Masking Strategies in Self-Supervised Pretraining for Zero-Shot Physiological Sequence Transfer

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How do different masking strategies in self-supervised pretraining affect the zero-shot transfer capabilities of physiological sequence models on the PhysioNet 2020 challenge dataset. 12 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.8/10. 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: How do different masking strategies in self-supervised pretraining affect the zero-shot transfer capabilities of physiological sequence models on the PhysioNet 2020 challenge dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.8/10.

3 Results

16 papers retrieved. 12 claims extracted; 0 independently verified. Quality review score: 2.8/10.

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 trained on a corpus comprising over 11 million ECG samples.	×	0.10
State space models (SSM) are confirmed as the superior architecture choice across all pretraining paradigms compared to	×	0.10
The CPC pretraining strategy shows the strongest and most transferable representations across diverse clinical tasks.	×	0.14
The data2vec pretraining strategy consistently lags behind other methodologies across all evaluation modes and scaling r	×	0.08
Scaling behavior is most clearly identified for the CPC and JEPA pretraining methodologies.	×	0.14
Lower pretraining loss correlates with small residual errors in downstream tasks.	×	0.03
The S4 backbone with a model dimension of 512 consistently outperforms larger dimensions (768, 1024) and alternative bac	×	0.04
All models in the study operate at a sampling rate of 240 Hz on 12-lead ECG inputs.	×	0.12
The default backbone adopted for the study is a 4-layer S4 model with a dimension of 512.	×	0.03
The CNN stem used in all models consists of four convolutional layers with batch normalization.	×	0.04
The Transformer backbone variant utilizes RoPE positional encoding and GELU activations.	×	0.02
The data2vec objective trains the model to predict the EMA teacher's averaged top-k contextualized layer representations	×	0.04

References

- http://arxiv.org/abs/2103.12676v2
- http://arxiv.org/abs/2106.15577v5
- http://arxiv.org/abs/2605.12241v1