PhysioJEPA: Joint Embedding Representations of Physiological Signals for Real Time Risk Estimation in the Intensive Care Unit

Benjamin Fox
Icahn School of Medicine at Mount Sinai
ben.fox@icahn.mssm.edu

Dung Hoang
Icahn School of Medicine at Mount Sinai
jolie.hoang@icahn.mssm.edu

Ankit Parekh
Icahn School of Medicine at Mount Sinai
ankit.parekh@mssm.edu

Girish Nadkarni
Icahn School of Medicine at Mount Sinai
girish.nadkarni@mountsinai.org

Ankit Sakhuja
Icahn School of Medicine at Mount Sinai
ankit.sakhuja@mssm.edu

Abstract

Self-supervised learning of multichannel, high-frequency physiological signals is largely unexplored, despite their potential for critical care applications. We present PhysioJEPA, the first joint embedding predicting architecture (JEPA) designed for multichannel physiological signals from critical care bedside monitoring devices. PhysioJEPA learns representations from 30-minute segments of physiological signals from three channels: arterial blood pressure, electrocardiography lead II, and photoplethysmography, using the MIMIC-III Waveform Database. Trained on over 10.7 million minutes of arterial blood pressure, electrocardiography lead II, and photoplethysmography from 4,282 intensive care unit stays (N=2,631 patients), PhysioJEPA's learned representations can be finetuned to estimate 5-minute risk of hypotension (AUROC = 0.83 [Confidence Interval or CI 0.83-0.84]) and shock index (AUROC = 0.95 [0.95-0.96]) improving performance over a supervised convolutional baseline (AUROC = 0.78 [0.78-0.78] and 0.95 [0.95-0.95] for hypotension and shock index, respectively). Furthermore, it generalized to an external cohort from the Mount Sinai Health System (AUROC = 0.78 [0.78-0.78] and 0.92 [0.92-0.93]). These results suggest that self-supervised JEPA representation learning is a promising approach for critical care signal data.

1 Introduction

Critical care monitoring via bedside devices generates vast amounts of high-frequency physiological signal data, yet current machine learning approaches primarily rely on aggregated signals and electronic health record features, limiting their ability for real-time risk estimation and to capture complex multichannel temporal relationships, which are essential for early warning systems in the intensive care unit (ICU) [Maheshwari et al., 2021, Moghadam et al., 2021, Yoon et al., 2020]. While recent advances in self-supervised representation learning have shown promise in learning physiological signals [Lutsker et al., 2024, Merrill and Althoff, 2022, Zhang et al., 2022, Yuan et al., 2024, Kolbeinsson et al., 2021], existing frameworks focus on shorter durations, single modalities, or

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Foundation Models for the Brain and Body.

low frequency data, missing opportunities to learn comprehensive physiological patterns. Critically, to date, no other work utilizes representation learning frameworks specifically designed for bedside monitoring devices.

Representation learning techniques continue to develop for time series data [Nie et al., 2023, Foumani et al., 2023, Zhang et al., 2022]. Recently, Joint Embedding Predictive Architectures (JEPA) have demonstrated superior performance in computer vision by learning representations through prediction in the embedding space rather than pixel reconstruction via masked autoregression [Assran et al., 2023, Bardes et al., 2024]. This approach is particularly well-suited for physiological signals, where avoiding reconstruction of noisy data while capturing temporal dependencies across multiple channels could yield more robust representations. The adaptation of JEPA to handle long, high-frequency time-series data represents a significant methodological innovation as modeling extended physiological recordings, especially in critical care settings, presents unique challenges due to massive data volume, complexity, and noise.

In this work, we introduce PhysioJEPA, the first JEPA-based model for multichannel physiological signals that: (1) learns representations from 30-minute, three-channel bedside monitoring signal segments through self-supervised training, (2) enables accurate prediction of critical care outcomes via task-specific fine-tuning, and (3) demonstrates cross-site generalizability through external validation. PhysioJEPA fills the gap in representation learning for bedside monitoring signal data and provides a foundation for future advances in critical care risk estimation.

2 Methods

2.1 PhysioJEPA Architecture

PhysioJEPA adapted the JEPA framework for physiological time series using PatchTST encoders [Nie et al., 2023], which were based on the original vision transformer (similar to JEPA). The architecture processed 30-minute segments of three physiological channels (arterial blood pressure [ABP], electrocardiography [ECG] lead II, and photoplethysmography [PPG] sampled at 125 Hz.

Self-Supervised Training: Following JEPA principles, we randomly selected non-overlapping context (10-40%) and targeted (10-30%) patches of tokenized time series data across all signal channels. The context encoder processed selected patches, and a predictor network estimated target embeddings from context patches. A target encoder, updated via exponential moving average, generated ground-truth embeddings for comparison to the predictor's output. A key difference in our JEPA architecture was that there are separate mask tokens for each signal channel, potentially allowing for more distinct encodings per channel. Additionally, we utilized depthwise convolutions to tokenize patches of timeseries data into 1-seconds patches, with the expected transformer dimension. Depthwise convolutions created separate kernels per input channel. Mean squared error loss was used to minimize the distance between predicted and target embeddings during training.

Implementation Details: Context and target encoders used 3-layer PatchTST encoders with 8 attention heads, dimension 512, and feedforward sizes of 2048. The predictor used 2 layers with 4 heads and a dimension of 256. The representation framework was trained for 100 epochs with a one-cycle learning rate scheduler and AdamW optimizer.

2.2 Data and Clinical Tasks

The MIMIC-III Waveform Database [Johnson et al., 2015, 2016] was used for training. For representation learning, data was split into 90% for training and 10% for validation. We then evaluated learned representations through fine-tuning attentive classifiers on two critical care risk estimation tasks:

5-Minute Hypotensive Risk: Risk estimation of hypotension (mean arterial pressure ≤ 65 mmHg or systolic blood pressure ≤ 90 mmHg) at a 5-minute forecast. A hypotensive event (i.e. label) was defined as five consecutive minutes below the hypotension thresholds, using the ABP signal channel and a peak detection algorithm adapted from Physionet's wfdb package [Goldberger et al., 2000, Moody et al., 2022]. Hypotensive events longer than 5-minutes were treated as a single event. Thus, for patients who had multiple events, there was at least 5 minutes in between every event. Data was split into training (80%), validation (10%), and testing (10%) using a proportional, subject-wise data

splitter. We chose hypotensive risk due to its common occurrence for patients admitted to the ICU (60-75% of people admitted to the ICU develop hypotension [Terwindt et al., 2022]), and there is a critical need for improved monitoring and proactive management of patients at risk.

5-Minute Shock Index Risk: Risk estimation of shock index (≥ 0.9) at a 5-minute forecast. Shock index was calculated as heart rate divided by systolic blood pressure for each minute of signal data [Cannon et al., 2009]. A shock index event (i.e. label) was defined as five consecutive minutes above the threshold. Heart rate was derived from the ABP signal channel, using Physionet's peak detection algorithms [Goldberger et al., 2000, Moody et al., 2022]. Data was split into training (80%), validation (10%), and testing (10%) using a proportional label, subject-wise data splitter. Shock index is the ratio of heart rate in beats per minute over systolic blood pressure and can be an important marker for shock and mortality [Koch et al., 2019]. A shock index of 0.5 to 0.7 is considered normal, while a higher shock index is more predictive of mortality [Cannon et al., 2009].

2.3 Fine Tuning Architecture

After representation learning, fine-tuning classifiers were trained with learned representations for each task. Notably, the weights of the context encoder from JEPA were frozen during this process. Attentive classifiers [Assran et al., 2023, Bardes et al., 2024] performed cross attention (4 attention heads) with a single query vector on generated JEPA representations, followed by a linear layer for classification. Classifiers were trained for 20 epochs with a one-cycle learning rate scheduler and the AdamW optimizer.

2.4 Baseline Models

For comparison, we trained two fully supervised baseline convolutional classifiers using a strong supervised time-series-based classifier presented by Wang et al. [2016] for the two tasks. A three-layer convolutional classifier with standard parameters was used (dimensions: 128, 256, and 128; kernel sizes: 7, 5, and 3). Classifiers again were trained for 20 epochs with a one-cycle learning rate scheduler and AdamW optimizer.

2.5 External Validation

To assess generalizability, we conducted external validation on the Mount Sinai Bedmaster dataset from 100 randomly selected patients (per task) from 6 separate adult ICUs from 2019 to 2024.

Models were evaluated using area under the receiver operating characteristic curve (AUROC), average precision, F1, recall, specificity, and sensitivity at 90% and 95% specificity (Sens@90%Spec, Sens@95%Spec, respectively).

3 Results

An overview of PhysioJEPA's architecture is shown in Figure 1. Of the 5,660 stays from the MIMIC-III Waveform Database with the required three signal channel data (ABP, ECG lead II, PPG), 1,378 were not used for representation learning due to 20% or more constant or NaN values in a single channel (in each 30-minute segment). We investigated this missingness further and found that the majority of the removed waveforms had $\geq 75\%$ missing or NaN values (total of 1,050 samples). Thus, our missingness threshold of 20% was removing samples that were majority missing or NaN values. Overall, PhysioJEPA was trained with 356,903 30-minute segments (total of 10,707,090 minutes) across 4,282 ICU stays (N=2,631 patients) with 3-channel signal data. Training was stopped after 100 epochs, and the last model was used for fine-tuning.

Hypotensive and shock-index labels were derived from the ABP signal channel. Each subject's ABP signal was processed in 1-minute segments to identify average heart rate in beats per minute and median systolic and diastolic blood pressures. Mean arterial pressure was calculated as $\frac{1}{3} * systolic + \frac{2}{3} * diastolic$. A rolling window was applied to the derived blood pressure and shock index minute-level labels to identify 5 minutes (or more) of continuous values meeting the threshold requirements for the task. Given this definition, it was possible (and likely) that multiple events were identified per patient (with at least five minutes in between events). Following, 30-minute input segments five minutes before the first minute of the identified events were extracted for training.

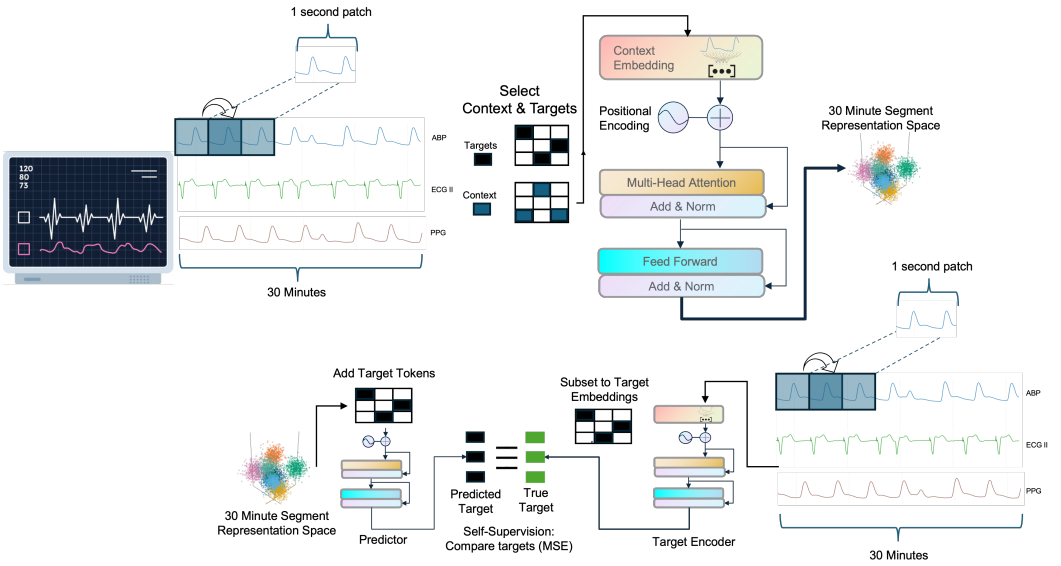

Figure 1: PhysioJEPA Architecture Overview: Bedside monitoring signal channels (ABP, ECG lead II, and PPG) are extracted in 30-minute segments sampled at 125 Hz and tokenized into 1-second patches. Random patches are selected by context and target masks, which are encoded using a PatchTST encoder [Nie et al., 2023]. Channel specific mask tokens are added to the context embeddings to indicate which patches the predictor must reconstruct, while the target encoder processes the full segment and subsets embeddings at the masked locations. Prediction is optimized by minimizing mean squared error between predicted and target embeddings.

	Hypotension				Shock Index			
Data Split	N ICU Stays	N Patients	Positive Events	Negative Events	N ICU Stays	N Patients	Positive Events	Negative Events
MIMIC-III Train	3280	2060	44897 (0.04)	1015429 (0.96)	3264	2062	53897 (0.05)	1087321 (0.95)
MIMIC-III Validation	381	237	5601 (0.05)	116307 (0.95)	335	216	6882 (0.07)	96229 (0.93)
MIMIC-III Test	341	229	4435 (0.04)	106273 (0.96)	413	245	8526 (0.05)	149215 (0.95)
Mount Sinai Bedmaster	99	99	2638 (0.03)	82626 (0.97)	98	98	1952 (0.05)	37089 (0.95)

Table 1: Hypotension and shock index event statistics for training, validation, testing, and external test sets.

Non-hypotensive and non-shock index samples included only patients who did not experience any hypotensive or shock index minute for their entire ICU stay. Dataset statistics are reported in Table 1.

3.1 Risk Estimation Results

Baseline Comparison: Two fully supervised convolutional classifiers achieved AUROC scores of 0.778 (95% bootstrapped confidence interval [CI]: 0.771–0.784) for 5-minute hypotension and 0.950 (95% CI: 0.948–0.952) for 5-minute shock index risk estimation on the held-out test set.

JEPA Representation Performance: Finetuning with frozen PhysioJEPA encoder weights outperformed baselines, AUROC scores of 0.833 (95% CI: 0.825–0.838) for 5-minute hypotension risk and 0.954 (95% CI: 0.952–0.955) for 5-minute shock index risk estimation. See Figure 2 for receiver operating characteristic curves.

At clinically relevant specificity thresholds, PhysioJEPA achieved superior performance: 55.2% sensitivity at 90% specificity and 40.7% at 95% specificity for 5-minute hypotension risk estimation, compared to baseline performance of 42.2% and 29.6%, respectively. For shock index, PhysioJEPA achieved 88.1% and 70.6% sensitivity at 90% and 95% specificity thresholds, outperforming baseline results of 85.2% and 67.6%. Additional performance metrics are detailed in Table 2.

Condition	Dataset	Model	AUROC	Avg Precision	F1	Recall	Specificity	Sens@90%Spec	Sens@95%Spec
Hypotension	Held-Out	JEPA	**0.83 (0.83 - 0.84)**	**0.26 (0.25 - 0.28)**	**0.12 (0.12 - 0.13)**	0.92 (0.92 - 0.93)	**0.45 (0.44 - 0.45)**	**0.55 (0.54 - 0.57)**	**0.41 (0.39 - 0.42)**
		Baseline	0.78 (0.77 - 0.78)	0.14 (0.13 - 0.15)	0.08 (0.08 - 0.09)	**1.00 (1.00 - 1.00)**	0.07 (0.07 - 0.07)	0.42 (0.41 - 0.44)	0.30 (0.28 - 0.31)
	External	JEPA	**0.78 (0.77 - 0.79)**	**0.10 (0.09 - 0.10)**	**0.08 (0.08 - 0.09)**	**0.95 (0.94 - 0.96)**	**0.34 (0.34 - 0.34)**	**0.34 (0.32 - 0.36)**	**0.20 (0.19 - 0.22)**
		Baseline	0.70 (0.69 - 0.71)	0.06 (0.05 - 0.06)	0.06 (0.06 - 0.07)	0.92 (0.91 - 0.93)	0.15 (0.15 - 0.15)	0.21 (0.20 - 0.23)	0.07 (0.06 - 0.07)
Shock Index	Held-Out	JEPA	**0.95 (0.95 - 0.96)**	0.47 (0.47 - 0.49)	**0.31 (0.30 - 0.31)**	**0.98 (0.98 - 0.98)**	**0.75 (0.75 - 0.75)**	**0.88 (0.88 - 0.89)**	**0.71 (0.70 - 0.72)**
		Baseline	0.95 (0.95 - 0.95)	**0.56 (0.55 - 0.57)**	0.30 (0.29 - 0.30)	0.98 (0.97 - 0.98)	0.74 (0.73 - 0.74)	0.85 (0.85 - 0.86)	0.68 (0.67 - 0.69)
	External	JEPA	**0.92 (0.92 - 0.93)**	**0.40 (0.38 - 0.41)**	**0.23 (0.22 - 0.23)**	**0.97 (0.96 - 0.98)**	**0.65 (0.65 - 0.66)**	**0.71 (0.70 - 0.74)**	**0.36 (0.34 - 0.37)**
		Baseline	0.78 (0.77 - 0.79)	0.16 (0.15 - 0.17)	0.17 (0.16 - 0.17)	0.84 (0.83 - 0.85)	0.57 (0.56 - 0.57)	0.52 (0.50 - 0.55)	0.20 (0.19 - 0.22)

Table 2: Comprehensive performance comparison of PhysioJEPA and supervised convolutional baseline model on hypotension and shock index risk estimation tasks across the held-out MIMIC-III test set and Mount Sinai Bedmaster external dataset. Best performance for each metric is shown in bold (95% bootstrapped confidence intervals).

3.2 External Validation Results

External validation on 100 randomly selected samples (for each task) from the Mount Sinai Bedmaster dataset (comprising 6 different adult ICUs) demonstrated strong cross-site generalizability without retraining. 3 samples were excluded due to our missingness constraints as detailed in the Methods. 5-minute hypotension risk estimation using the JEPA architecture and classifier trained on MIMIC-III data achieved an AUROC of 0.781 (95% CI: 0.773–0.788) compared to 0.695 (95% CI: 0.686–0.705) for the baseline classifier. Similarly, 5-minute shock index risk estimation achieved an AUROC of 0.923 (95% CI: 0.917–0.927) compared to 0.782 (95% CI: 0.771–0.794) for the baseline classifier.

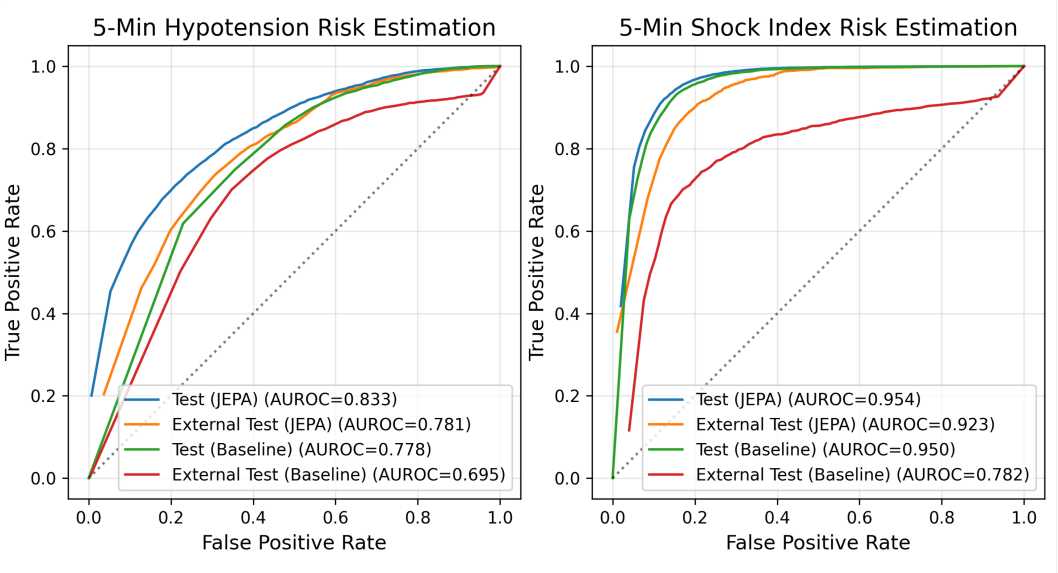

Figure 2: Receiver operating characteristic curves for 5-minute hypotension (left) and shock index (right) risk estimation across the held-out test set and the Mount Sinai Bedmaster external test set.

.

4 Discussion and Conclusion

PhysioJEPA demonstrates that JEPA-based self-supervised learning effectively captures complex multichannel physiological signal relationships for critical care risk estimation. The strong performance on both tasks, particularly the substantial improvements over supervised baselines on the external test set, establishes the potential for real-world ICU deployment and cross-site adaptability.

Clinical Impact: The 5-minute prediction horizon provides actionable early warning capabilities for preventing adverse events. PhysioJEPA's superior performance compared to supervised baselines demonstrates the clinical value of self-supervised representation learning. The model architecture relies solely on bedside monitoring data, enabling deployment across diverse ICU environments without requiring electronic health record integration.

Technical Contributions: Our work extends JEPA to physiological time series and shows that embedding-space prediction can learn relevant features of physiological signal data and generalize better to unseen datasets for predictive tasks, compared to supervised learning. Furthermore, given that the representations are generated without being fine-tuned to specific tasks, they may be applicable to other critical care outcomes, which merits further investigation.

Comparison to Existing Approaches: Unlike prior hypotensive risk models that were based on aggregated features [Maheshwari et al., 2021, Moghadam et al., 2021, Yoon et al., 2020], PhysioJEPA leverages raw, unaggregated multichannel bedside monitoring data to generate features through self-supervised representation learning. This approach eliminates the need for handcrafted feature extraction while achieving competitive performance.

Limitations and Future Work: Future work should evaluate noninvasive risk estimation by removing the ABP channel. Longer forecast horizons (e.g., 10–15 minutes), varying input segment lengths, alternative preprocessing, additional datasets, and task-specific fine-tuning should also be explored to enhance training and assess generalization. Overall, PhysioJEPA provides a foundation for self-supervised representation learning of multichannel physiological signals, supporting more robust and generalizable models for critical care risk estimation.

References

Kamal Maheshwari, Sai Buddi, Zhongping Jian, Jos Settels, Tetsuya Shimada, Barak Cohen, Daniel I. Sessler, and Feras Hatib. Performance of the Hypotension Prediction Index with non-invasive arterial pressure waveforms in non-cardiac surgical patients. *Journal of Clinical Monitoring and Computing*, 35(1):71–78, feb 1 2021. ISSN 1573-2614. doi: 10.1007/s10877-020-00463-5.

Mina Chookhachizadeh Moghadam, Ehsan Masoumi, Samir Kendale, and Nader Bagherzadeh. Predicting hypotension in the ICU using noninvasive physiological signals. *Computers in Biology and Medicine*, 129:104120, feb 1 2021. ISSN 0010-4825. doi: 10.1016/j.compbiomed.2020.104120.

Joo Heung Yoon, Vincent Jeanselme, Artur Dubrawski, Marilyn Hravnak, Michael R. Pinsky, and Gilles Clermont. Prediction of hypotension events with physiologic vital sign signatures in the intensive care unit. *Critical Care*, 24(1):1–9, 12 2020. ISSN 1364-8535. doi: 10.1186/s13054-020-03379-3. number: 1 publisher: BioMed Central.

Guy Lutsker, Gal Sapir, Anastasia Godneva, Smadar Shilo, Jerry R. Greenfield, Dorit Samocha-Bonet, Shie Mannor, Eli Meirom, Gal Chechik, Hagai Rossman, and Eran Segal. From Glucose Patterns to Health Outcomes: A Generalizable Foundation Model for Continuous Glucose Monitor Data Analysis. aug 20 2024. doi: 10.48550/arXiv.2408.11876. URL `http://arxiv.org/abs/2408.11876`. arXiv:2408.11876 [cs, q-bio].

Mike A. Merrill and Tim Althoff. Self-supervised Pretraining and Transfer Learning Enable Flu and COVID-19 Predictions in Small Mobile Sensing Datasets. jun 2 2022. doi: 10.48550/arXiv.2205.13607. URL `http://arxiv.org/abs/2205.13607`. arXiv:2205.13607 [cs].

Xiang Zhang, Ziyuan Zhao, Theodoros Tsiligkaridis, and Marinka Zitnik. Self-Supervised Contrastive Pre-Training For Time Series via Time-Frequency Consistency. oct 15 2022. doi: 10.48550/arXiv.2206.08496. URL `http://arxiv.org/abs/2206.08496`. arXiv:2206.08496 [cs].

Hang Yuan, Shing Chan, Andrew P. Creagh, Catherine Tong, Aidan Acquah, David A. Clifton, and Aiden Doherty. Self-supervised learning for human activity recognition using 700,000 person-days of wearable data. *npj Digital Medicine*, 7(1):1–10, apr 12 2024. ISSN 2398-6352. doi: 10.1038/s41746-024-01062-3. publisher: Nature Publishing Group.

Arinbjörn Kolbeinsson, Piyusha Gade, Raghu Kainkaryam, Filip Jankovic, and Luca Foschini. Self-supervision of wearable sensors time-series data for influenza detection. dec 27 2021. doi: 10.48550/arXiv.2112.13755. URL `http://arxiv.org/abs/2112.13755`. arXiv:2112.13755 [cs].

Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. mar 5 2023. doi: 10.48550/arXiv.2211.14730. URL `http://arxiv.org/abs/2211.14730`. arXiv:2211.14730 [cs].

Navid Mohammadi Foumani, Chang Wei Tan, Geoffrey I. Webb, and Mahsa Salehi. Improving Position Encoding of Transformers for Multivariate Time Series Classification. may 26 2023. doi: 10.48550/arXiv.2305.16642. URL http://arxiv.org/abs/2305.16642. arXiv:2305.16642 [cs] version: 1.

Mahmoud Assran, Quentin Duval, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael Rabbat, Yann LeCun, and Nicolas Ballas. Self-Supervised Learning from Images with a Joint-Embedding Predictive Architecture. apr 13 2023. doi: 10.48550/arXiv.2301.08243. URL http://arxiv.org/abs/2301.08243. arXiv:2301.08243 [cs].

Adrien Bardes, Quentin Garrido, Jean Ponce, Xinlei Chen, Michael Rabbat, Yann LeCun, Mahmoud Assran, and Nicolas Ballas. Revisiting Feature Prediction for Learning Visual Representations from Video, February 2024. URL http://arxiv.org/abs/2404.08471. arXiv:2404.08471 [cs].

Alistair Johnson, Tom Pollard, and Roger Mark. Mimic-III Clinical Database, 2015. URL https://physionet.org/content/mimiciii/1.4/. [Online; accessed 2024-03-06].

Alistair E. W. Johnson, Tom J. Pollard, Lu Shen, Li-wei H. Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G. Mark. Mimic-III, a freely accessible critical care database. *Scientific Data*, 3(1):160035, may 24 2016. ISSN 2052-4463. doi: 10.1038/sdata.2016.35. number: 1 publisher: Nature Publishing Group.

Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), June 2000. ISSN 1524-4539. doi: 10.1161/01.cir.101.23.e215. URL http://dx.doi.org/10.1161/01.cir.101.23.e215.

George Moody, Tom Pollard, and Benjamin Moody. Wfdb software package, 2022. URL https://physionet.org/content/wfdb/10.7.0/.

Lotte E. Terwindt, Jaap Schuurmans, Björn J. P. van der Ster, Carin A. G. C. L. Wensing, Marijn P. Mulder, Marije Wijnberge, Thomas G. V. Cherpanath, Wim K. Lagrand, Alain A. Karlas, Mark H. Verlinde, Markus W. Hollmann, Bart F. Geerts, Denise P. Veelo, and Alexander P. J. Vlaar. Incidence, Severity and Clinical Factors Associated with Hypotension in Patients Admitted to an Intensive Care Unit: A Prospective Observational Study. *Journal of Clinical Medicine*, 11(22): 6832, nov 18 2022. ISSN 2077-0383. doi: 10.3390/jcm11226832. PMID: 36431308 PMCID: PMC9696980.

Chad M. Cannon, Carla C. Braxton, Mendy Kling-Smith, Jonathan D. Mahnken, Elizabeth Carlton, and Michael Moncure. Utility of the shock index in predicting mortality in traumatically injured patients. *The Journal of Trauma*, 67(6):1426–1430, December 2009. ISSN 1529-8809. doi: 10.1097/TA.0b013e3181bbf728.

Erica Koch, Shannon Lovett, Trac Nghiem, Robert A Riggs, and Megan A Rech. Shock index in the emergency department: utility and limitations. *Open Access Emergency Medicine : OAEM*, 11:179–199, August 2019. ISSN 1179-1500. doi: 10.2147/OAEM.S178358. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6698590/.

Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. *CoRR*, abs/1611.06455, 2016. URL http://arxiv.org/abs/1611.06455.