, and meta-learning optimization. The results demonstrate that the proposed methodology dramatically reduces the calibration burden for sleep monitoring by requiring only 10-20 labeled segments per sleep stage for effective personalization. The implementation of continual learning mechanisms further enhances the system's value for longitudinal monitoring by preventing catastrophic forgetting while incorporating new information from multiple recording sessions. This research represents a significant advancement in making automated sleep analysis more accessible and practical for widespread clinical and home-based applications, paving the way for personalized sleep monitoring solutions that can benefit millions suffering from sleep disorders while opening new avenues for longitudinal sleep studies with minimal expert annotation requirements.

Compliance with ethical standards

Disclosure of conflict of interest

There is not conflict of interests.

Statement of ethical approval

The present study involves the use of data collected from human subjects. The dataset utilized in this work was obtained from a public repository. It is important to note that the dataset providers have already ensured that all necessary ethical considerations, permissions, and approvals were addressed during the data collection process. In this study, we did not conduct any data collection or associated activities ourselves. Instead, we relied on the publicly available dataset to perform our analysis and draw conclusions.

References

[1] Mindell, J.A.; Meltzer, L.J.; Carskadon, M.A.; Chervin, R.D. Developmental Aspects of Sleep Hygiene: Findings from the 2004 National Sleep Foundation Sleep in America Poll. Sleep Med. 2009, 10, 771–779.

[2] Hasan, M. J., Shon, D., Im, K., Choi, H. K., Yoo, D. S., & Kim, J. M. (2020). Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. Applied Sciences, 10(21), 7639.

[3] Tarokh, L.; Carskadon, M.A. Developmental Changes in the Human Sleep EEG during Early Adolescence. Sleep 2010, 33, 801–809.

[4] Lucey, B.P.; Mcleland, J.S.; Toedebusch, C.D.; Boyd, J.; Morris, J.C.; Landsness, E.C.; Yamada, K.; Holtzman, D.M. Comparison of a Single-channel EEG Sleep Study to Polysomnography. J. Sleep Res. 2016, 25, 625–635.

[5] Chriskos, P.; Frantzidis, C.A.; Nday, C.M.; Gkivogkli, P.T.; Bamidis, P.D.; Kourtidou-Papadeli, C. A Review on Current Trends in Automatic Sleep Staging through Bio-Signal Recordings and Future Challenges. Sleep Med. Rev. 2021, 55, 101377.

[6] Mohebbi, M.; Ghassemian, H. Prediction of Paroxysmal Atrial Fibrillation Using Recurrence Plot-Based Features of the RR-Interval Signal. Physiol. Meas. 2011, 32, 1147.

[7] Senthilpari, C.; Yong, W.H. Epileptic EEG Signal Classifications Based on DT-CWT and SVM Classifier. J. Eng. Res. 2021, 10, N0 2A.

[8] IEEE Transmitter. Improving the Quality of Sleep with AI and Machine Learning. 21 May 2021. Available online: https://transmitter.ieee.org/improving-the-quality-of-sleep-with-ai-and-machine-learning (accessed on 10 October 2022).

[9] Imtiaz, S.A. A Systematic Review of Sensing Technologies for Wearable Sleep Staging. Sensors 2021, 21, 1562.

[10] Imtiaz, S.A.; Rodriguez-Villegas, E. Automatic Sleep Staging Using State Machine-Controlled Decision Trees. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) , Milan, Italy, 25–29 August 2015; IEEE: Manhattan, NY, USA, 2015; pp. 378–381.

[11] Sen, B.; Peker, M.; Cavusoglu, A.; Celebi, F.V. A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. J. Med. Syst. 2014, 38, 18.

[12] Memar, P.; Faradji, F. A Novel Multi-Class EEG-Based Sleep Stage Classification System. IEEE Trans. Neural Syst. Rehabil. Eng. 2017, 26, 84–95.

[13] Santaji, S.; Desai, V. Analysis of EEG Signal to Classify Sleep Stages Using Machine Learning. Sleep Vigil. 2020, 4, 145–152.

[14] Bhusal, A.; Alsadoon, A.; Prasad, P.W.C.; Alsalami, N.; Rashid, T.A. Deep Learning for Sleep Stages Classification: Modified Rectified Linear Unit Activation Function and Modified Orthogonal Weight Initialisation. Multimed. Tools Appl. 2022, 81, 9855–9874.

[15] Tao, Y.; Yang, Y.; Yang, P.; Nan, F.; Zhang, Y.; Rao, Y.; Du, F. A Novel Feature Relearning Method for Automatic Sleep Staging Based on Single-Channel EEG. Complex Intell. Syst. 2022.

[16] Yulita, I.N.; Fanany, M.I.; Arymurthy, A.M. Sleep Stage Classification Using Convolutional Neural Networks and Bidirectional Long Short-Term Memory. In Proceedings of the 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS) , Bali, Indonesia, 5 June 2017; IEEE: Manhattan, NY, USA, 2017; pp. 303–308.

[17] Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.-K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000, 101, e215–e220.

[18] Hasib Fardin, Hasan Md Imran, Hamdadur Rahman, Anamul Haque Sakib, Md Ismail Hossain Siddiqui. Robust and explainable poultry disease classification via MaxViT with attention-guided visualization. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1848–59. Available from: https://journalijsra.com/node/1054.

[19] Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Amira Hossain, Hasib Fardin, Al Shahriar Uddin Khondakar Pranta. Custom CNN for acoustic emission classification in gas pipelines. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1760–8. Available from: https://journalijsra.com/node/1046.

[20] Mohammad Rasel Mahmud, Al Shahriar Uddin Khondakar Pranta, Anamul Haque Sakib, Abdullah Al Sakib, Md Ismail Hossain Siddiqui. Robust feature selection for improved sleep stage classification. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1790–7. Available from: https://journalijsra.com/node/1049.

[21] Md Ariful Islam, Mohammad Rasel Mahmud, Anamul Haque Sakib, Md Ismail Hossain Siddiqui, Hasib Fardin. Time domain feature analysis for gas pipeline fault detection using LSTM. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1769–77. Available from: https://journalijsra.com/node/1047

[22] Mohammad Rasel Mahmud, Hasib Fardin, Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Abdullah Al Sakib. Hybrid deep learning for interpretable lung cancer recognition across computed tomography and histopathological imaging modalities. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1798–810. Available from: https://journalijsra.com/node/1050.

[23] Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Sanjida Akter, Jesika Debnath, Mohammad Rasel Mahmud. Comparative analysis of traditional machine learning Vs deep learning for sleep stage classification. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1778–89. Available from: https://journalijsra.com/node/1048.

[24] Sanjida Akter, Mohammad Rasel Mahmud, Md Ariful Islam, Md Ismail Hossain Siddiqui, Anamul Haque Sakib. Efficient and interpretable monkeypox detection using vision transformers with explainable visualizations. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1811–22. Available from: https://journalijsra.com/node/1051.

[25] Anamul Haque Sakib, Md Ismail Hossain Siddiqui, Sanjida Akter, Abdullah Al Sakib, Mohammad Rasel Mahmud. LEVit-Skin: A balanced and interpretable transformer-CNN model for multi-class skin cancer diagnosis. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1860–73. Available from: https://journalijsra.com/node/1055.

[26] Hamdadur Rahman, Hasan Md Imran, Amira Hossain, Md Ismail Hossain Siddiqui, Anamul Haque Sakib. Explainable vision transformers for real time chili and onion leaf disease identification and diagnosis. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1823–33. Available from: https://journalijsra.com/node/1052.

[27] Siddiqui IH, Al Sakib A, Sakib AH, Fardin H, Debnath J. Dual-branch CrossViT for ovarian cancer diagnosis: Integrating and explainable AI for real-time clinical applications. Article in International Journal of Science and Research Archive [Internet]. 2025 [cited 2025 May 14];2025(01):1834–47. Available from: https://doi.org/10.30574/ijsra.2025.15.1.1164.

[28] Haque R, Khan MA, Rahman H, Khan S, Siddiqui MIH, Limon ZH, et al. Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis. Comput Biol Med [Internet]. 2025 Jun 1 [cited 2025 May 13];191:110166. Available from: https://www.sciencedirect.com/science/article/pii/S0010482525005177.

[29] Masum A Al, Limon ZH, Islam MA, Rahman MS, Khan M, Afridi SS, et al. Web Application-Based Enhanced Esophageal Disease Diagnosis in Low-Resource Settings. 2024 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON) [Internet]. 2024 Nov 28 [cited 2025 May 13];153–8. Available from: https://ieeexplore.ieee.org/document/10962580.

[30] Ahmed MR, Haque R, Rahman SMA, Afridi S, Abir MFF, Hossain MF, et al. Towards Automated Detection of Tomato Leaf Diseases. Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024. 2024;387–92.

[31] Haque R, Al Sakib A, Hossain MF, Islam F, Ibne Aziz F, Ahmed MR, et al. Advancing Early Leukemia Diagnostics: A Comprehensive Study Incorporating Image Processing and Transfer Learning. BioMedInformatics 2024, Vol 4, Pages 966-991 [Internet]. 2024 Apr 1 [cited 2025 May 13];4(2):966–91. Available from: https://www.mdpi.com/2673-7426/4/2/54/htm.

[32] Noman A Al, Hossain A, Sakib A, Debnath J, Fardin H, Sakib A Al, et al. ViX-MangoEFormer: An Enhanced Vision Transformer–EfficientFormer and Stacking Ensemble Approach for Mango Leaf Disease Recognition with

Explainable Artificial Intelligence. Computers 2025, Vol 14, Page 171 [Internet]. 2025 May 2 [cited 2025 May 13];14(5):171. Available from: https://www.mdpi.com/2073-431X/14/5/171/htm.

[33] Hosen MD, Bin Mohiuddin A, Sarker N, Sakib MS, Al Sakib A, Dip RH, et al. Parasitology Unveiled: Revolutionizing Microorganism Classification Through Deep Learning. Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024. 2024;1163–8.

[34] Al Noman A, Fardin H, Chhabra G, Sultana S, Haque R, Ahmed MR, et al. Monkeypox Lesion Classification: A Transfer Learning Approach for Early Diagnosis and Intervention. Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2024. 2024;247–54.

[35] Hasan J, Hasan K, Al Noman A, Hasan S, Sultana S, Arafat MA, et al. Transforming Leukemia Classification: A Comprehensive Study on Deep Learning Models for Enhanced Diagnostic Accuracy. PEEIACON 2024 - International Conference on Power, Electrical, Electronics and Industrial Applications. 2024;266–71.

[36] Sohaib M, Ghaffar A, Shin J, Hasan MJ, Suleman MT. Automated Analysis of Sleep Study Parameters Using Signal Processing and Artificial Intelligence. Int J Environ Res Public Health [Internet]. 2022 Oct 1 [cited 2025 May 23];19(20).

[37] Sohaib M, Hasan MJ, Shah MA, Zheng Z. A robust self-supervised approach for fine-grained crack detection in concrete structures. Scientific Reports 2024 14:1 [Internet]. 2024 Jun 2 [cited 2025 Apr 17];14(1):1–20.

[38] Hasan MJ, Elyan E, Yan Y, Ren J, Sarker MMK. Segmentation Framework for Heat Loss Identification in Thermal Images: Empowering Scottish Retrofitting and Thermographic Survey Companies. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2024 [cited 2025 Feb 2];14374 LNAI:220–8.

[39] Sohaib M, Hasan MJ, Chen J, Zheng Z. Generalizing infrastructure inspection: step transfer learning aided extreme learning machine for automated crack detection in concrete structures. Meas Sci Technol [Internet]. 2024 Feb 21 [cited 2025 May 23];35(5):055402.