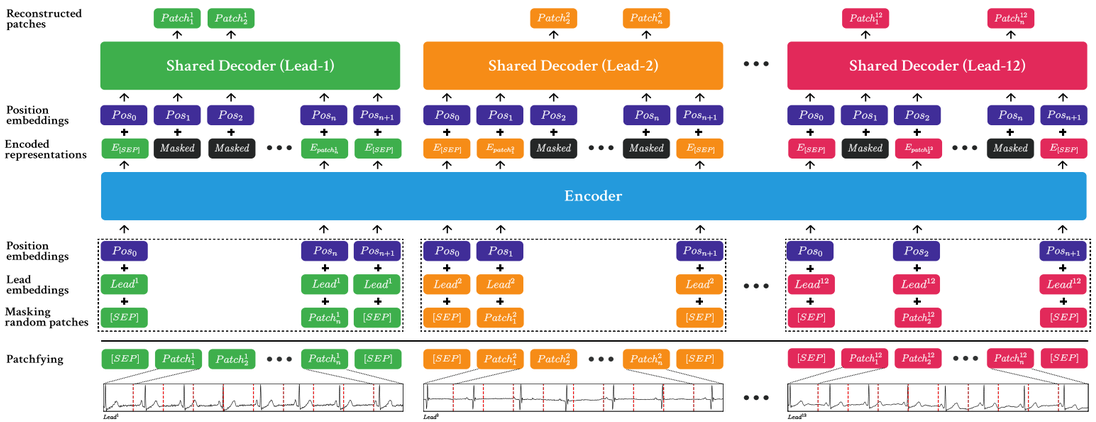
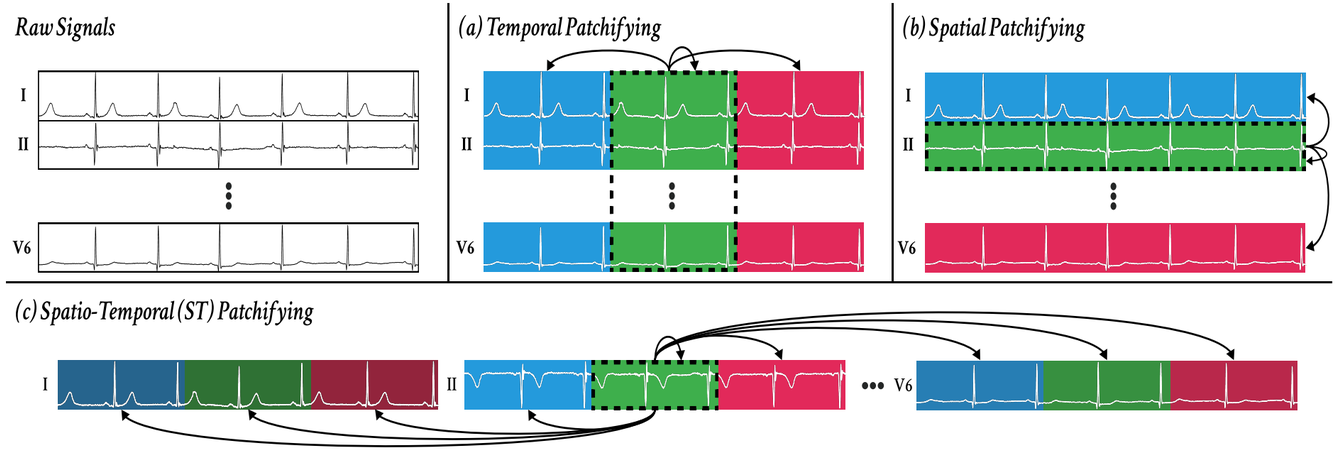
Guiding Masked Representation Learning to Capture Spatio-Temporal Relationship of Electrocardiogram

Yeongyeon Na* · Minje Park* · Yunwon Tae* · Sunghoon Joo†
* Equal contribution, † Corresponding author.

We propose the **simple but effective** ECG-specific generative self-supervised learning framework, named **ST-MEM** (Spatio-Temporal Masked Electrocardiogram Modeling).



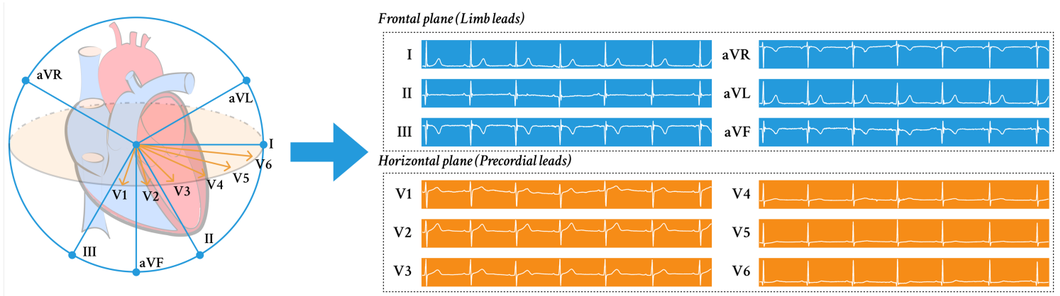
- ST-MEM includes **lead indicators** (lead-wise shared decoder, learnable lead embeddings, and separation embedding).



- ST-MEM can learn general representation by capturing spatio-temporal relationship of ECGs through **spatio-temporal patchifying.**

Introduction

What is ECGs? The ECG is a non-invasive heart measurement to observe the electrical signals over time and diagnose diseases. A standard 12-lead ECG is the most common measurement setting that provides spatial and temporal information regarding the heart.



Why do we need representation learning in ECGs?

- High cost of labeling ECGs (hiring cardiologists → $$$)
- Diverse heart diseases → Diverse downstream tasks
- Small number of input features from mobile ECG devices such as smartwatches

Experiments

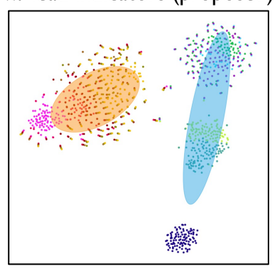
Fine-tuning results of arrhythmia and myocardial infarction (MI) classification tasks. The experiment is conducted based on 12-lead ECG data on unseen data (i.e., not used during the pre-training stage).

Methods	*PTB-XL*			*CPSC2018*		
	1%	5%	100%	1%	5%	100%
Supervised	0.676 ± 0.011	0.736 ± 0.020	0.905 ± 0.004	0.600 ± 0.095	0.609 ± 0.111	0.958 ± 0.002
MoCo v3	0.797 ± 0.006	0.826 ± 0.015	0.913 ± 0.002	0.791 ± 0.045	0.903 ± 0.019	0.967 ± 0.003
CMSC	0.648 ± 0.064	0.773 ± 0.023	0.877 ± 0.003	0.625 ± 0.013	0.732 ± 0.038	0.938 ± 0.006
MTAE	0.707 ± 0.024	0.713 ± 0.001	0.910 ± 0.001	0.670 ± 0.032	0.756 ± 0.013	0.961 ± 0.001
MTAE+RLM	0.730 ± 0.030	0.730 ± 0.003	0.911 ± 0.004	0.708 ± 0.020	0.726 ± 0.011	0.960 ± 0.002
MLAE	0.793 ± 0.007	0.838 ± 0.018	0.915 ± 0.001	0.860 ± 0.013	0.922 ± 0.007	0.973 ± 0.002
(250 Hz) CPC†	0.740 ± 0.057	0.838 ± 0.024	0.933 ± 0.001	0.754 ± 0.015	0.898 ± 0.026	0.974 ± 0.002
(100 Hz) CPC†	0.773 ± 0.014	0.842 ± 0.043	**0.934 ± 0.002**	0.762 ± 0.058	0.917 ± 0.016	0.973 ± 0.003
ST-MEM (Ours)	**0.815 ± 0.012**	**0.878 ± 0.011**	0.933 ± 0.003	**0.897 ± 0.025**	**0.952 ± 0.004**	**0.980 ± 0.001**

Robustness of any lead combinations.

Methods	*PTB-XL*			*CPSC2018*			*PhysioNet2017*
	12-lead	6-lead	1-lead	12-lead	6-lead	1-lead	1-lead
MTAE+RLM	0.911 ± 0.004	0.888 ± 0.002	0.795 ± 0.003	0.960 ± 0.002	0.931 ± 0.017	0.909 ± 0.006	0.857 ± 0.005
MLAE	0.915 ± 0.001	0.890 ± 0.001	0.797 ± 0.001	0.973 ± 0.002	0.959 ± 0.002	0.925 ± 0.001	0.861 ± 0.003
ST-MEM (Ours)	**0.933 ± 0.003**	**0.903 ± 0.007**	**0.804 ± 0.005**	**0.980 ± 0.001**	**0.973 ± 0.002**	**0.937 ± 0.006**	**0.866 ± 0.003**

w/ lead indicators (proposed)



w/o lead indicators (ablation)