Scaling Wearable Foundation Models

Girish Narayanswamy*, Xin Liu*, Kumar Ayush, Yuzhe Yang, Xuhai Xu, Shun Liao, Jake Garrison, Shyam Tailor, Jake Sunshine, Yun Liu, Tim Althoff, Shrikanth Narayanan, Pushmeet Kohli, Jiening Zhan, Mark Malhotra, Shwetak Patel, Samy Abdel-Ghaffar, Daniel McDuff*

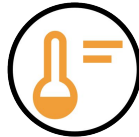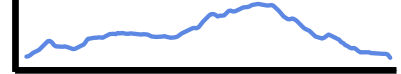* Corresponding: girishvn@uw.edu, {xliucs, dmcduff}@google.com

Google Research
Google DeepMind

Background

Wearables Health Monitors

Wearables Health Monitors
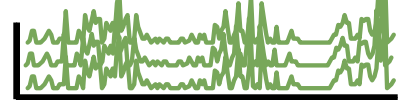
Wearables at Scale
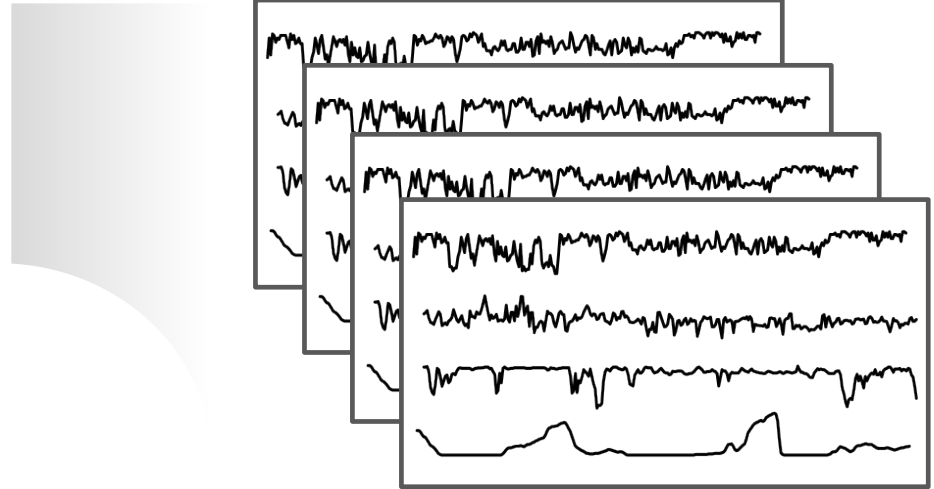
Ubiquity of wearables

Wearables at Scale

Ubiquity of wearables

Unlabeled sensor data

Self-Supervised Learning for Sensor Data

Large-scale Training of Foundation Models for Wearable Biosignals

Salar Abbaspourazad, Oussama Elachqar, Andrew Miller, Saba Emrani, Udhyakumar Nallasamy, Ian Shapiro

Published: 16 Jan 2024, Last Modified: 20 Apr 2024 · ICLR 2024 poster · Everyone · Revisions · BibTeX

RelCon: Relative Contrastive Learning for a Motion Foundation Model for Wearable Data

Maxwell A Xu, Jaya Narain, Gregory Darnell, Haraldur T Hallgrimsson, Hyewon Jeong, Darren Forde, Richard Andres Fineman, Karthik Jayaraman Raghuram, James Matthew Rehg, Shirley You Ren

Published: 22 Jan 2025, Last Modified: 15 Mar 2025 · ICLR 2025 Poster · Everyone · Revisions · BibTeX · CC BY 4.0

SleepFM: Multi-modal Representation Learning for Sleep Across Brain Activity, ECG and Respiratory Signals

Rahul Thapa, Bryan He, Magnus Ruud Kjaer, Hyatt Moore IV, Gauri Ganjoo, Emmanuel Mignot, James Zou

Published: 01 May 2024, Last Modified: 24 Jun 2024 · ICML 2024 Poster · Everyone · Revisions · BibTeX · CC BY 4.0

PaPaGei: Open Foundation Models for Optical Physiological Signals

Arvind Pillai, Dimitris Spathis, Fahim Kawsar, Mohammad Malekzadeh

Advancing Location–Invariant and Device–Agnostic Motion Activity Recognition on Wearable Devices

Rebecca Adaimi, Abdelkareem Bedri, Jun Gong, Richard Kang, Joanna Arreaza–Taylor, Gerri–Michelle Pascual, Michael Ralph, Gierad Laput

Self-supervised learning for human activity recognition using 700,000 person-days of wearable data

Hang Yuan, Shing Chan, Andrew P. Creagh, Catherine Tong, Aidan Acquah, David A. Clifton & Aiden Doherty ✉

npj Digital Medicine **7**, Article number: 91 (2024) | Cite this article

Self-Supervised Pretraining and Transfer Learning Enable Flu and COVID-19 Predictions in Small Mobile Sensing Datasets

Mike A. Merrill, Tim Althoff

Published: 31 Dec 2022, Last Modified: 05 Sept 2024 · CHIL 2023 · Everyone · Revisions · BibTeX · CC BY-SA 4.0

Scaling Laws

Scaling Laws for Neural Language Models

Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, Dario Amodei

Deep Learning Scaling is Predictable, Empirically

Joel Hestness, Sharan Narang, Newsha Ardalani, Gregory Diamos, Heewoo Jun, Hassan Kianinejad, Md. Mostofa Ali Patwary, Yang Yang, Yanqi Zhou

On Data Scaling in Masked Image Modeling

Zhenda Xie, Zheng Zhang, Yue Cao, Yutong Lin, Yixuan Wei, Qi Dai, Han Hu; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 10365-10374

Scaling Vision Transformers

Xiaohua Zhai, Alexander Kolesnikov, Neil Houlsby, Lucas Beyer; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 12104-12113

Scaling Laws for Generative Mixed-Modal Language Models

Armen Aghajanyan, Lili Yu, Alexis Conneau, Wei-Ning Hsu, Karen Hambardzumyan, Susan Zhang, Stephen Roller, Naman Goyal, Omer Levy, Luke Zettlemoyer Proceedings of the 40th International Conference on Machine Learning, PMLR 202:265-279, 2023.

Scaling Laws Visualized

Loss measure

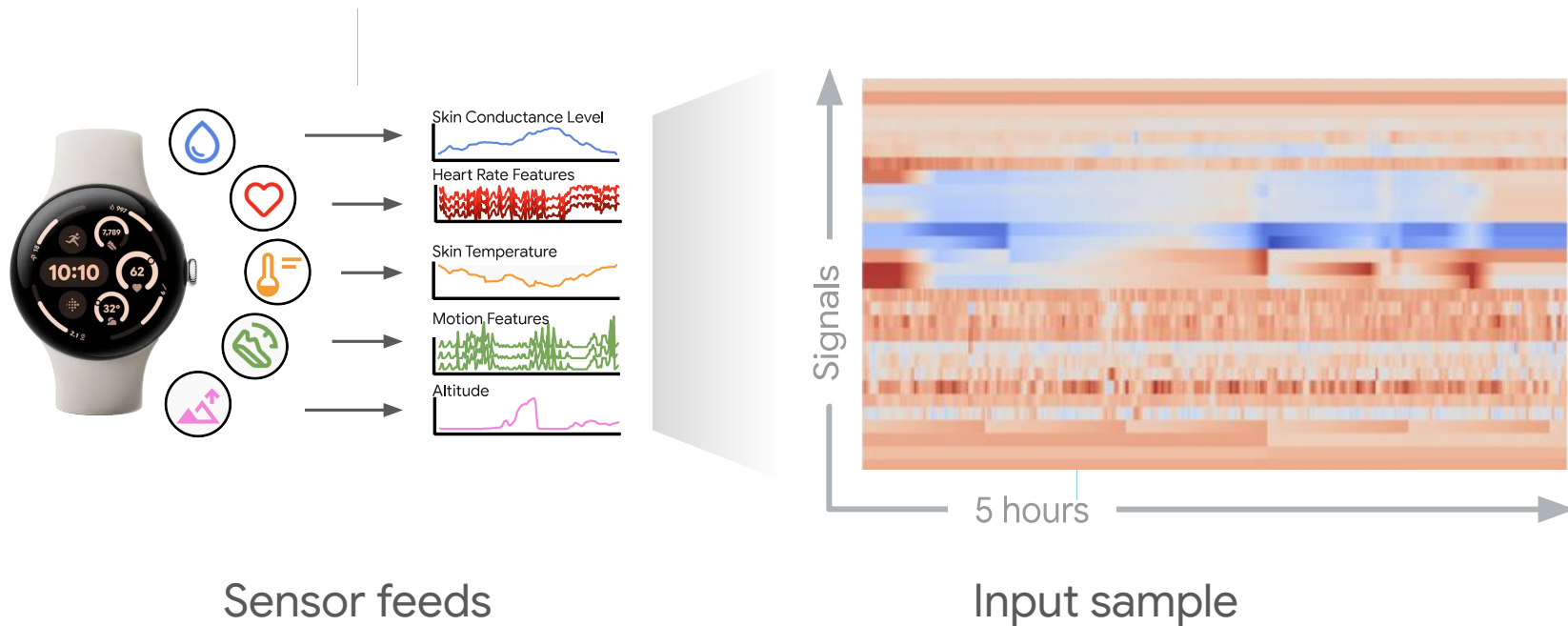Resource size
(compute / data / model parameters)
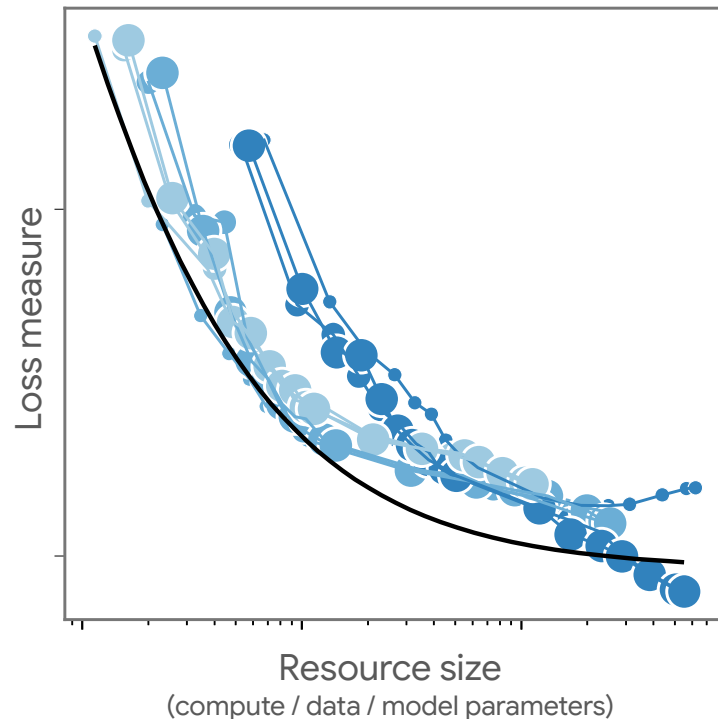
Do **scaling laws** apply to models trained on **wearable sensor data**?

Methods

Data for Large Sensor Models

40M Hours of wearable sensor data from **165K** subjects

Sensor feeds

Input sample

Data for Large Sensor Models

Study	#Subjects (000s)	#Hours (000s)	Sensors						Generative
			ECG	PPG	ACC	SCL	TMP	ALT	
[1]	0.05	0.20	✗	✗	✓	✗	✗	✗	✓
[2]	141	400	✓	✓	✗	✗	✗	✗	✗
[3]	100	15,700	✗	✗	✗	✗	✗	✗	✗
LSM	165	40,000	✗	✓	✓	✓	✓	✓	✓

[1] Adaimi et al. "Advancing location-invariant and device agnostic motion activity recognition on wearable devices." (2024)
[2] Abbaspourazad et al. "Large-scale training of foundation models for wearable biosignals." (2023)
[3] Yuan et al. "Self-supervised learning for human activity recognition using 700,000 persondays of wearable data." (2024)

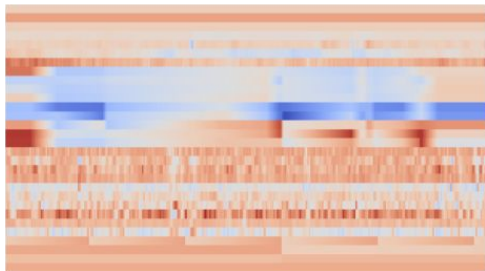# Training Large Sensor Models
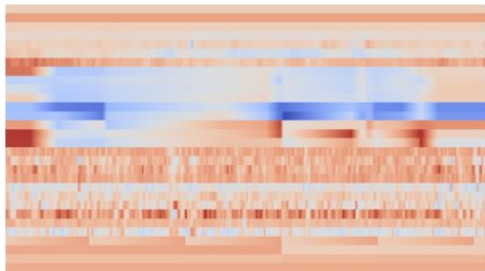
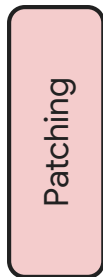**Sensor input**   **Patched Input**
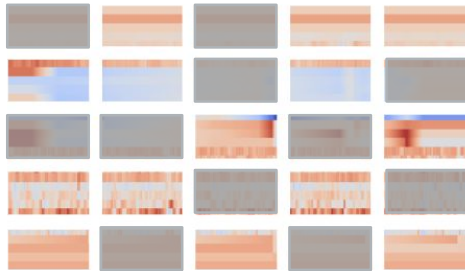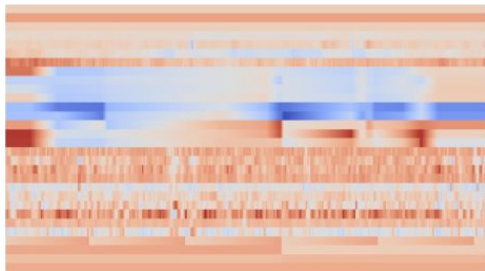
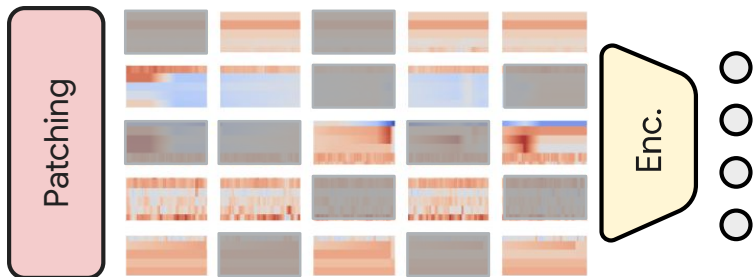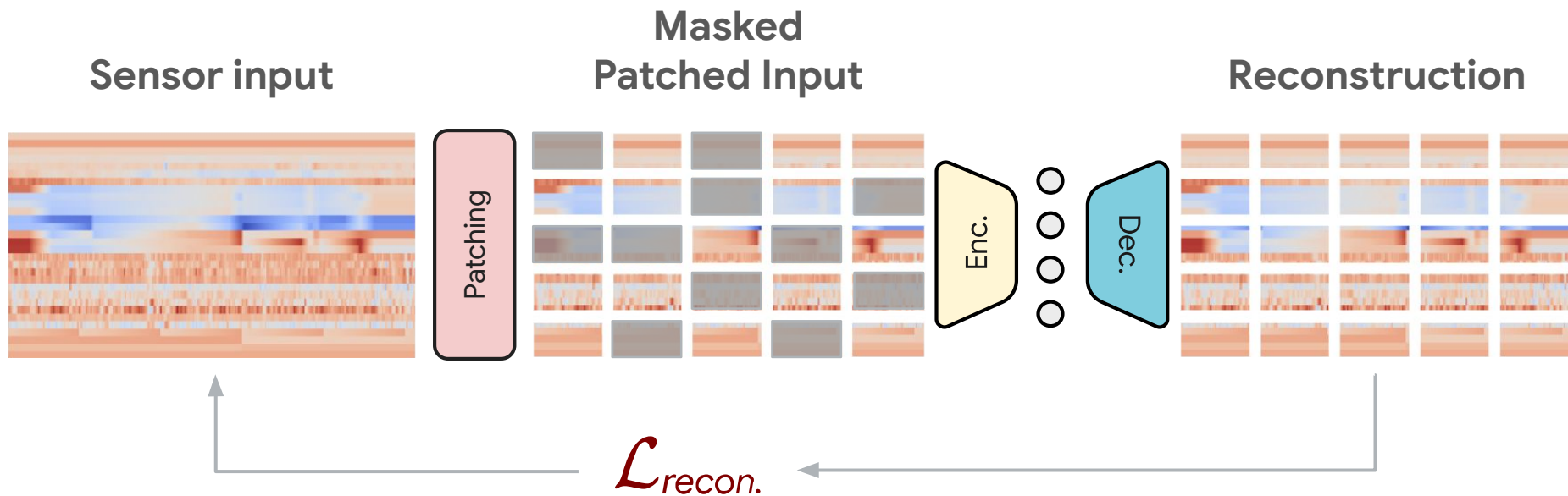# Training Large Sensor Models

Sensor input

Masked Patched Input

Patching

Training Large Sensor Models

Sensor input

Masked Patched Input

Patching

Enc.

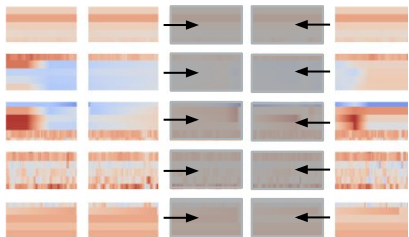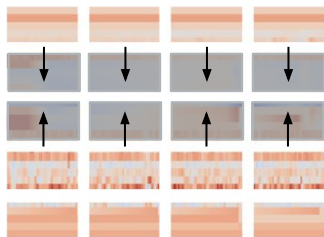# Training Large Sensor Models
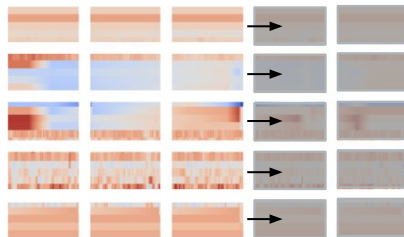
Evaluating Large Sensor Models

Generative

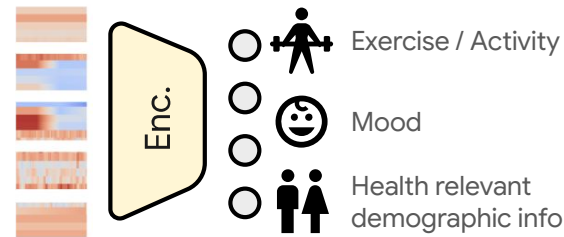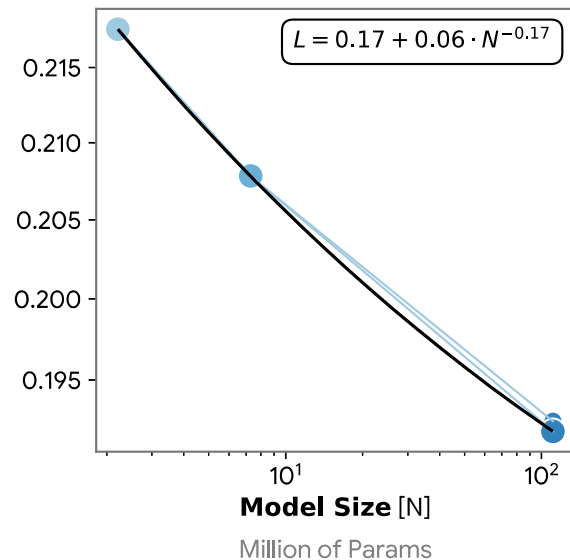Temporal Imputation
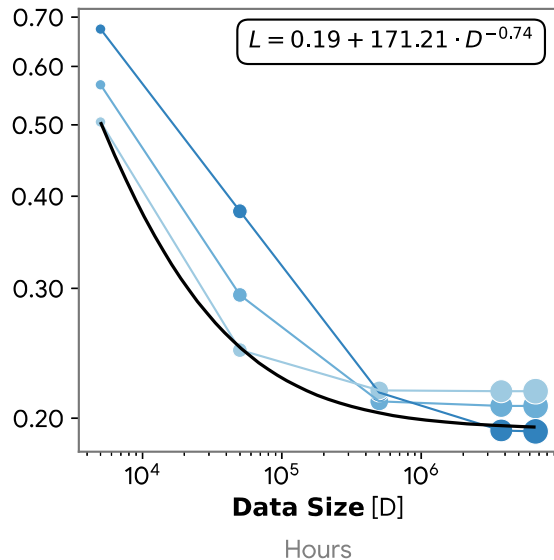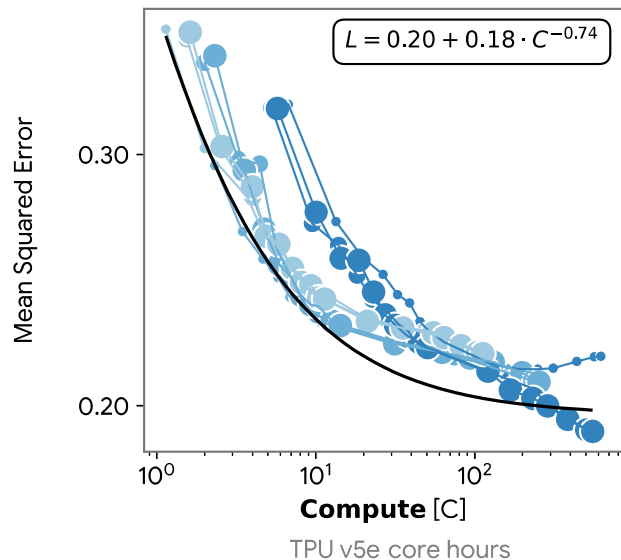
Sensor Imputation

Temporal Extrapolation

Discriminative

Detection / Classification Tasks

Enc.

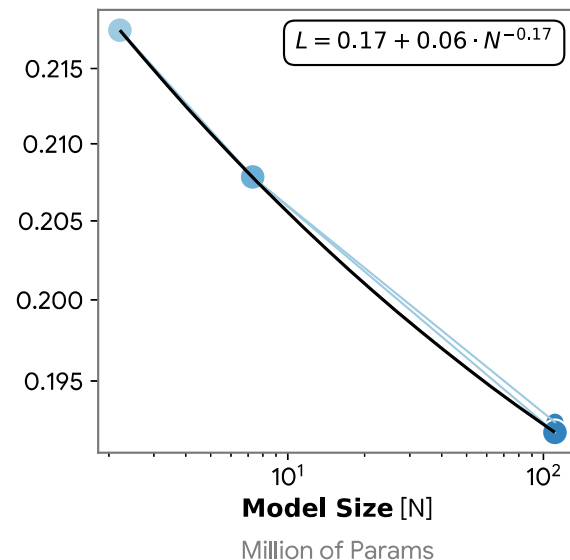Exercise / Activity

Mood

Health relevant demographic info

Results

Scaling

Figure showing three scaling plots of Mean Squared Error.

Left plot — Mean Squared Error vs Compute [C] (TPU v5e core hours):
$$L = 0.20 + 0.18 \cdot C^{-0.74}$$

Middle plot — Mean Squared Error vs Data Size [D] (Hours):
$$L = 0.19 + 171.21 \cdot D^{-0.74}$$

Right plot — Mean Squared Error vs Model Size [N] (Million of Params):
$$L = 0.17 + 0.06 \cdot N^{-0.17}$$

Data Size (Hours):
- 0.005M
- 0.05M
- 0.5M
- 3.8M
- 6.6M

Model Params:
- ViT 2M
- ViT 7M
- ViT 110M

Scaling: Compute

$L = 0.20 + 0.18 \cdot C^{-0.74}$

$L = 0.19 + 171.21 \cdot D^{-0.74}$

$L = 0.17 + 0.06 \cdot N^{-0.17}$

Mean Squared Error

Compute [C]
TPU v5e core hours

Data Size [D]
Hours

Model Size [N]
Million of Params

Data Size (Hours):
- 0.005M
- 0.05M
- 0.5M
- 3.8M
- 6.6M

Model Params:
- ViT 2M
- ViT 7M
- ViT 110M

Scaling: Data Size

$L = 0.20 + 0.18 \cdot C^{-0.74}$

$L = 0.19 + 171.21 \cdot D^{-0.74}$

$L = 0.17 + 0.06 \cdot N^{-0.17}$

Mean Squared Error
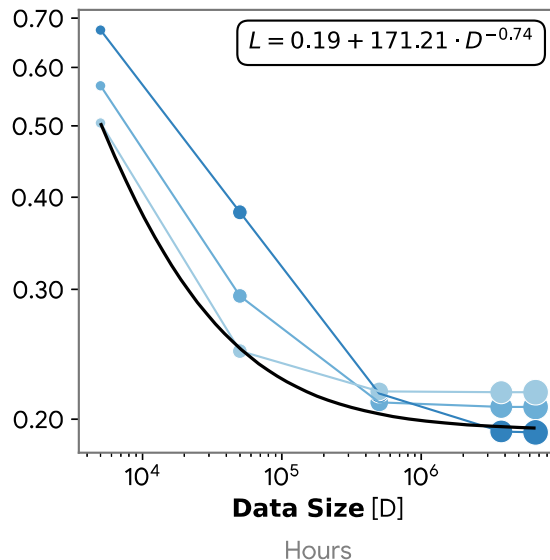
Compute [C]
TPU v5e core hours

Data Size [D]
Hours

Model Size [N]
Million of Params
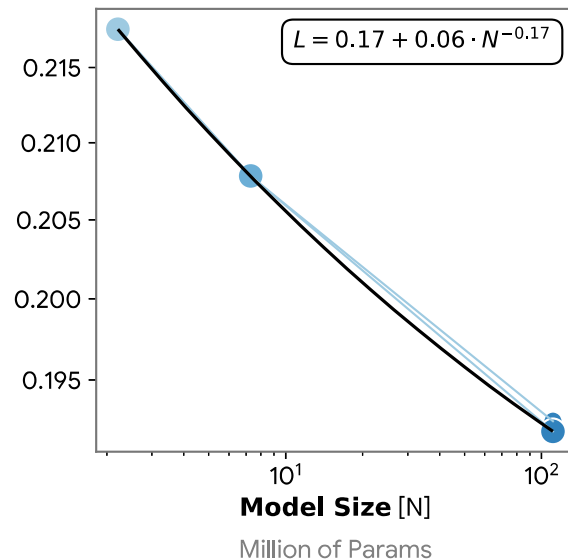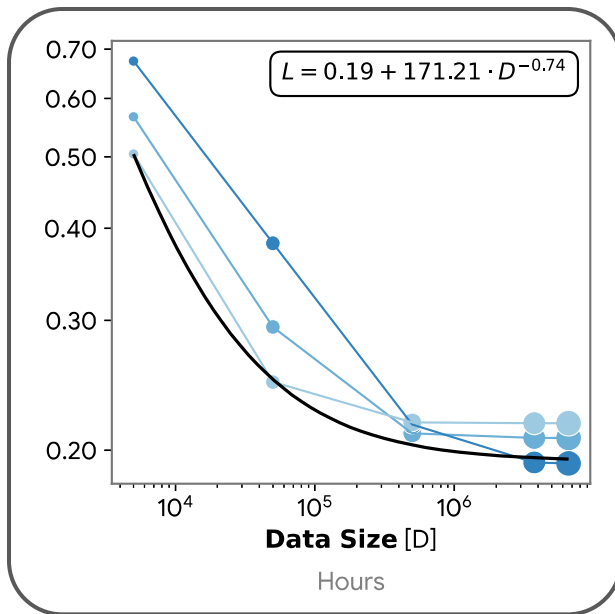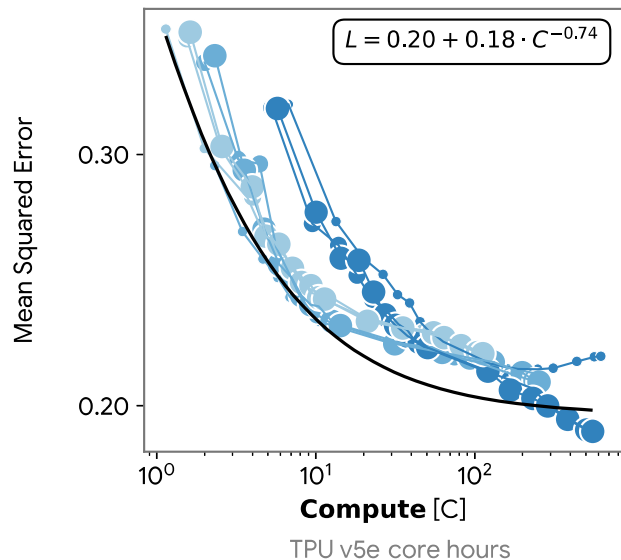
Data Size (Hours):   Model Params:
- 0.005M   ViT 2M
- 0.05M   ViT 7M
- 0.5M   ViT 110M
- 3.8M
- 6.6M

Scaling: Model Size

Generative Task: Signal Level Reconstructions

Example 1   Example 2   Example 3

Skin temp. · Heart Rate · RR 80th %ile · RR 20th %ile · RR Entropy · Step Count · Log Energy · Acc. Covariance · Acc. Mean · Acc. Kurtosis

Time (minutes)

■ Reconstructed Signal   ■ Original Signal   ☐ Masked in Input
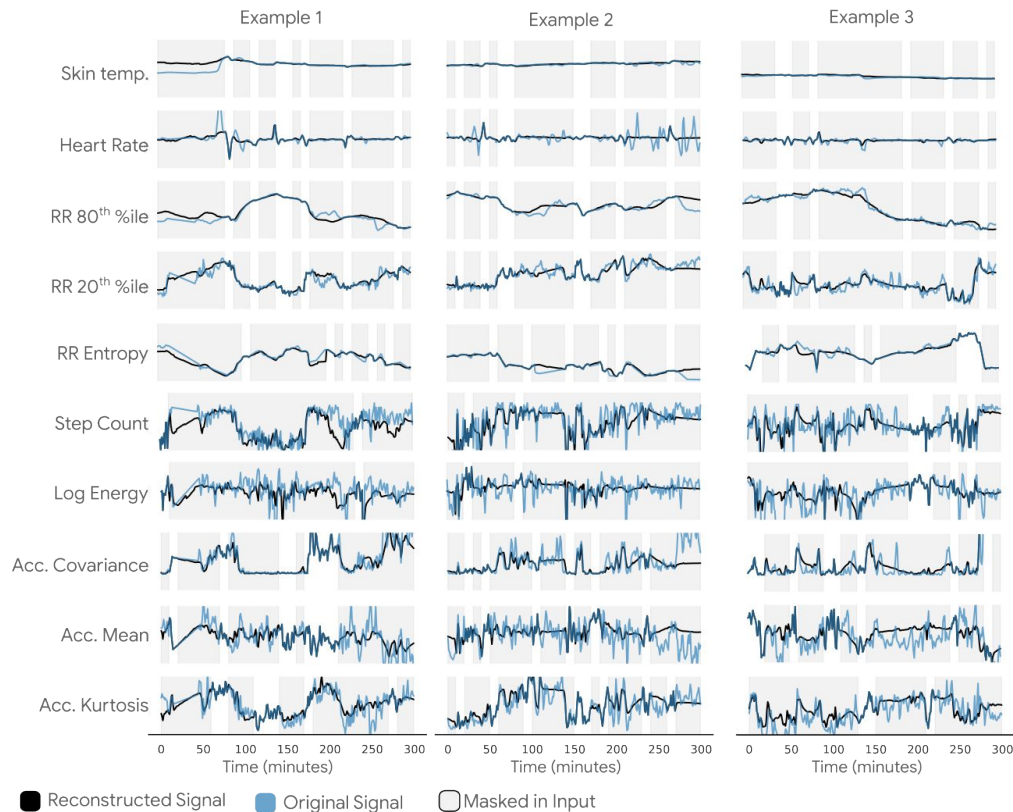
Generative Task: Image Level Reconstructions

	Sample 1	Sample 2	Sample 3
Original			
Random Imputation			
Temporal Extrapolation			
Temporal Interpolation			
Sensor Imputation			
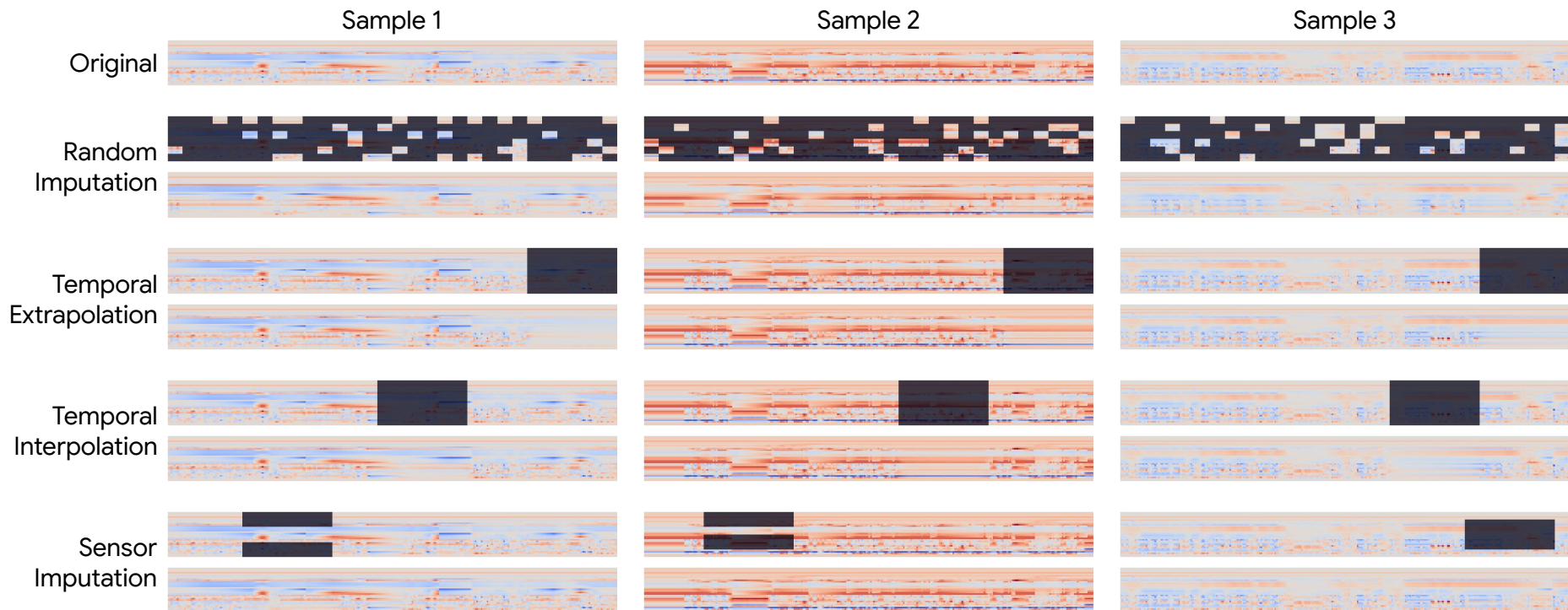
Generative Task: Comparison to Baseline

Generative Task: The Effect of Scaling

(a) Data Scaling Across Model Sizes

Random Imputation 80% (Test Loss) · Temporal Extrapolation 60 Minutes (Zero Shot) · Temporal Interpolation 60 Minutes (Zero Shot)

(b) Model Scaling Across Data Sizes
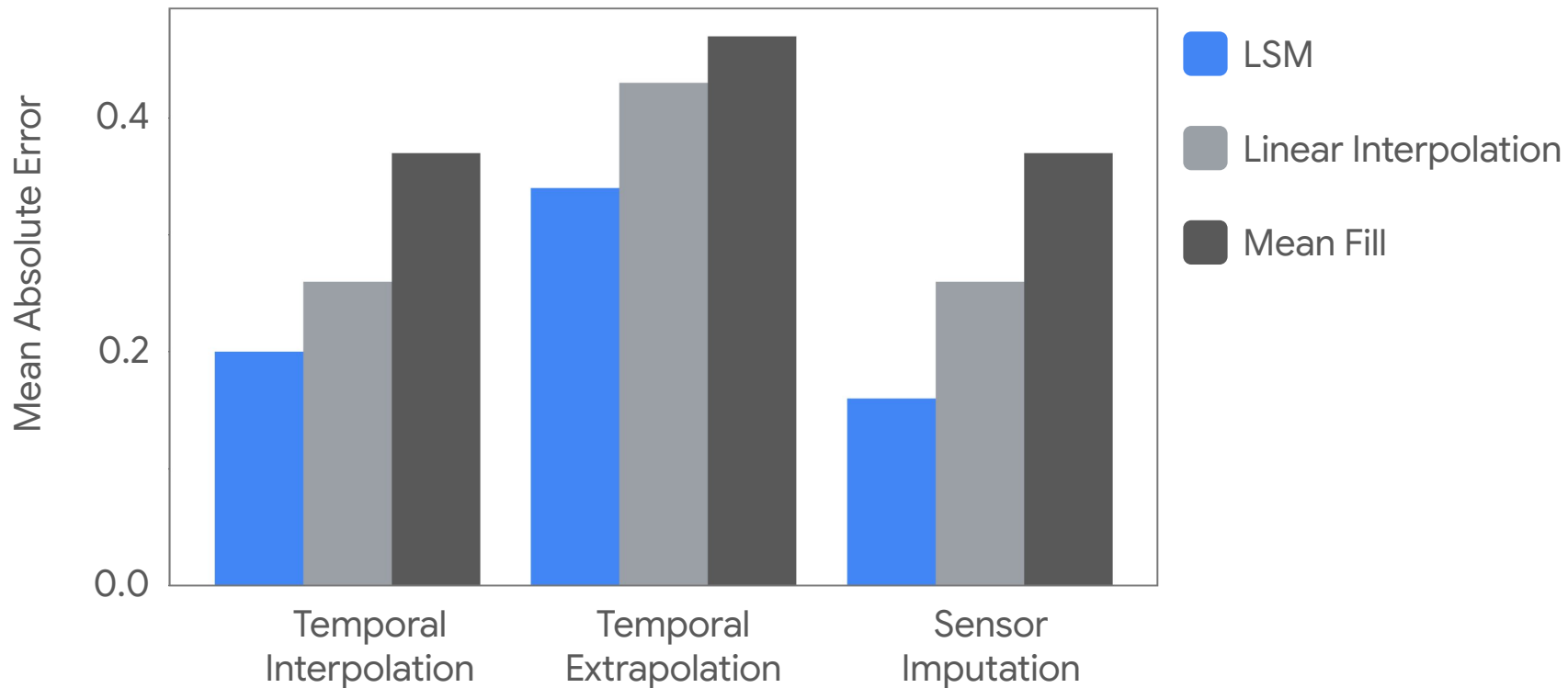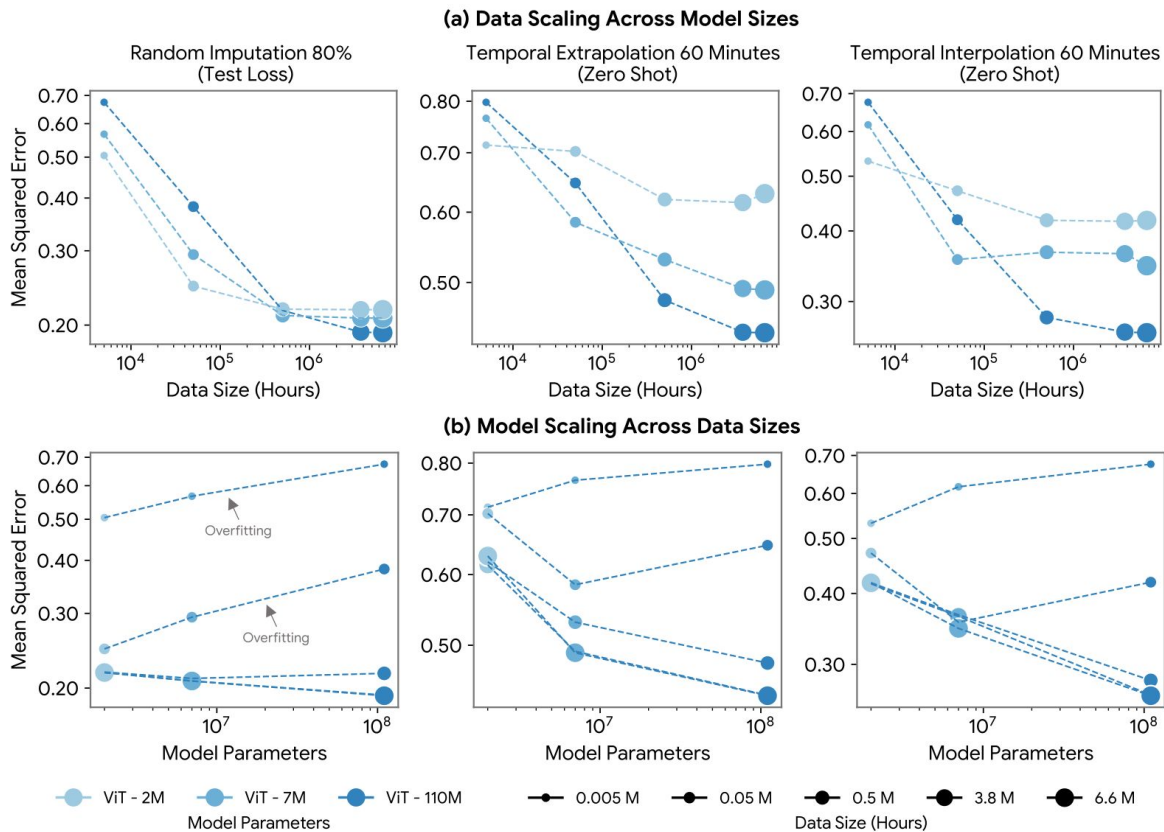
Generative Task: The Effect of Scaling

(A) Compute Scaling

	Sample 1	Sample 2	Sample 3
Original			
Masked Sample			
500 TPU v5e Core Hours			
10 TPU v5e Core Hours			

(B) Data Scaling

	Sample 1	Sample 2	Sample 3
Original			
Masked Sample			
6.6 M Hours			
0.005 M Hours			

(C) Model Scaling

	Sample 1	Sample 2	Sample 3
Original			
Masked Sample			
LSM-Base			
LSM-Tiny			
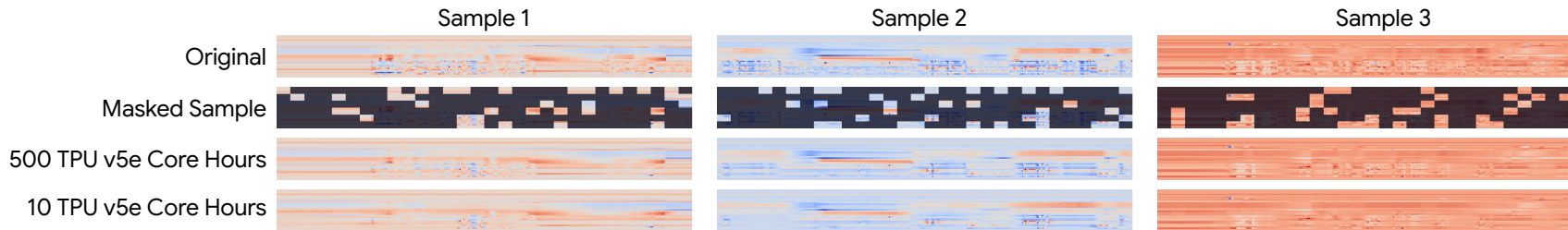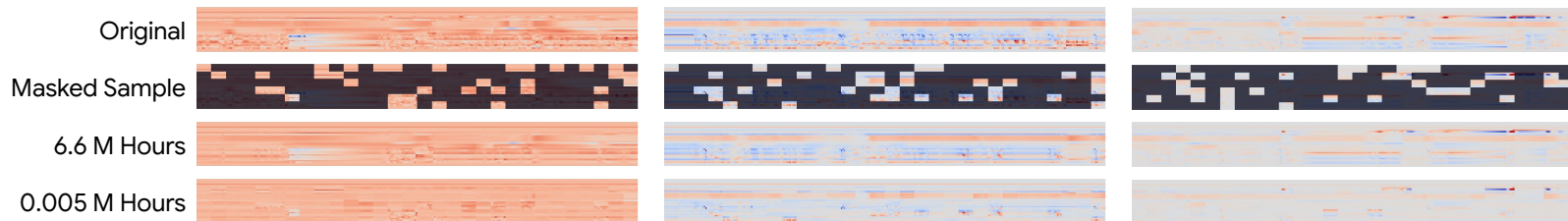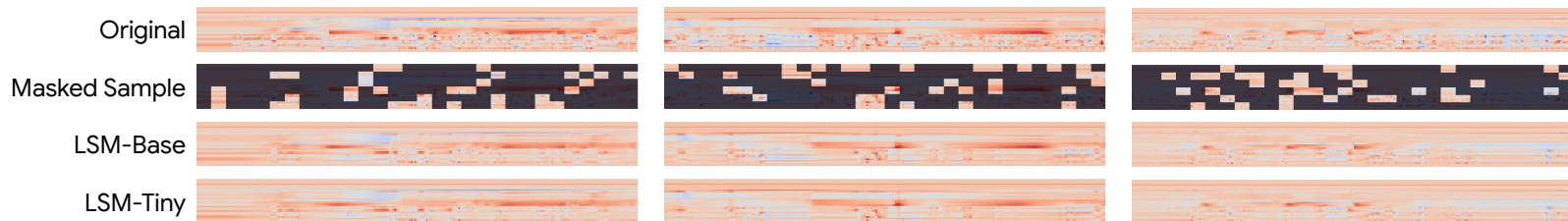
Discriminative Task: Label Efficient Learner

Few-Shot Activity Recognition

Legend:
- LSM + FT
- LSM + Linear Probe
- Baseline

Y-axis: Accuracy (%)
X-axis: Number of Shots
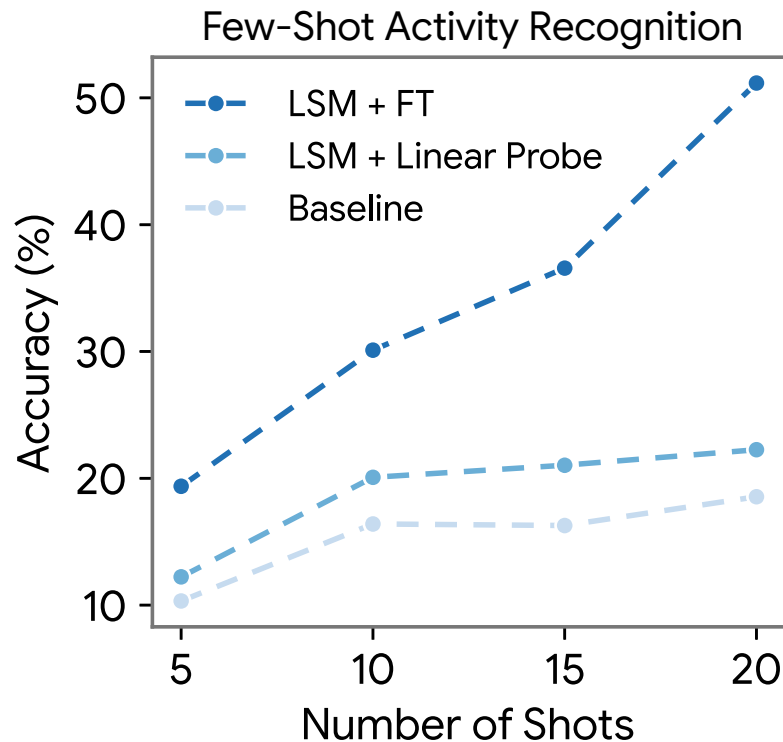
Discriminative Task: Comparison to Baseline
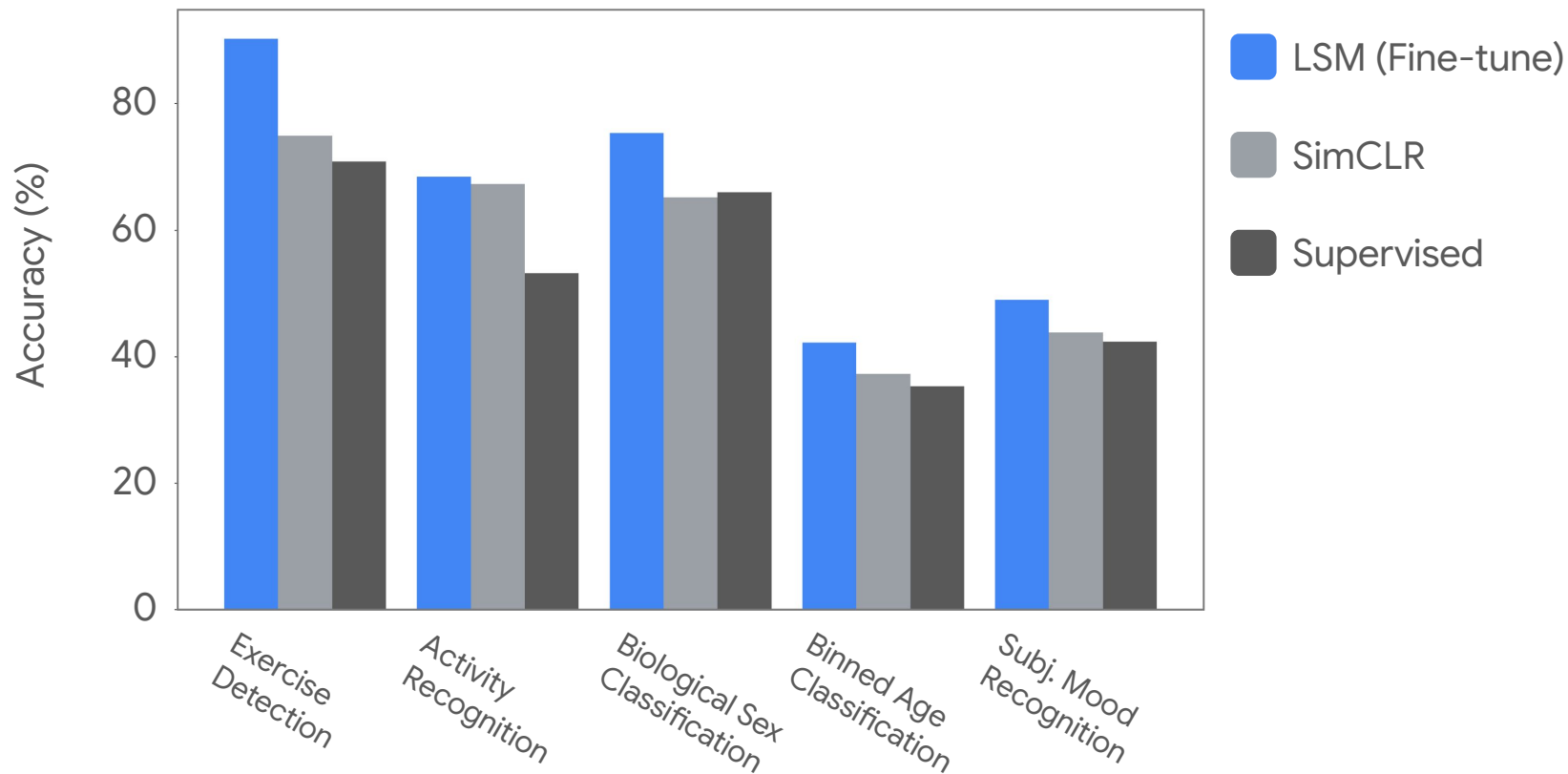
Discriminative Task: The Effect of Scaling
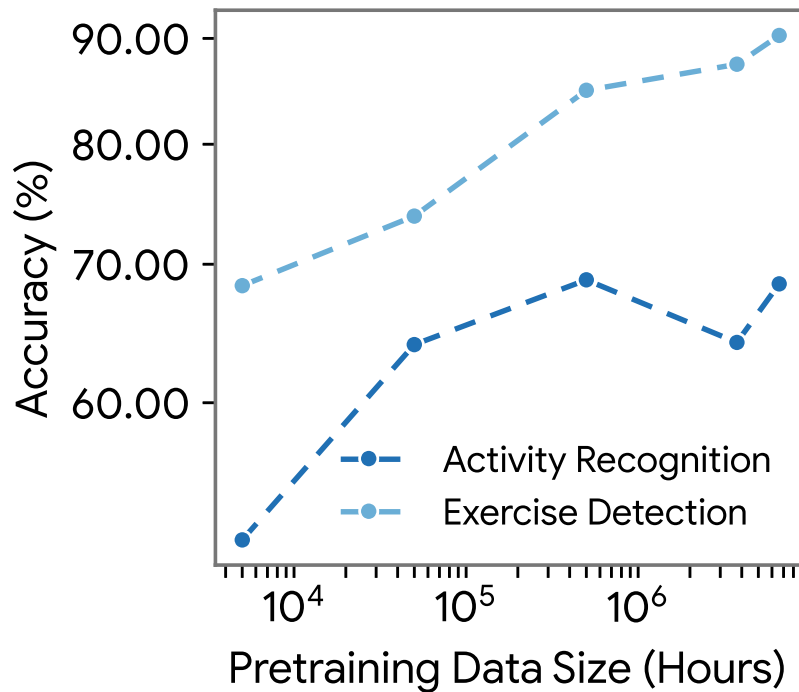
Discriminative Task: The Effect of Scaling

Conclusion

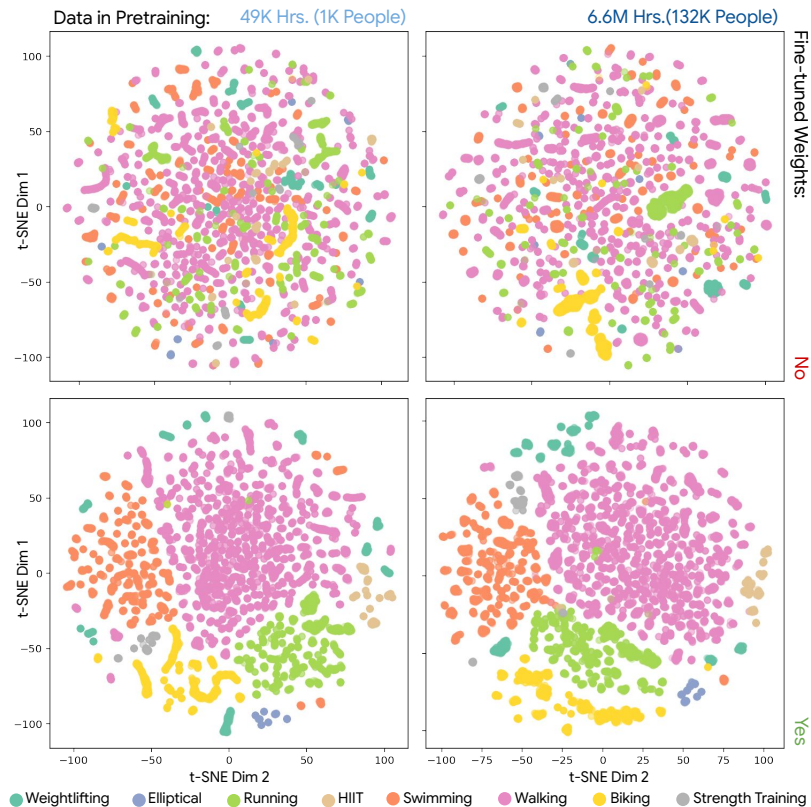# Summary of Contributions

1. Large Sensor Model (LSM)
   - First multimodal wearable sensor foundation model
2. Largest wearable modeling study to date
   - Pretrained on 40M hours of data
3. Prove scaling in the wearable sensor domain
4. Illustrate the downstream capabilities of LSM
   - Generative: imputation and forecasting
   - Discriminative: exercise, activity, mood, health related demographics

Scaling Wearable Foundation Models

Girish Narayanswamy*, Xin Liu*, Kumar Ayush, Yuzhe Yang, Xuhai Xu, Shun Liao, Jake Garrison, Shyam Tailor, Jake Sunshine, Yun Liu, Tim Althoff, Shrikanth Narayanan, Pushmeet Kohli, Jiening Zhan, Mark Malhotra, Shwetak Patel, Samy Abdel-Ghaffar, Daniel McDuff*

* Corresponding: girishvn@uw.edu, {xliucs, dmcduff}@google.com

Google Research
Google DeepMind