Title: MVIC-Net: Multi-view interactive convolutional networks for ECG-based sleep apnea detection with explainable AI

Author names: Tanzir Ahamed , Md. Shihab Uddin , Tanvir Ahammed , Md Akteruzzaman , Md. Hasanuzzaman Dipu , M. Humayet Islam , Md. Ali Hossain , Md Masum Billah

Abstract: Deep learning models, particularly Convolutional Neural Networks (CNNs), have advanced automated Sleep Apnea (SA) detection from single-lead Electrocardiogram (ECG) signals. However, current approaches often face two limitations: (1) reliance on single-view representations (e.g., time-series or one image type) that can fail to capture the diverse feature set inherent in complex ECG dynamics; and (2) image-based methods that frequently employ standard architectures which may not optimally extract intricate local patterns crucial for SA identification. To address these challenges, we propose MVIC-Net, a Multi-View Interactive Convolutional Network framework that uniquely processes four complementary ECG representations simultaneously: 1D numeric data, 2D reshaped numeric data, Continuous Wavelet Transform (CWT) images, and circular plot images. We introduce the Convolutional and Interactive Learning Neural Network (CINet), an architecture inspired by SCINet and specifically engineered for enhanced ECG image feature extraction. CINet utilizes a hierarchical downsampling and interactive learning mechanism: it recursively splits feature maps, processes subsequences through distinct convolutional filters, and interactively combines them to mitigate information loss while capturing multi-resolution features effectively. Our MVIC-Net model integrates pre-trained CINet and standard CNN encoders, fusing their outputs via concatenation before final classification. Empirical validation on the Apnea-ECG database demonstrates superior performance, achieving 99.25% accuracy. This significantly surpasses individual-view models (with CINet reaching 96.32% accuracy on image-based views), pre-trained ResNet152V2 baselines, and a multi-model ensemble approach. To enhance clinical trust, we incorporate Explainable AI (XAI) via Grad-CAM, providing transparency into CINet's decision-making process. Our results establish the effectiveness of combining multi-view learning with interactive convolutional architectures for robust physiological signal classification.

Keywords: Sleep apnea ECG preprocessing ECG classification Image processing XAI CINet Interactive CNN Ensemble learning Multi-view classification

10.1016/j.knosys.2026.115660