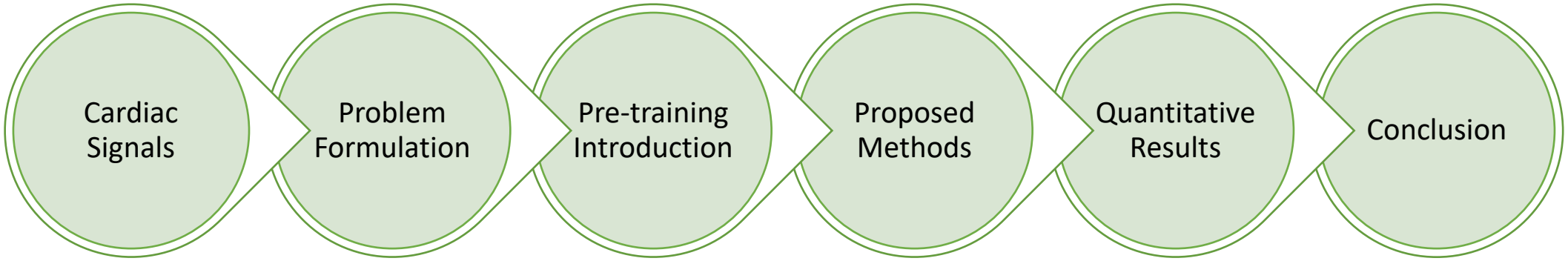
CLOCS: Contrastive Learning of Cardiac Signals

Invited Talk:
ML4CVD
Broad Institute

Dani Kiyasseh
dani.kiyasseh@eng.ox.ac.uk
August 25th, 2020

Roadmap


Cardiac Signals
Problem Formulation
Pre-training Introduction
Proposed Methods
Quantitative Results
Conclusion

The Power of Cardiac Signals

Letter | Published: 07 January 2019

Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network

Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H. Tison, Codie Bourn, Mintu P. Turakhia & Andrew Y. Ng

Letter | Published: 11 May 2020

Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network

Sushravya Raghunath, Alvaro E. Ulloa Cerna, Linyuan Jing, David P. vanMaanen, Joshua Stough, Dustin N. Hartzel, Joseph B. Leader, H. Lester Kirchner, Martin C. Stumpe, Ashraf Hafez, Arun Nemani, Tanner Carbonati, Kipp W. Johnson, Katelyn Young, Christopher W. Good, John M. Pfeifer, Aalpen A. Patel, Brian P. Delisle, Amro Alsaid, Dominik Beer, Christopher M. Haggerty & Brandon K. Fornwalt

Letter | Published: 07 January 2019

Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram

Zachi I. Attia, Suraj Kapa, Francisco Lopez-Jimenez, Paul M. McKie, Dorothy J. Ladewig, Gaurav Satam, Patricia A. Pellikka, Maurice Enriquez-Sarano, Peter A. Noseworthy, Thomas M. Munger, Samuel J. Asirvatham, Christopher G. Scott, Rickey E. Carter & Paul A. Friedman

JAMA Cardiology | Original Investigation

Development and Validation of a Deep-Learning Model to Screen for Hyperkalemia From the Electrocardiogram

Conner D. Galloway, MSc; Alexander V. Valys, BS; Jacqueline B. Shreibati, MD; Daniel L. Treiman, BS; Frank L. Petterson, PhD; Vivek P. Gundotra; David E. Albert, MD; Zachi I. Attia, MSc; Rickey E. Carter, PhD; Samuel J. Asirvatham, MD; Michael J. Ackerman, MD, PhD; Peter A. Noseworthy, MD; John J. Dillon, MD; Paul A. Friedman, MD

Problem Formulation

Rate of data generation far exceeds
that of labelling by expert annotators

Presence of small,
labelled datasets

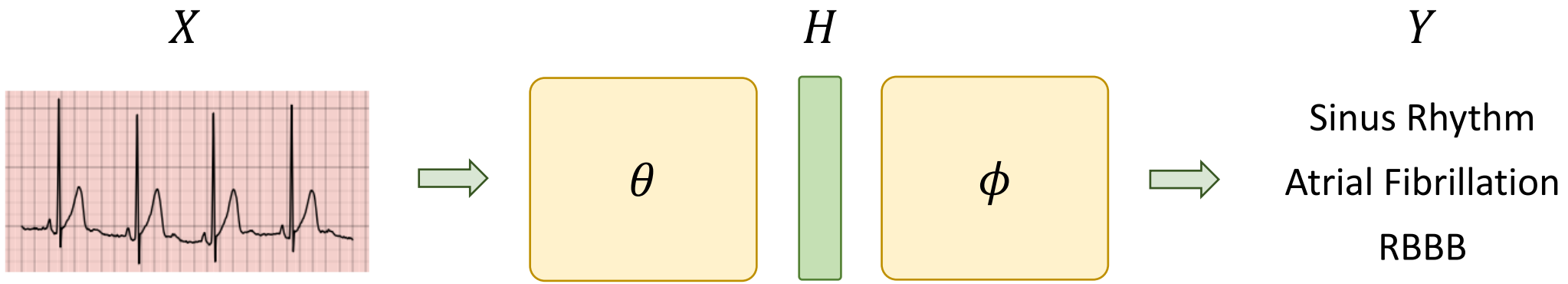
Presence of large,
unlabelled datasets

Research Question

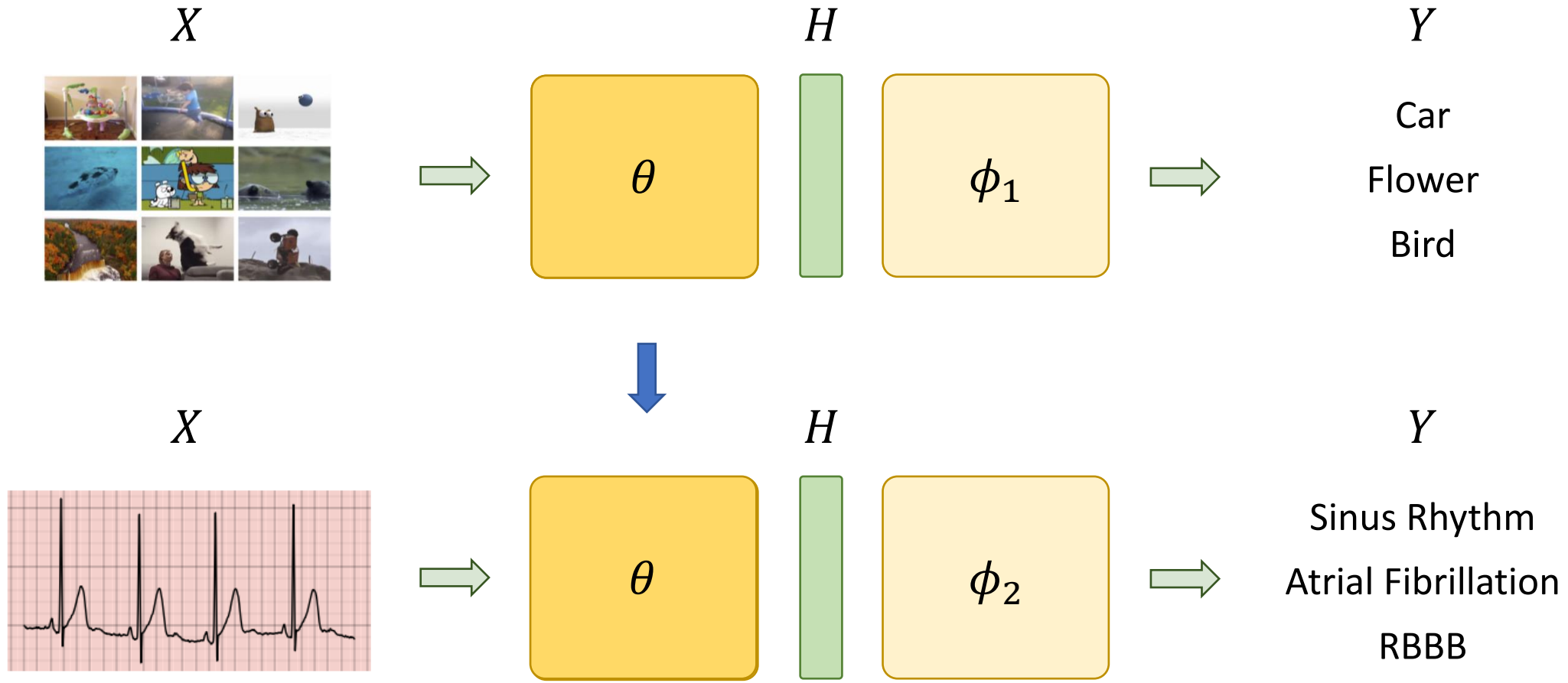
How can we derive reliable clinical insights from **small, labelled datasets** while exploiting **large, unlabelled datasets**?

Cardiac Arrhythmia Classification

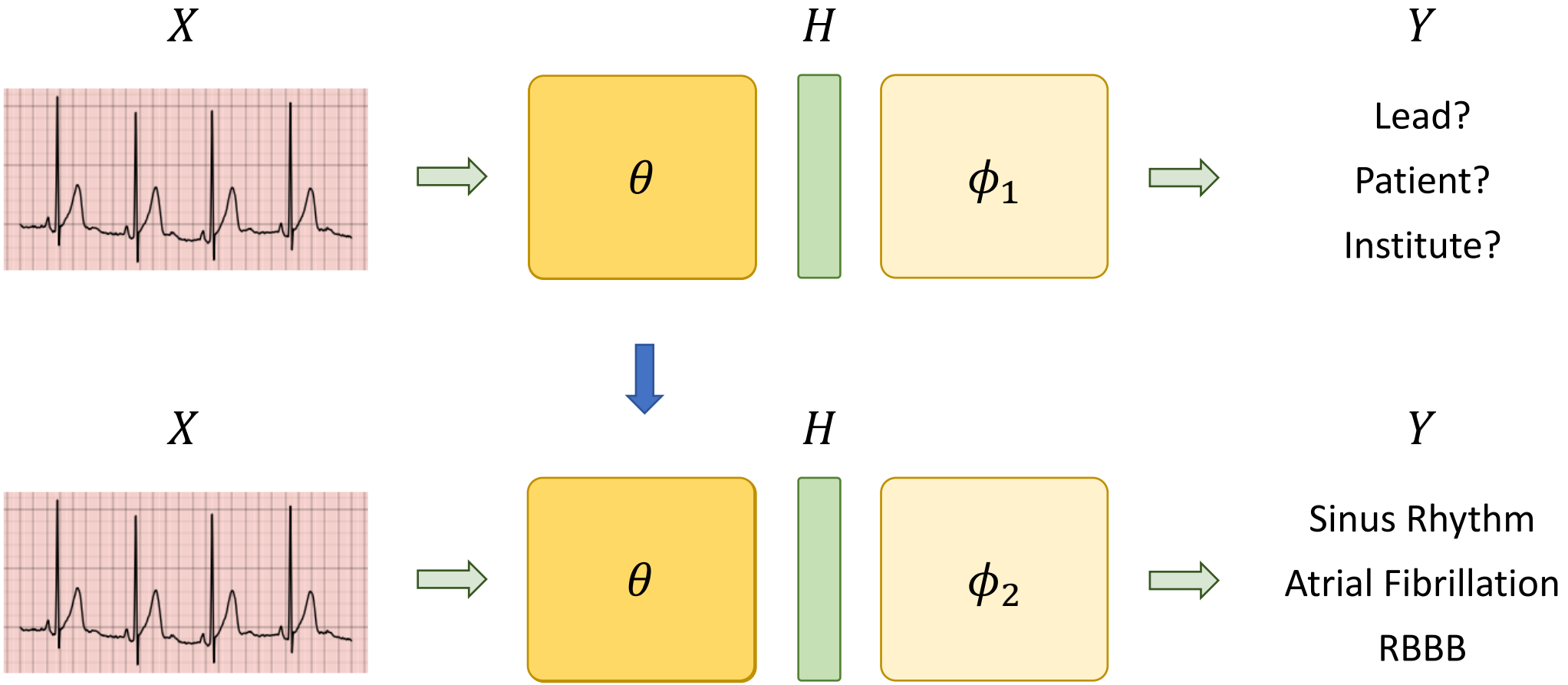
$$f_\theta : X \in R^D \rightarrow H \in R^E \qquad g_\phi : H \in R^E \rightarrow Y \in R^C$$



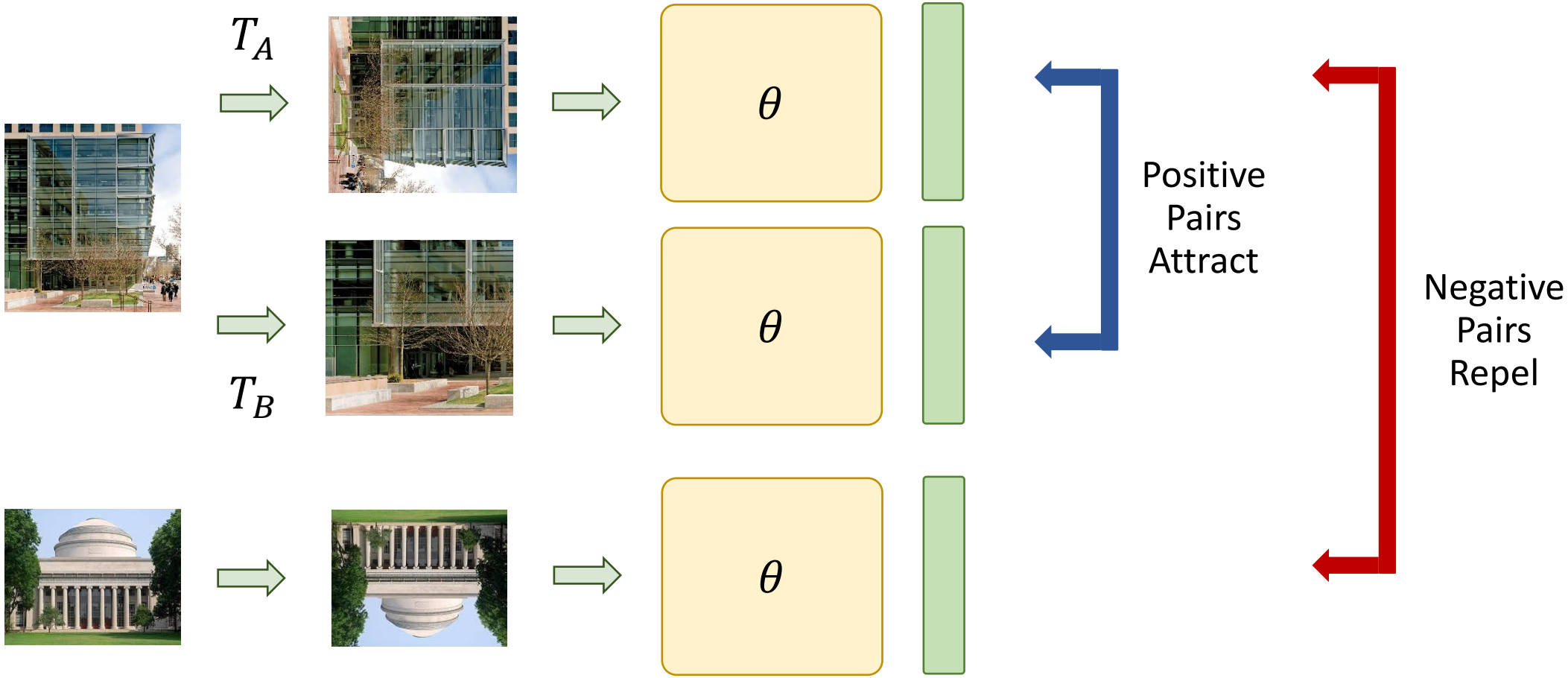
Supervised Pre-training



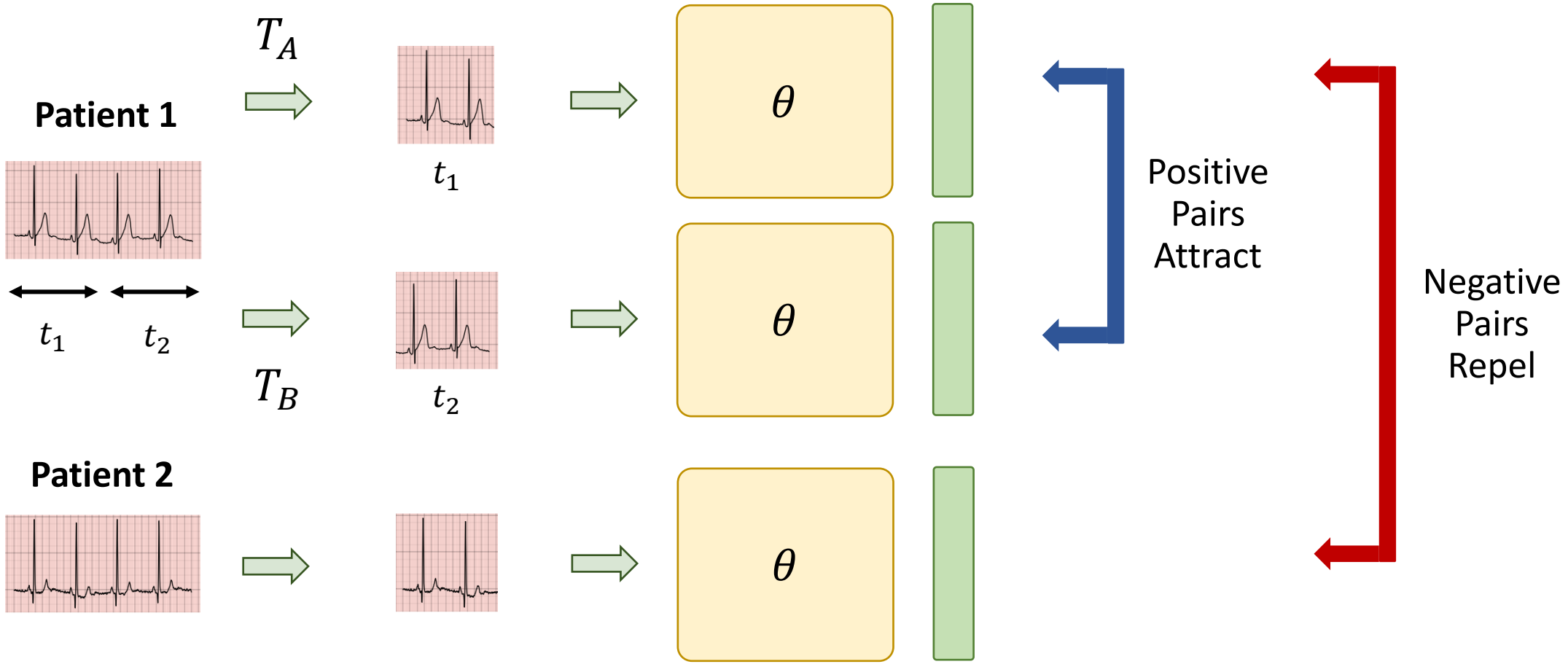
Self-supervised Pre-training



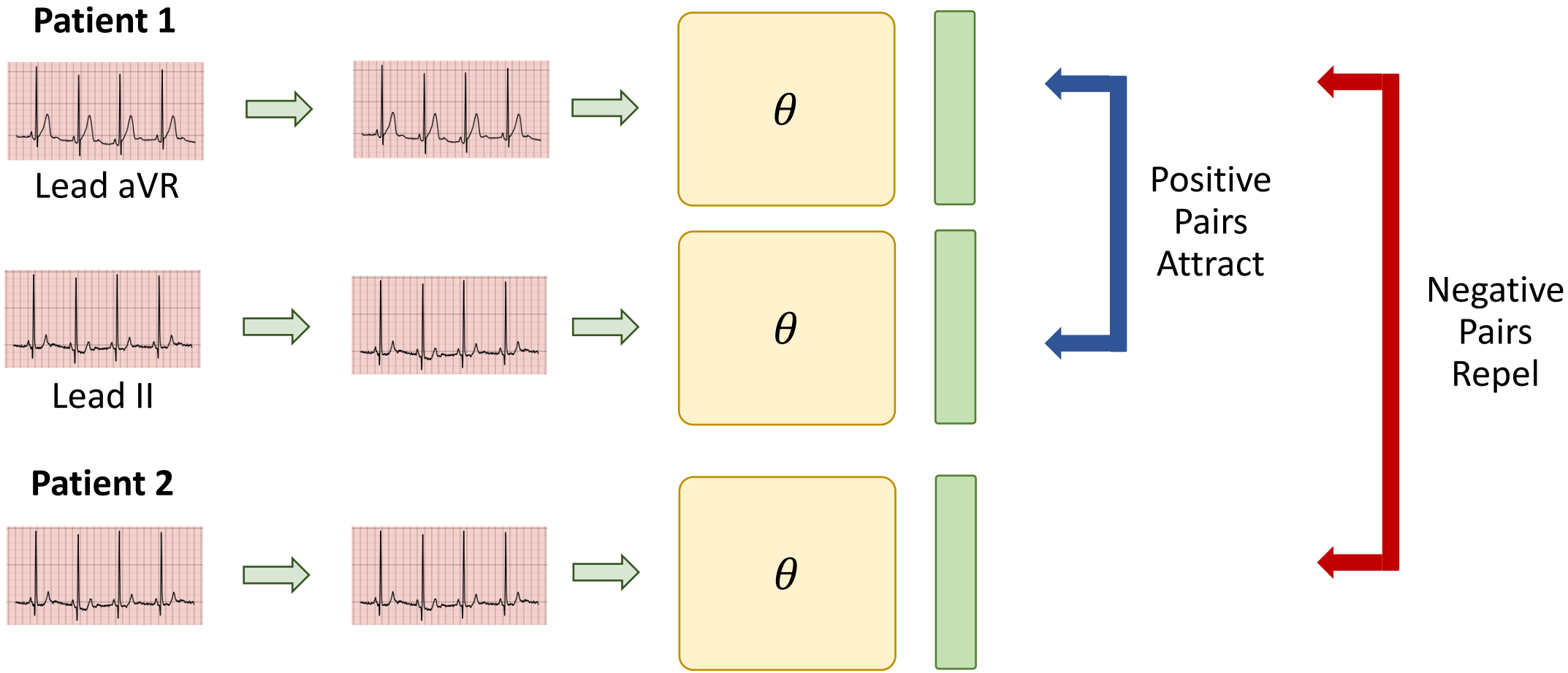
Contrastive Learning - SimCLR



CMSC – Contrastive Multi-segment Coding



CMLC – Contrastive Multi-lead Coding



Noise Contrastive Estimation (NCE) Loss

For representations, h, and two transformations, A and B

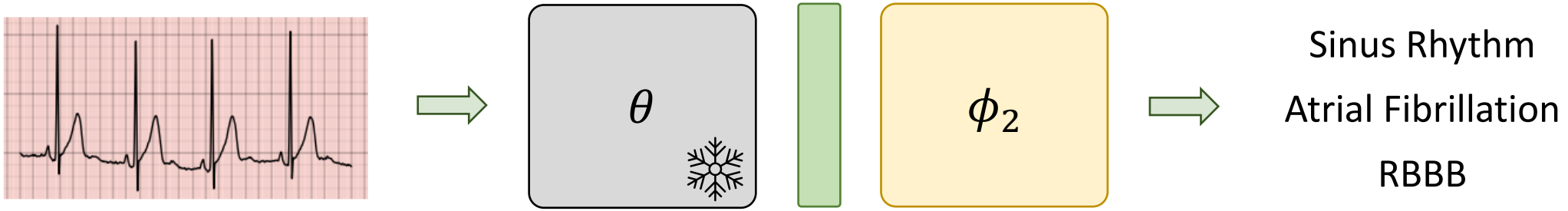
$$\mathcal{L}_{NCE} = -E_{i \in P}\left[\log\left(\frac{e^{s\left(h_A^i, h_B^i\right)}}{\sum_j e^{s\left(h_A^i, h_B^j\right)}}\right)\right]$$

Cosine-similarity of two representations

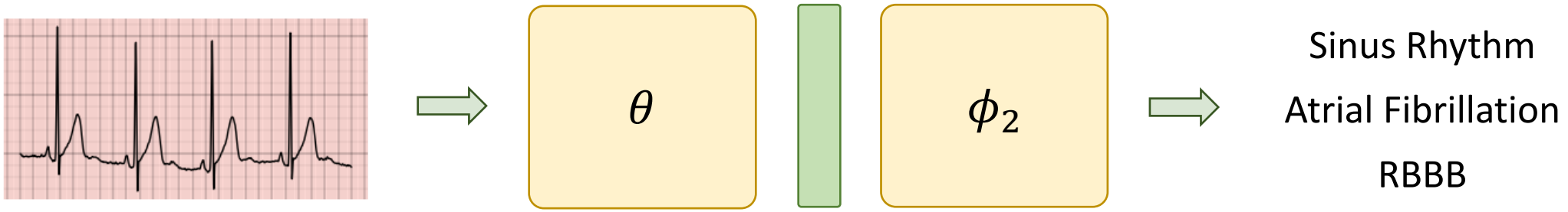
$$s\left(h_A^i, h_B^i\right) = \frac{h_A^i \cdot h_B^i}{\| h_A^i \| \| h_B^i \|} \cdot \frac{1}{\tau}$$

Evaluation Scenarios

Linear Evaluation (only update ϕ_2)



Transfer Capability of Representations (update both θ and ϕ_2)



Datasets

Dataset	# of Patients	# of Leads	# of Classes
PhysioNet 2020[1]	6,877	12	9
Chapman[2]	10,646	12	11
PhysioNet 2017[3]	8,528 (recordings)	1	4
Cardiology[4]	328	1	12

Always used for self-supervised pre-training

Predominantly used for evaluation

[1] Perez Alday, E. A., et al. (2020). Classification of 12-lead ECGs: the PhysioNet - Computing in Cardiology Challenge 2020 (version 1.0.0). *PhysioNet*.

[2] Zheng, J. et al. (2020). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Scientific Data, 7(1):1–8.

[3] Clifford, G. et al. (2017). Afib classification from a short single lead ECG recording: the physionet/computing in cardiology challenge 2017. In 2017 Computing in Cardiology, pages 1–4.

[4] Hannun, A. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine, 25(1):65.

Linear Evaluation

- CMSC outperforms CMLC and CMSMLC
- CMSC outperforms state-of-the-art SimCLR by 15.8% on Chapman

Table 1: Test AUC of the linear evaluation of the representations at $F = 0.5$, after having pre-trained on Chapman or PhysioNet 2020 with $E = 128$. Pre-training and evaluating multi-lead datasets* using 4 leads (II, V2, aVL, aVR). Mean and standard deviation are shown across 5 seeds.

Dataset	Chapman*	PhysioNet 2020*
SimCLR	0.738 ± 0.034	0.615 ± 0.014
CMSC	$\mathbf{0.896 \pm 0.005}$	$\mathbf{0.715 \pm 0.033}$
CMLC	0.870 ± 0.022	0.596 ± 0.008
CMSMLC	0.847 ± 0.024	0.680 ± 0.008

Transfer Capabilities of Representations

- CMSC outperforms state-of-the-art SimCLR in 5/6 of all experiments

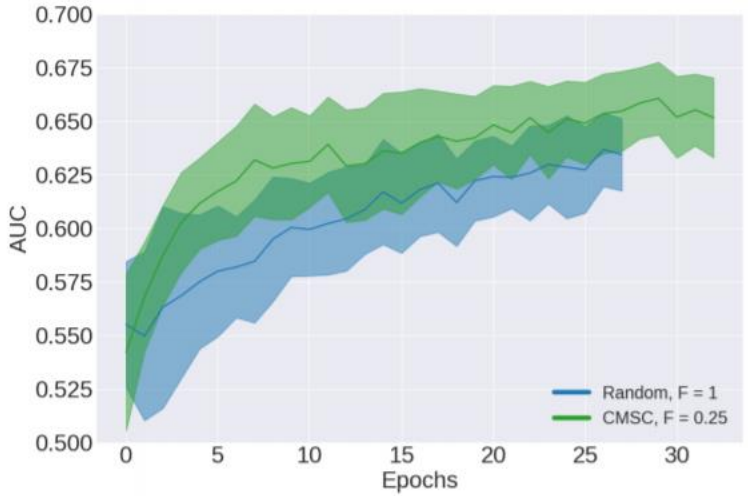
- Greatest benefit observed for small datasets, e.g., Cardiology, PhysioNet 2017

Table 2: Test AUC in the fine-tuning scenario at $F = 0.5$, after having pre-trained on Chapman or PhysioNet 2020 with $E = 128$. Pre-training, fine-tuning, and evaluating multi-lead datasets* using 4 leads. Mean and standard deviation are shown across 5 seeds.

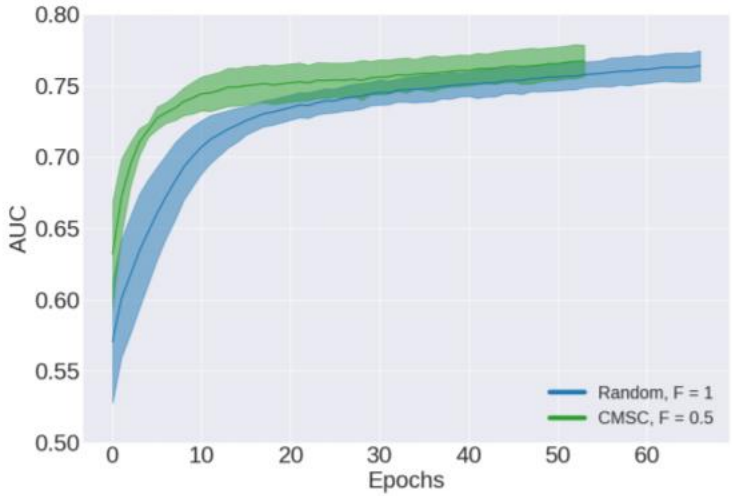
Pretraining Dataset	Chapman*			PhysioNet 2020*		
Downstream Dataset	Cardiology	PhysioNet 2017	PhysioNet 2020*	Cardiology	PhysioNet 2017	Chapman*
Random Init	0.678 ± 0.011	0.763 ± 0.005	0.803 ± 0.008	0.678 ± 0.011	0.763 ± 0.005	0.907 ± 0.006
SimCLR	0.676 ± 0.011	0.772 ± 0.010	0.823 ± 0.011	0.658 ± 0.027	0.762 ± 0.009	0.923 ± 0.010
CMSC	0.695 ± 0.024	0.773 ± 0.013	$\mathbf{0.830 \pm 0.002}$	$\mathbf{0.714 \pm 0.014}$	0.7600 ± 0.013	$\mathbf{0.932 \pm 0.008}$
CMLC	0.665 ± 0.016	0.767 ± 0.013	0.810 ± 0.011	0.675 ± 0.013	0.762 ± 0.007	0.910 ± 0.012
CMSMLC	$\mathbf{0.717 \pm 0.006}$	$\mathbf{0.774 \pm 0.004}$	0.814 ± 0.009	0.698 ± 0.011	$\mathbf{0.774 \pm 0.012}$	0.930 ± 0.012

Doing More with Less

- CMSC outperforms a random initialization with 1/4th the number of labels



(a) PhysioNet 2020 → Cardiology

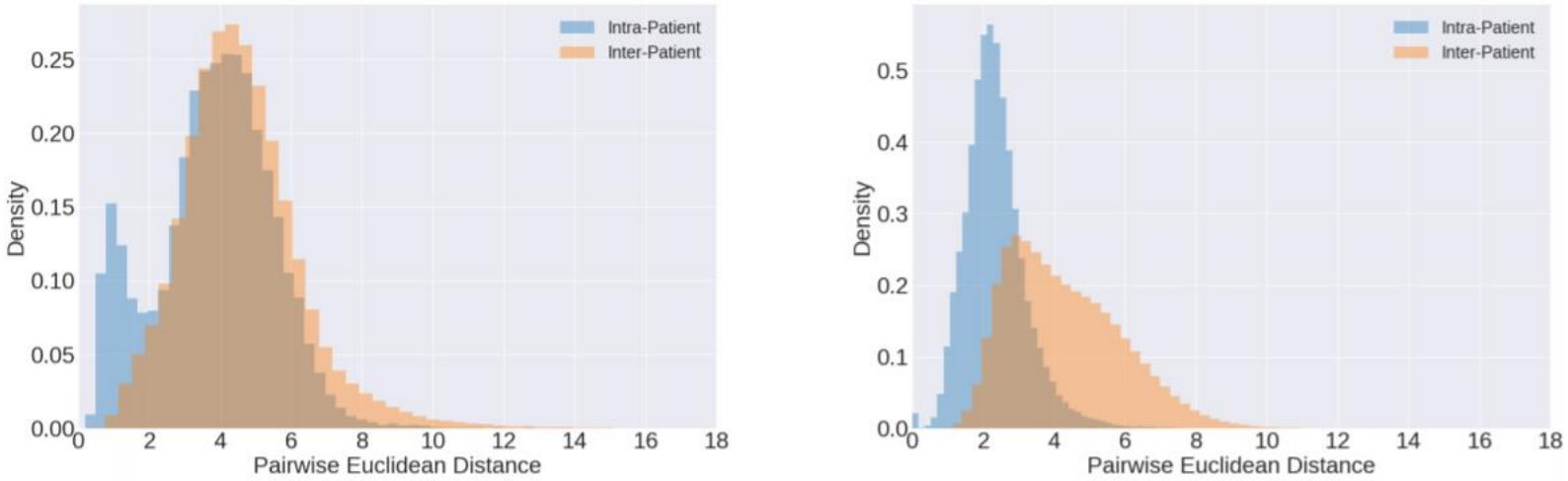


(b) Chapman → PhysioNet 2017

Figure 3: Validation AUC starting with a random and self-supervised initialization exposed to different amounts of labelled training data. Results are averaged across 5 seeds. Shaded area represents one standard deviation.

Patient-specific Representations

- CMSC learns representations that are unique to each patient



(a) SimCLR (b) CMSC

Figure 5: Distribution of pairwise Euclidean distance between representations ($E = 128$) belonging to the same patient and those belonging to different patients. Self-supervision was performed on PhysioNet 2020. Notice the lower average intra-patient values and improved separability with CMSC.

Takeaways

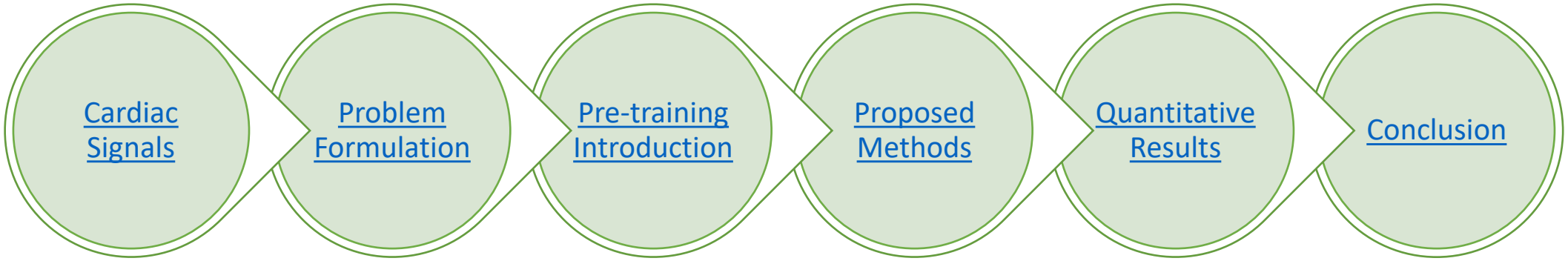
- Self-supervised pre-training with CLOCS results in **patient-specific representations**
- CLOCS pre-training can **improve the label-efficiency** of learning on downstream medical tasks 4x
- CLOCS pre-training is most beneficial when downstream datasets are small and lack sufficient labels

Future Work

- **Quantification of patient similarity** for educational purposes and disease discovery

- **Exploiting additional cardiac signals** for self-supervised learning before transferring to ECG datasets

Glossary


Cardiac Signals
Problem Formulation
Pre-training Introduction
Proposed Methods
Quantitative Results
Conclusion