Optimal Temporal Context and Lead Redundancy in Self-Supervised ECG Representation Learning

Introduction: Self-supervised learning (SSL) can learn electrocardiographic representations without manual labels, but the influence of temporal context and lead redundancy on out-of-domain generalization has not been well defined yet. We evaluated how ECG window duration and lead configuration affect downstream performance and robustness of SSL representations.

Methods: A SimSiam self-supervised encoder with a bidirectional GRU backbone was pretrained to learn temporal ECG representations using window durations from 1–10 s. Pretraining was performed on unlabeled recordings from PTB-XL and the PTB Diagnostic Database (PTB-DB). Learned representations were evaluated using a fixed linear classifier (MLP) for myocardial infarction detection. The classifier was trained on PTB-XL and evaluated both on PTB-XL and on PTB-DB as an independent out-of-domain test set. Window-level predictions were aggregated at the patient level, and experiments were repeated across multiple random seeds. Additional analyses compared full 12-lead recordings with an 8-lead subset excluding linearly derivable leads.

Results: Intermediate window durations (~3–4 s) consistently produced the strongest and most stable representations, achieving the highest validation AUC and robust test performance across seeds. Very short windows (1–2 s) showed reduced sensitivity, suggesting insufficient temporal context, while longer windows (≥6–10 s) provided no improvement and frequently degraded performance despite increased computational cost. Validation and independent test AUC values followed similar trends across configurations, indicating that the SSL encoder captured generalizable ECG structure rather than dataset-specific patterns. Performance using 8 independent leads matched that of full 12-lead inputs, suggesting limited benefit from redundant or derived leads.

Conclusions: Self-supervised ECG representations generalize across datasets when trained with appropriate temporal context. Window durations of approximately 3–4 s provide an optimal balance between contextual information and robustness. Furthermore, removing redundant leads maintains performance while improving efficiency. These findings provide practical guidance for scalable and resource-efficient SSL-based ECG analysis.

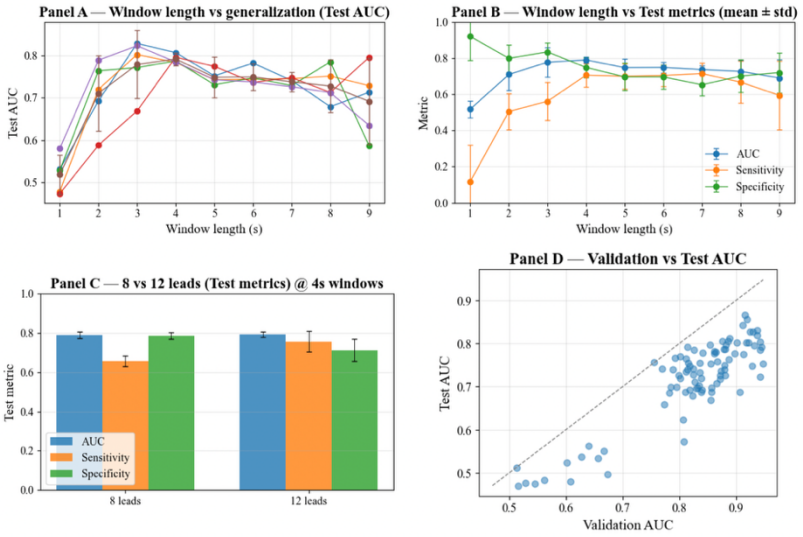

Figure. Effect of ECG window length and lead set on out-of-domain generalization. (A) Test AUC versus window length, showing variability across runs. (B) Mean ± SD test AUC, sensitivity, and specificity as a function of window length. (C) Comparison of 8 versus 12 leads at 4 sec windows. (D) Validation versus test AUC highlighting the generalization gap across configurations.