

Self-Supervised Pretraining Strategies for Few-Shot ECG Classification in Transformers and CNNs

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the impact of self-supervised pretraining strategies (e.g., MoCo, DINO) on the performance of Transformers compared to CNNs in few-shot ECG classification, evaluated using AUC-ROC and. 14 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Pretraining Strategies and Scaling for ECG Foundation Models: A Systematic Study. Research question: What is the impact of self-supervised pretraining strategies (e.g., MoCo, DINO) on the performance of Transformers compared to CNNs in few-shot ECG classification, evaluated using AUC-ROC and robustness to label scarcity?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.5/10.

3 Results

11 papers retrieved. 14 claims extracted; 0 independently verified. Quality review score: 3.5/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The study covers five different pretraining methodologies.	×	0.12
The pretraining corpus comprises over 11 million samples.	×	0.04
State space models (SSM) are the superior architecture choice across all pretraining paradigms compared to CNNs and Tran	×	0.12
CPC shows the strongest and most transferable representations across diverse clinical tasks among the five methodologies	×	0.12
data2vec consistently lags behind other methodologies across all evaluation modes and scaling regimes.	×	0.04
Scaling behavior is most clearly identified for CPC and JEPA methodologies.	×	0.09
Lower pretraining loss correlates with small residual errors in downstream tasks.	×	0.03
Recent ECG foundation models have been pre-trained on tens of millions of recordings.	×	0.14
	×	0.00
All models operate at 240 Hz on 12-lead ECG inputs.	×	0.15
The S4 backbone with model dimension 512 consistently outperforms larger and alternative configurations in ablation stud	×	0.02
The study adopts a 4-layer S4 backbone with dimension 512 as the default configuration.	×	0.01
The study investigates five self-supervised pre-training objectives spanning contrastive, predictive, and clustering-base	×	0.11
data2vec trains the model to predict the EMA teacher’s averaged top-k contextualized layer representations at masked pos	×	0.01

References

- <http://arxiv.org/abs/2203.07057v2>
- <http://arxiv.org/abs/1911.06045v3>
- <http://arxiv.org/abs/2605.12241v1>