

NEURAL INFORMATION
PROCESSING SYSTEMS



UniMTS: Unified Pre-training for Motion Time Series

Xiyuan Zhang, Diyan Teng, Ranak Roy Chowdhury, Shuheng Li,
Dezhi Hong, Rajash K. Gupta, Jingbo Shang

University of California, San Diego

Introduction

Background

- **Motion Time Series Classification** (or Human Activity Recognition) identifies human activities using time series readings from wearable devices
- Applications: healthcare, motion tracking, smart home automation, etc.



Healthcare



Motion Tracking

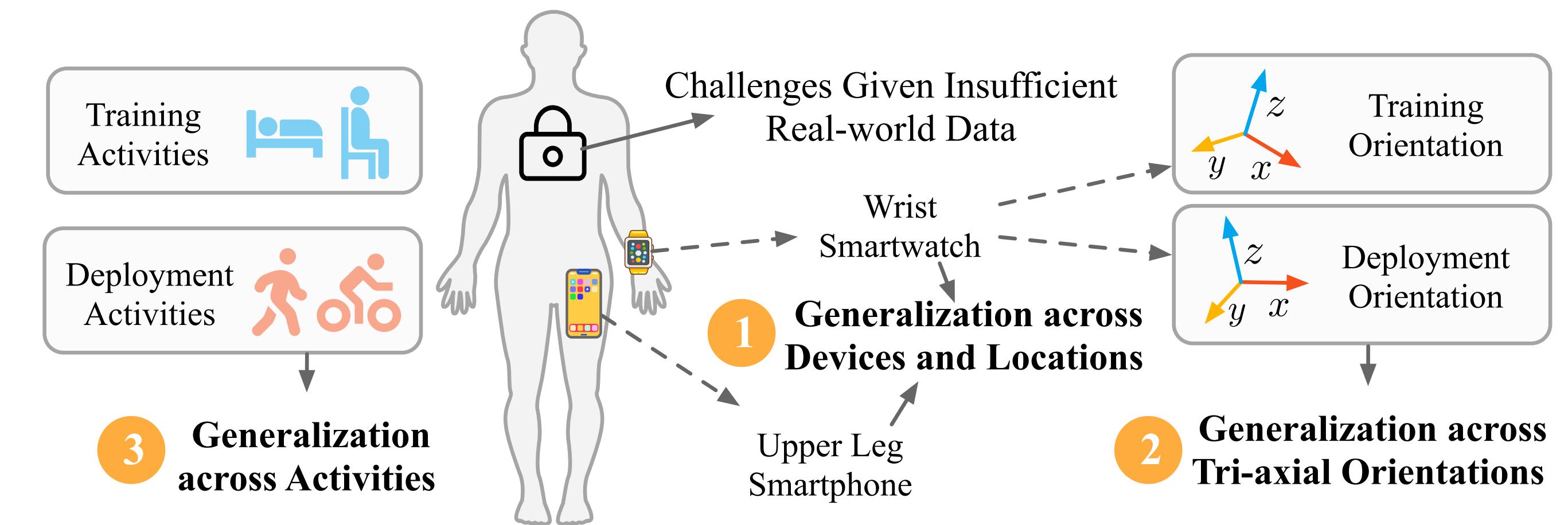


Smart Home Automation

Introduction

Challenges

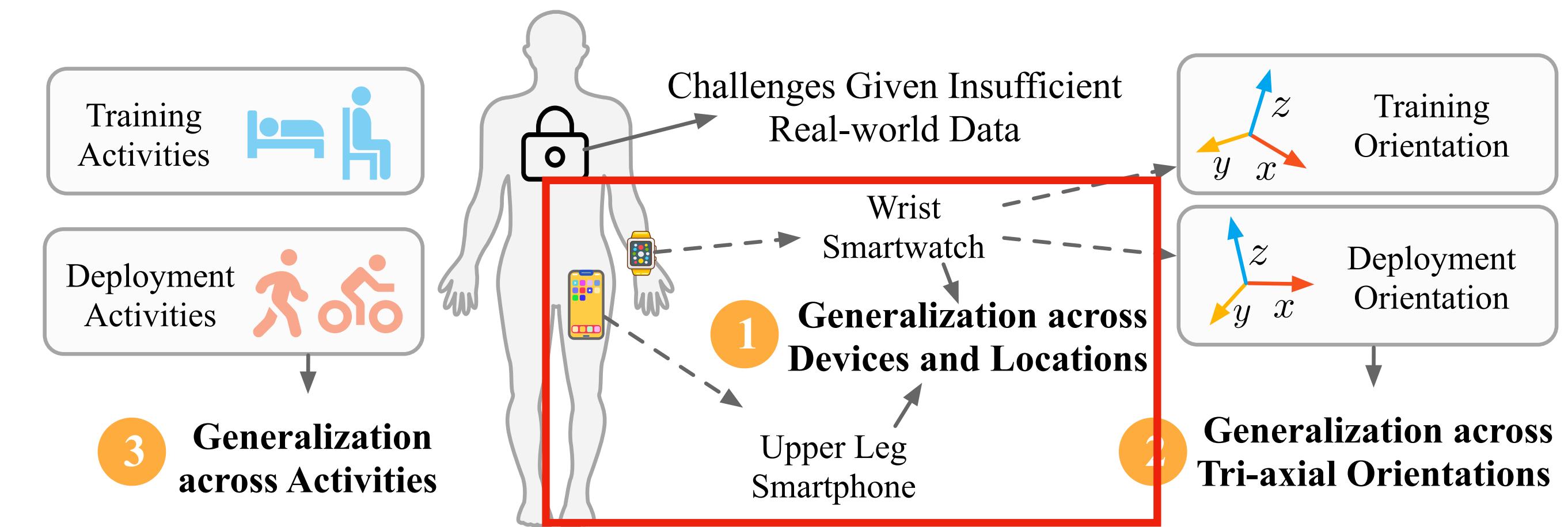
- Insufficient training samples
- Do not generalize across
 - Devices or Locations
 - Orientations
 - Activities



Introduction

Challenges

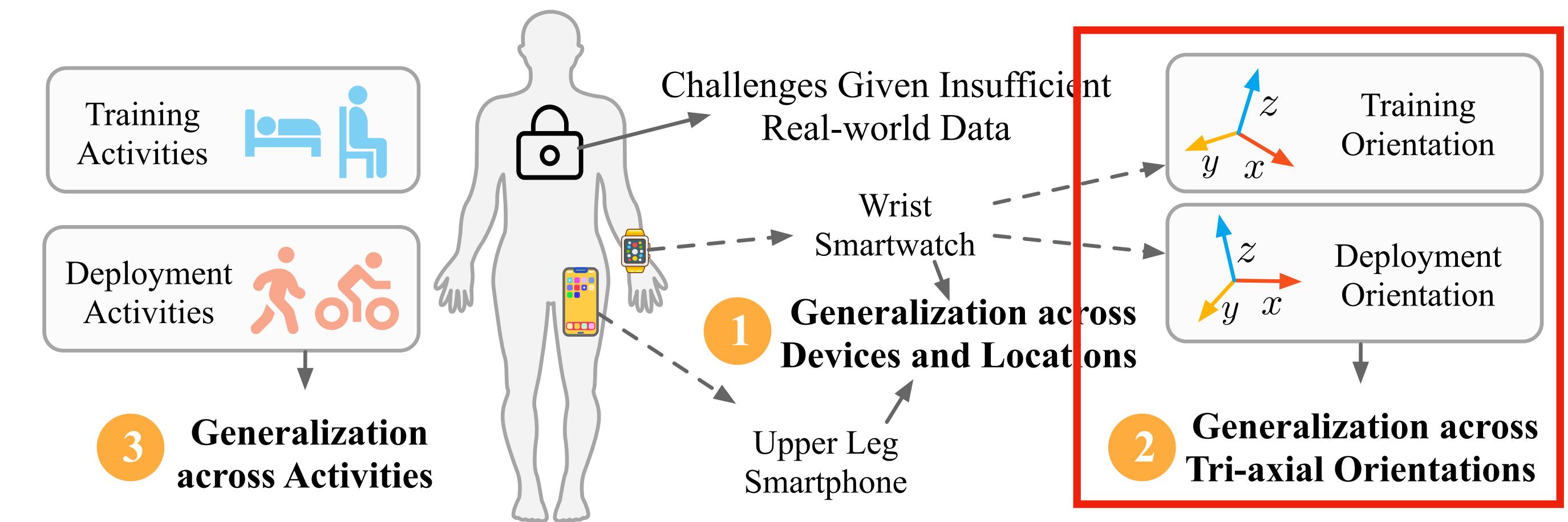
- Insufficient training samples
- Do not generalize across
 - Devices or Locations
 - Orientations
 - Activities



Introduction

Challenges

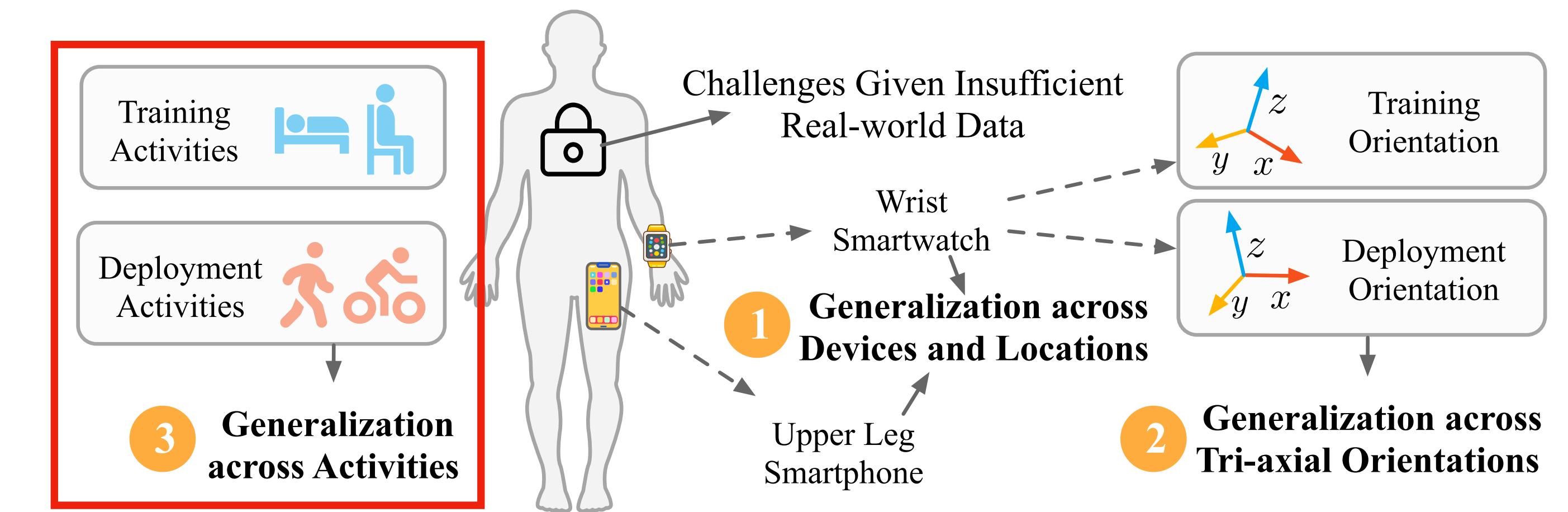
- Insufficient training samples
- Do not generalize across
 - Devices or Locations
 - Orientations
 - Activities



Introduction

Challenges

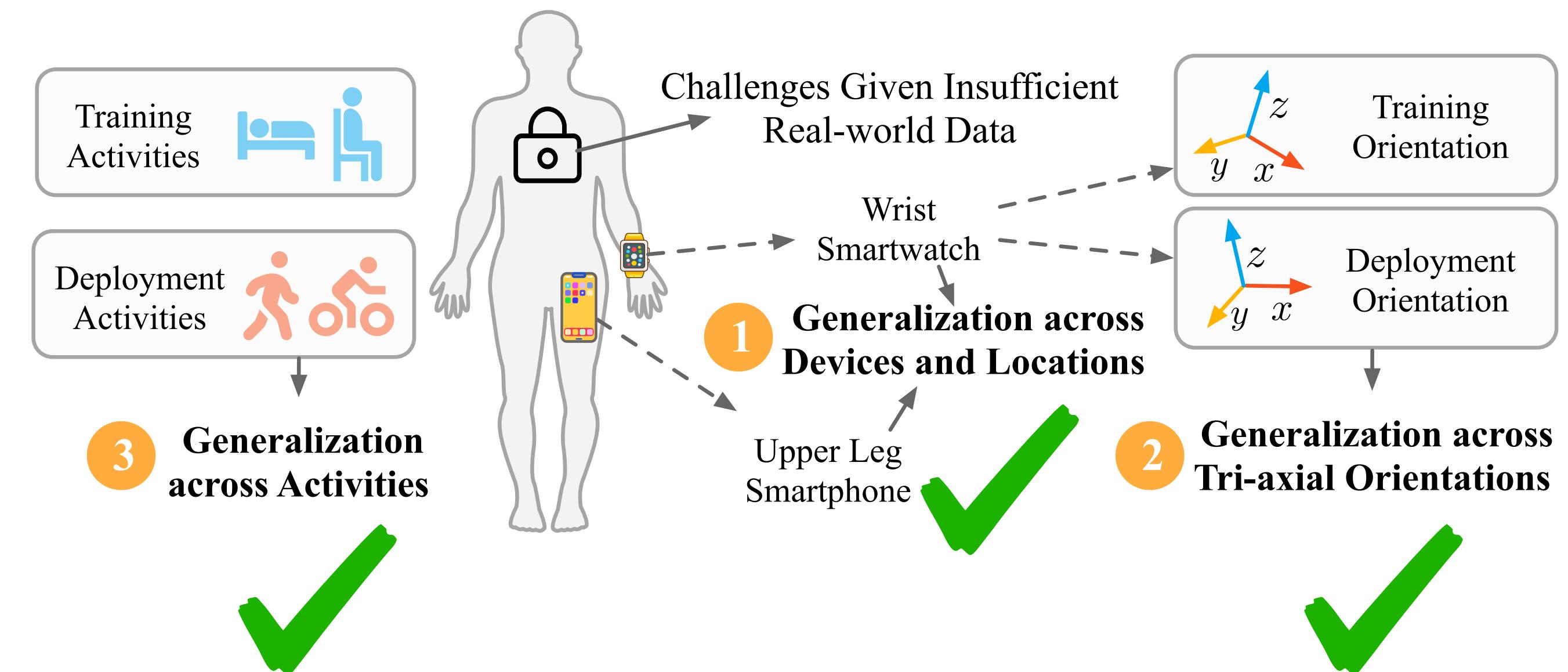
- Insufficient training samples
- Do not generalize across
 - Devices or Locations
 - Orientations
 - Activities



Introduction

Goal

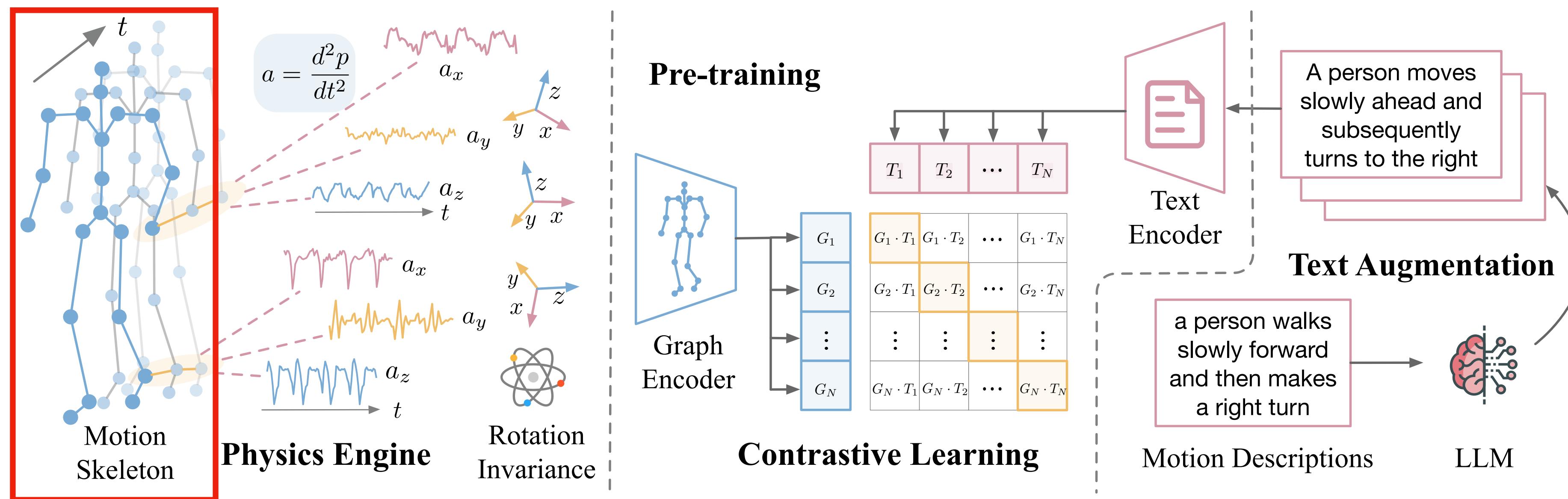
- A single framework that addresses all the above three challenges
- Generalize across
 - Devices or Locations
 - Orientations
 - Activities



Method

Pre-training

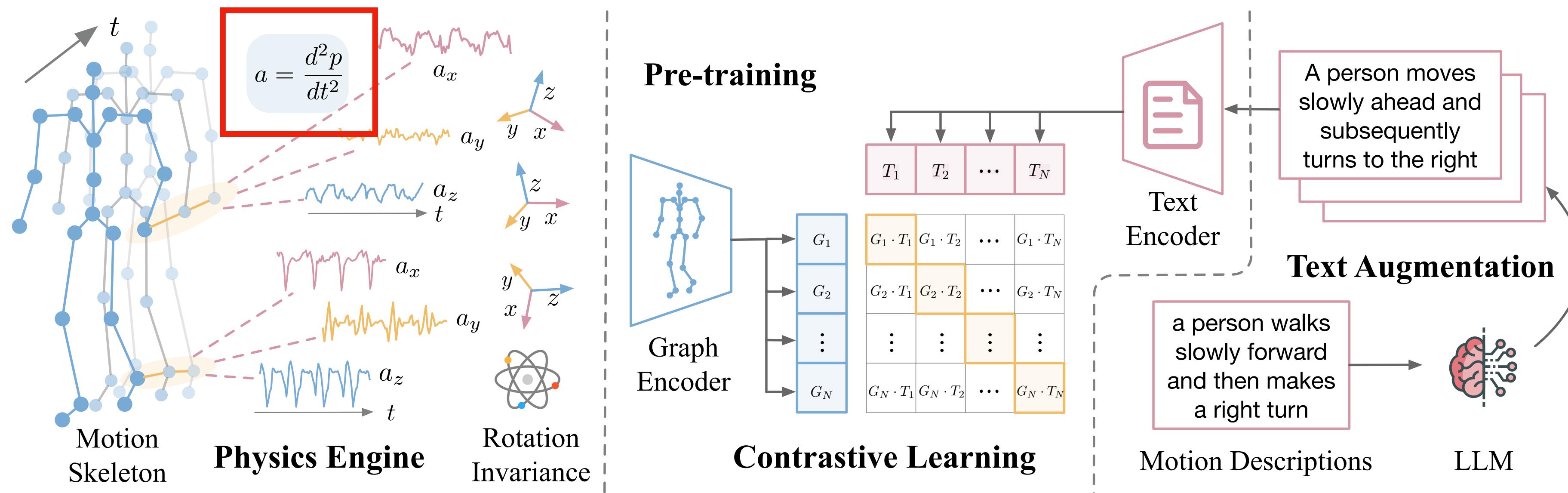
- Physics engine: simulate synthetic IMU data + rotation-invariant augmentation
- Contrastive learning: graph encoder + text encoder with text augmentation



Method

Pre-training

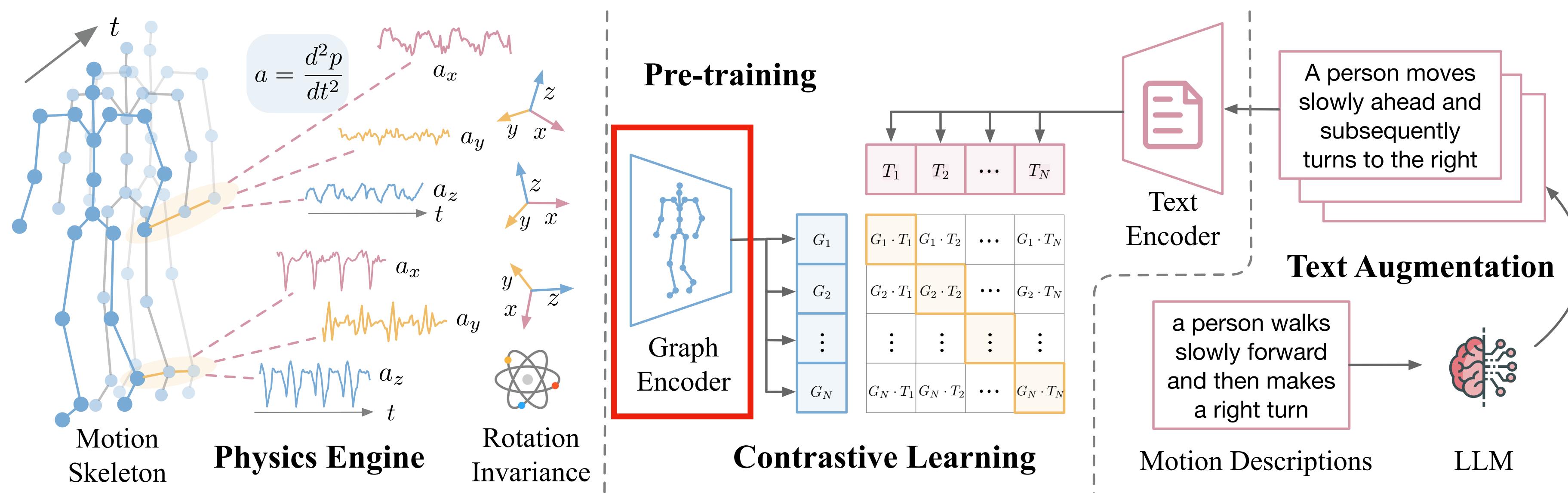
- Physics engine: simulate synthetic IMU data + rotation-invariant augmentation
- Contrastive learning: graph encoder + text encoder with text augmentation



Method

Pre-training

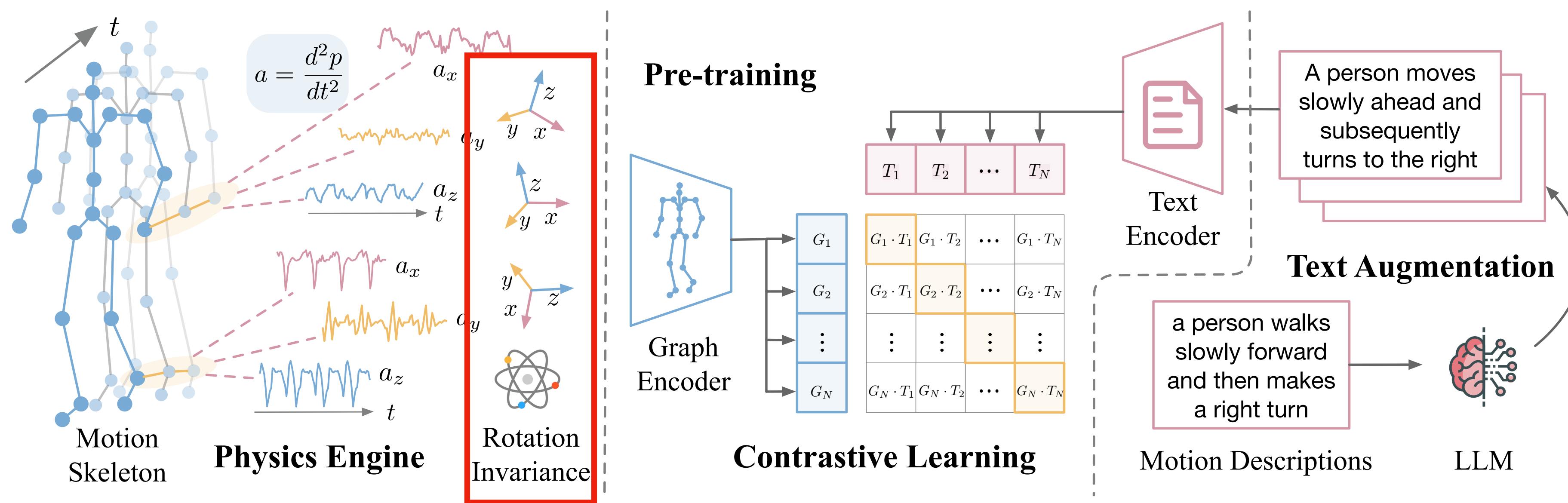
- Physics engine: simulate synthetic IMU data + rotation-invariant augmentation
- Contrastive learning: graph encoder + text encoder with text augmentation



Method

Pre-training

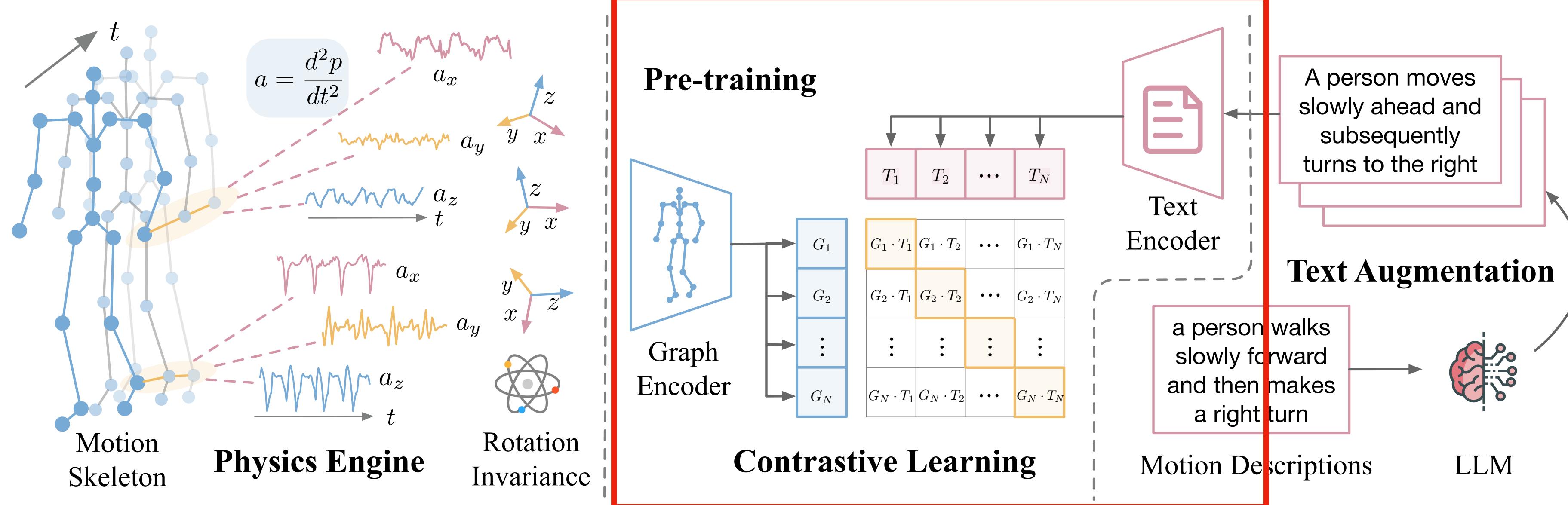
- Physics engine: simulate synthetic IMU data + rotation-invariant augmentation
- Contrastive learning: graph encoder + text encoder with text augmentation



Method

Pre-training

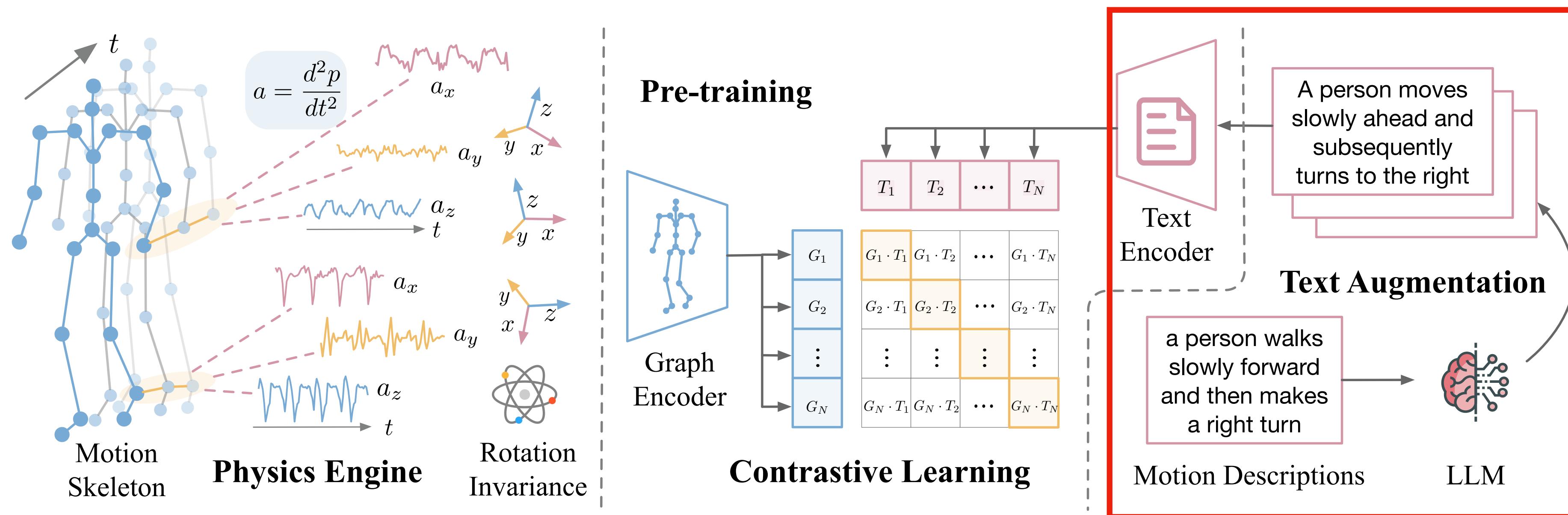
- Physics engine: simulate synthetic IMU data + rotation-invariant augmentation
- Contrastive learning: graph encoder + text encoder with text augmentation



Method

Pre-training

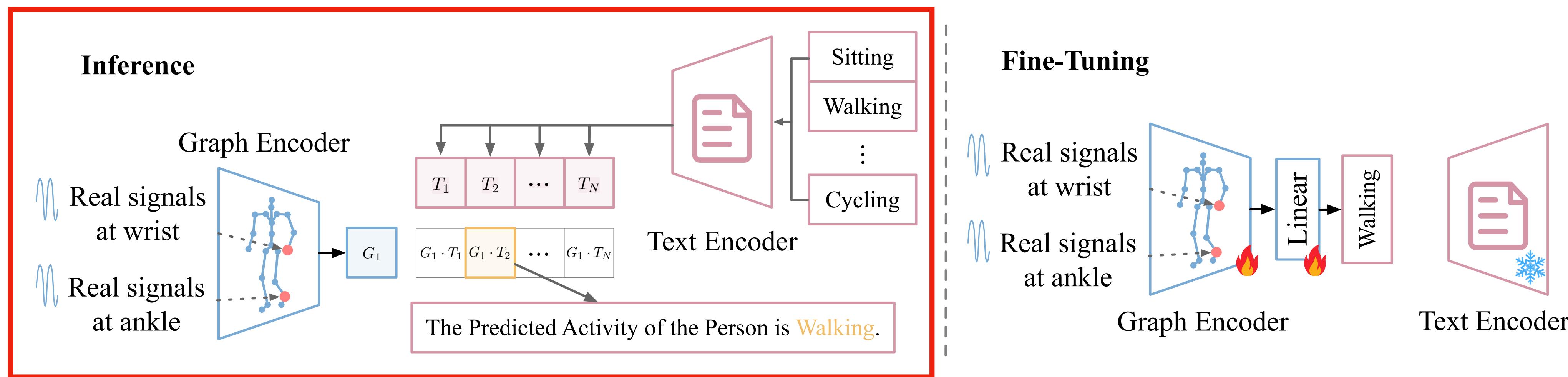
- Physics engine: simulate synthetic IMU data + rotation-invariant augmentation
- Contrastive learning: graph encoder + text encoder with text augmentation



Method

Fine-tuning and Inference

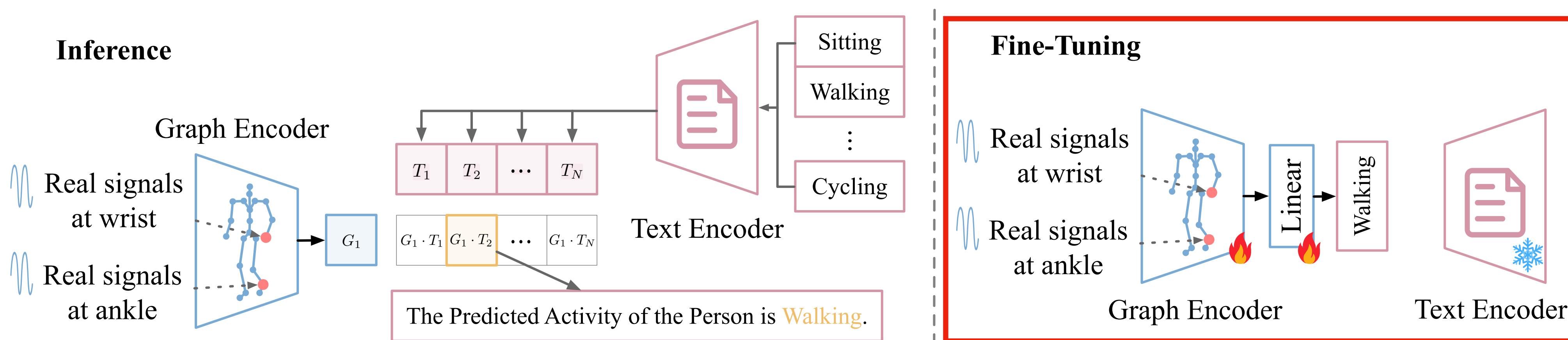
- Inference: compare the similarity scores with all label candidates
- Fine-tuning: cross entropy



Method

Fine-tuning and Inference

- Inference: compare the similarity scores with all label candidates
- Fine-tuning: cross entropy



Experiments

Datasets

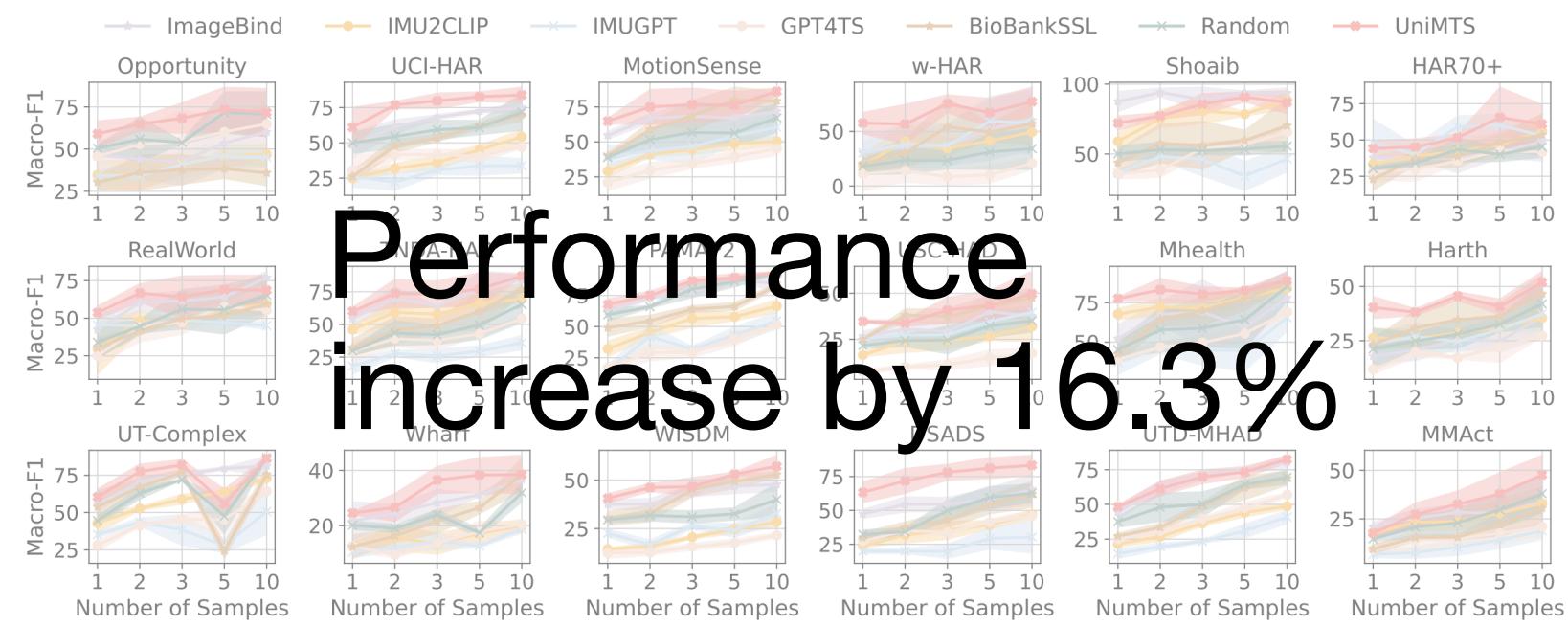
- State-of-the-art performance on 18 HAR benchmark datasets
 - 8 easy datasets (<10), 8 medium datasets (10-20), 2 hard datasets (> 20)

Experiments

Main Results

- State-of-the-art performance on 18 HAR benchmark datasets
 - 8 easy datasets (<10), 8 medium datasets (10-20), 2 hard datasets (> 20)

Dataset	Metrics	Performance increase by 340%																		
		Opportunity	UCI-HAR	MotionSense	w-HAR	Shoaib	HAR70+	RealWorld	TNDA-HAR	PAMAP2	USC-HAD	Mhealth	Harth	UT-Complex	Wharf	WISDM	DSADS	UTD-MHAD	MMAct	Average
Number of Classes	4	6	6	7	7	7	8	8	12	12	12	12	13	14	18	19	27	35		
Level	Easy										Medium									
ImageBind [16]	Acc	50.2	14.1	18.1	13.1	19.4	0.0	18.9	19.3	14.6	11.8	7.9	9.8	9.7	3.2	7.1	2.1	3.3	2.9	12.5
	F1	30.0	5.0	16.4	8.6	15.5	0.0	10.1	13.7	8.1	7.3	1.7	6.1	6.3	1.9	4.5	1.2	1.6	1.5	7.8
	R@2	83.6	21.4	42.8	54.1	35.5	0.2	23.9	32.8	15.6	25.0	21.3	22.5	17.8	6.8	13.7	6.9	5.1	3.7	24.0
IMU2CLIP [38]	Acc	25.9	17.8	15.5	6.6	16.7	5.6	6.1	8.5	1.9	12.8	7.9	2.5	7.3	1.4	4.3	4.6	3.7	6.4	8.6
	F1	10.3	11.4	8.6	3.3	9.2	1.8	4.5	4.2	1.1	9.3	1.3	1.1	3.4	0.7	2.7	1.1	2.1	1.6	4.3
	R@2	65.5	35.1	39.1	19.7	34.6	22.2	19.7	22.4	9.8	19.8	13.4	4.8	16.3	5.5	9.6	8.3	6.5	10.1	20.1
IMUGPT [28]	Acc	10.1	11.1	1.8	67.2	12.1	0.0	16.0	14.2	8.9	6.0	0.8	4.8	11.6	2.7	8.3	7.5	3.7	2.0	11.1
	F1	0.4	0.0	3.5	58.8	6.2	0.0	0.0	6.0	1.5	6.9	0.5	1.1	8.3	1.8	6.6	2.0	0.3	0.7	5.7
	R@2	3.7	8.2	40.4	67.2	32.1	0.0	3.7	25.5	1.3	31.1	0.5	1.4	17.7	13.9	14.6	8.8	5.1	23.9	
HARGPT [24]	Acc	28.8	15.0	11.6	4.9	21.0	34.3	12.7	13.7	11.1	9.5	10.4	28.8	7.5	5.9	5.5	5.8	3.3	2.3	12.9
	F1	17.3	12.7	5.6	3.1	12.4	10.6	5.3	5.4	2.1	3.6	7.4	7.5	4.4	1.4	3.5	3.4	1.5	1.1	6.0
	R@2	47.0	31.4	35.9	11.5	38.6	51.9	31.7	25.2	23.1	17.6	27.4	48.6	6.3	11.4	1.8	12.1	3.3	5.1	25.3
LLaVA [30]	Acc	40.1	16.3	22.0	0.0	16.0	1.3	16.4	13.9	10.1	11.1	8.5	16.3	2.7	3.1	5.6	5.3	5.1	4.0	12.0
	F1	14.8	5.5	6.2	0.0	4.8	2.7	8.3	2.9	1.2	2.1	5.0	0.6	0.7	0.6	0.1	0.3	0.0	3.4	0.0
	R@2	67.6	54.6	34.4	0.0	35.3	43.4	28.4	25.2	18.7	19.6	32.5	16.4	9.0	3.6	10.6	10.5	7.4	7.5	22.4
UniMTS	Acc	45.9	35.2	45.2	59.0	63.6	68.2	43.6	59.1	47.2	30.5	70.7	68.9	34.8	18.2	27.8	31.5	22.8	10.2	43.5
	F1	42.2	22.0	33.7	42.9	57.2	34.8	36.7	53.7	43.6	27.8	61.8	41.1	29.2	13.7	25.5	23.7	18.5	10.0	34.3
	R@2	80.0	53.1	57.2	60.7	82.1	86.6	64.0	77.5	63.2	45.4	78.7	85.0	44.2	38.6	47.1	46.0	32.6	18.7	58.9
w/o rot aug	Acc	37.7	18.6	25.1	36.1	19.4	53.4	55.5	35.6	32.5	20.7	27.4	59.4	9.9	3.6	13.5	21.5	13.5	4.3	27.1
	F1	30.4	8.2	11.2	26.4	13.5	27.7	41.1	29.6	28.3	10.5	24.0	20.5	4.9	4.8	10.6	13.4	7.2	4.7	17.6
	R@2	74.3	40.1	64.3	52.5	40.1	76.5	76.3	54.4	48.8	29.7	49.4	61.2	28.5	6.8	23.3	34.0	32.1	7.6	44.0
w/o text aug	Acc	52.7	36.3	43.4	57.4	55.6	61.0	40.7	40.9	38.4	29.6	59.2	62.8	30.3	28.2	31.3	29.3	6.5	12.6	39.8
	F1	39.4	21.3	34.7	41.4	49.5	32.8	29.7	32.7	33.9	21.6	49.0	26.7	21.2	19.6	27.2	22.2	4.7	10.3	28.8
	R@2	70.9	39.6	63.7	57.4	78.7	77.4	60.4	63.5	48.8	49.1	63.4	77.9	41.4	43.2	45.1	41.7	10.7	23.2	53.1
w/o graph	Acc	41.4	37.6	18.9	23.0	50.9	18.8	42.7	33.8	30.0	22.4	28.7	34.8	23.4	0.5	12.4	19.8	1.9	11.2	25.1
	F1	20.5	28.5	21.6	17.9	38.6	9.1	23.7	28.9	23.6	23.0	19.3	15.3	11.5	1.1	8.5	16.0	2.7	7.7	17.6
	R@2	63.7	66.7	54.2	34.4	66.1	35.3	51.1	62.1	41.8	41.7	43.9	51.6	35.6	8.2	23.7	29.1	4.7	18.0	40.7



Few-shot Setting

Dataset	Metrics	Opportunity	UCI-HAR	MotionSense	w-HAR	Shoaib	HAR70+	RealWorld	TINDA-HAR	PAMAP2	USC-HAD	Mhealth	Harth	UT-Complex	Wharf	WISDM	DSADS	UTD-MHAD	MMAct	Average
Number of Classes		4	6	6	7	7	7	8	8	12	12	12	12	13	14	18	19	27	35	
Level		Easy								Medium								Hard	Avg	
ImageBind [16]	Acc	82.8	90.9	94.0	95.1	98.5	91.8	83.8	95.0	94.3	52.1	82.9	93.0	97.0	71.4	77.2	90.0	74.0	65.8	85.0(12.2)
	F1	84.6	90.7	93.1	61.9	98.5	67.7	86.1	94.9	94.6	52.2	83.2	62.7	96.9	43.3	76.8	89.3	71.6	58.3	78.1(16.5)
IMU2CLIP [38]	Acc	71.9	87.1	79.3	88.5	98.2	87.8	79.9	97.1	90.7	55.7	94.5	93.9	90.6	52.3	62.7	88.7	62.8	66.9	80.5(14.3)
	F1	70.0	87.0	77.2	75.1	98.1	64.4	80.5	97.1	91.2	52.0	95.0	65.7	90.2	35.6	62.8	88.2	59.6	60.3	75.0(17.0)
TST [68]	Acc	89.1	91.6	91.4	75.4	73.5	95.8	68.9	94.1	88.9	66.2	86.0	98.8	92.9	59.1	70.2	80.0	61.4	54.5	79.9(13.6)
	F1	91.1	91.4	89.6	79.3	72.3	62.9	67.2	94.1	89.8	60.9	83.8	72.7	92.8	34.5	69.1	76.1	59.6	48.4	74.2(16.3)
TARNet [11]	Acc	82.6	91.1	75.0	79.7	67.0	67.2	60.9	92.5	92.4	57.6	82.0	94.7	94.4	63.2	58.5	30.7	78.1	56.7	75.1(17.5)
	F1	74.4	91.2	70.0	41.0	61.1	56.9	54.1	92.4	70.0	48.7	77.5	47.8	94.4	38.1	58.5	23.0	75.5	55.9	65.9(20.5)
TS2Vec [67]	Acc	87.7	92.1	94.2	90.0	84.4	84.4	74.6	90.1	91.1	50.1	84.4	98.0	85.5	73.2	69.5	84.3	80.0	53.4	84.5(13.9)
	F1	83.3	92.1	93.9	100.0	87.2	88.1	68.8	95.9	92.9	53.6	93.8	58.3	98.5	47.9	68.8	82.3	76.7	53.4	79.8(16.7)
BioBankSSL [60]	Acc	85.9	92.7	99.4	93.4	99.1	93.2	75.5	89.4	87.3	65.8	95.1	98.0	96.4	81.4	82.3	65.8	80.5	62.5	85.8(11.5)
	F1	86.1	92.9	99.2	81.3	99.1	85.0	68.0	89.6	87.3	72.6	94.0	68.1	96.1	62.7	82.3	81.1	78.4	58.4	82.7(12.1)
THAT [29]	Acc	80.1	86.8	87.1	77.1	92.0	95.8	67.7	97.7	95.5	51.9	89.0	97.6	86.1	60.0	60.3	82.1	71.1	55.0	80.1(14.7)
	F1	84.5	86.7	86.5	82.0	92.4	95.8	62.9	97.7	95.5	53.0	89.9	71.9	90.0	23.5	60.0	79.4	70.8	52.3	74.5(17.5)
IMUGPT [28]	Acc	84.8	87.0	86.2	85.3	61.7	98.3	62.2	87.7	85.6	41.1	71.3	98.3	86.5	54.6	67.8	71.9	57.7	51.9	74.4(16.3)
	F1	84.9	87.0	85.2	48.8	61.8	78.0	56.0	87.5	83.8	42.4	63.6	65.2	85.8	24.6	67.4	71.0	53.2	49.0	66.4(17.7)
TimesNet [60]	Acc	80.0	88.5	90.5	88.5	82.4	97.5	65.4	93.2	88.0	58.3	81.7	97.7	92.7	41.4	67.4	77.3	62.3	50.9	78.0(16.2)
	F1	82.4	88.6	88.3	79.5	82.6	77.2	62.2	93.1	88.0	48.7	79.1	61.4	92.8	30.2	67.1	78.5	61.0	45.7	72.6(17.3)
GPT4TS [72]	Acc	83.6	88.4	92.3	70.5	73.5	97.8	66.5	95.7	92.6	61.7	87.8	97.2	95.3	51.8	63.8	79.9	66.1	52.6	78.7(15.2)
	F1	86.1	88.5	92.1	54.8	74.1	76.6	56.6	95.7	93.4	58.9	85.2	58.2	95.2	34.4	64.5	78.0	63.7	48.8	72.5(17.7)
SHARE [71]	Acc	89.0	92.2	99.6	96.7	77.5	97.7	66.0	95.9	94.1	72.9	97.0	99.1	98.7	63.2	77.0	92.0	73.0	62.1	85.8(13.2)
	F1	90.2	92.1	99.6	85.5	78.0	77.5	57.1	95.8	94.8	66.7	97.1	74.0	98.8	40.6	76.5	91.9	69.3	56.5	80.1(16.5)
UniMTS	Acc	88.2	92.1	99.8	98.4	99.7	96.9	84.0	99.6	98.0	58.3	97.0	98.0	99.1	82.7	81.3	94.3	85.6	77.1	90.6(10.6)
	F1	89.1	92.2	99.8	98.8	99.7	87.9	81.2	99.6	98.1	63.1	97.3	86.7	99.2	51.7	80.9	94.2	83.2	71.7	87.5(13.4)
Random	Acc	87.7	89.3	95.5	95.1	66.1	98.1	74.2	98.7	96.2	48.5	82.9	98.6	98.5	76.8	79.4	90.3	74.4	73.9	84.7(13.5)
	F1	88.9	89.4	95.5	60.1	65.0	85.9	74.0	98.7	96.8	51.9	78.2	69.1	98.5	50.8	79.3	90.1	69.7	71.6	78.5(15.1)

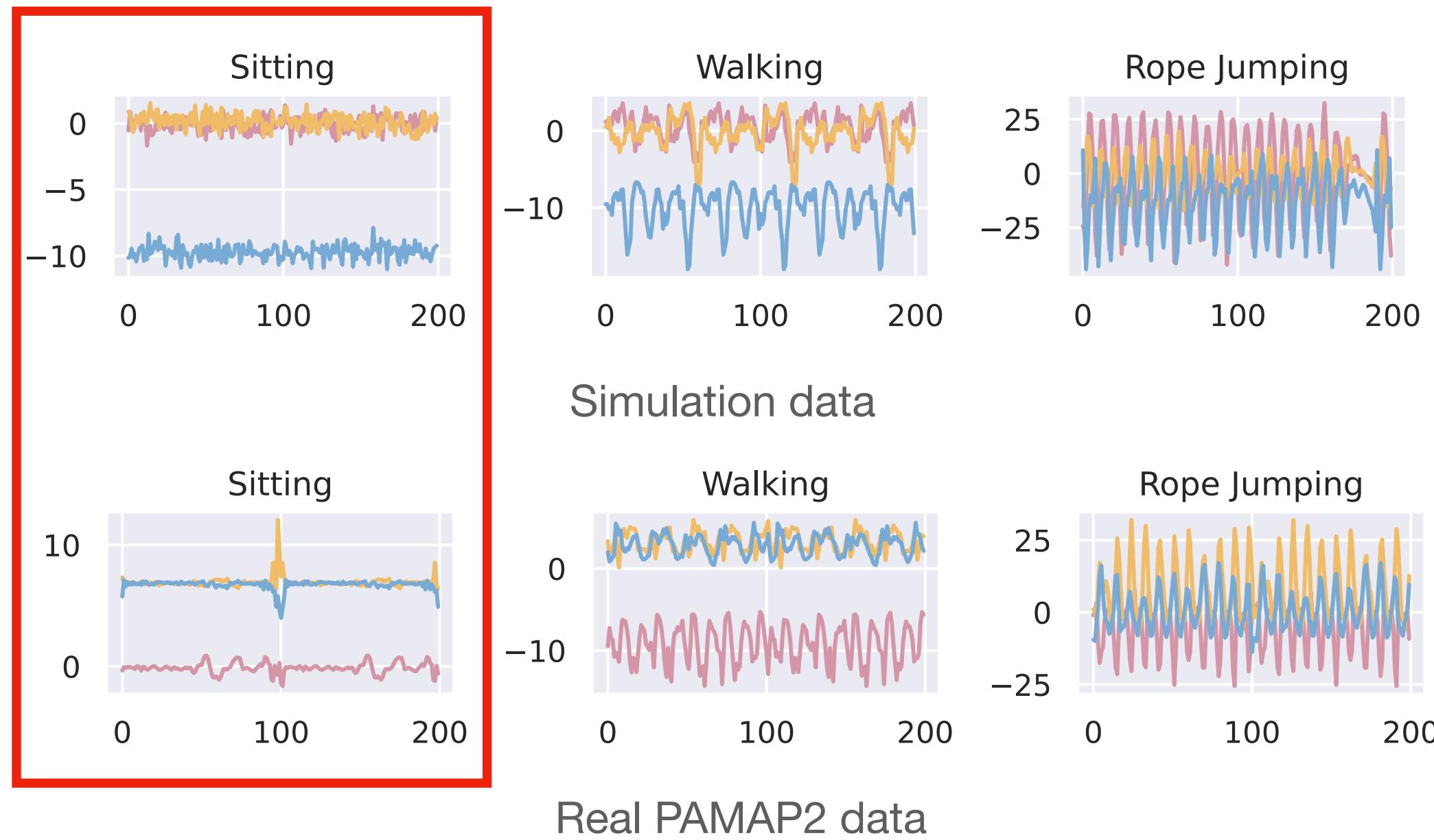
Full-shot setting

Zero-shot Setting

Experiments

Visualizations

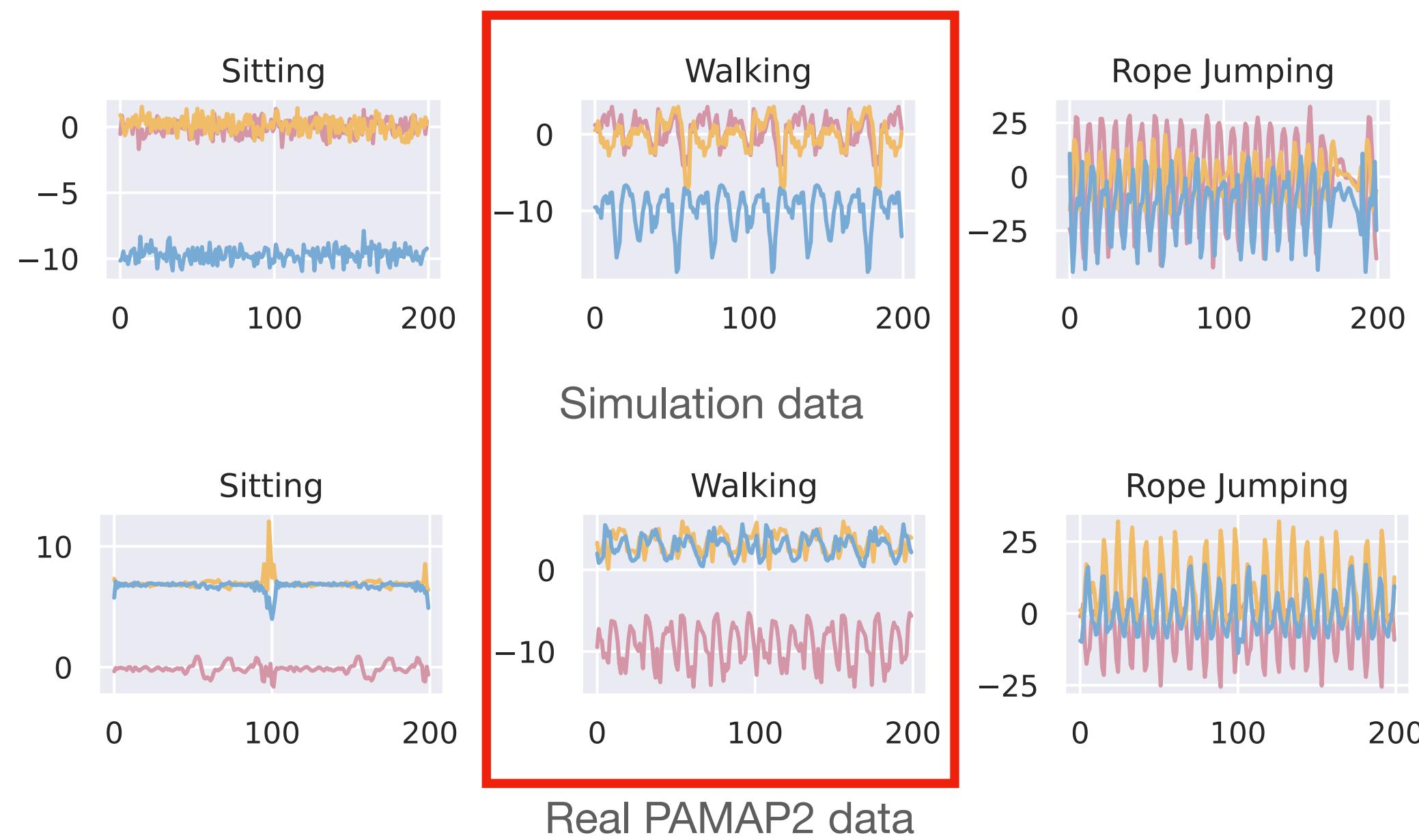
- Simulated data have similar patterns as real data



Experiments

Visualizations

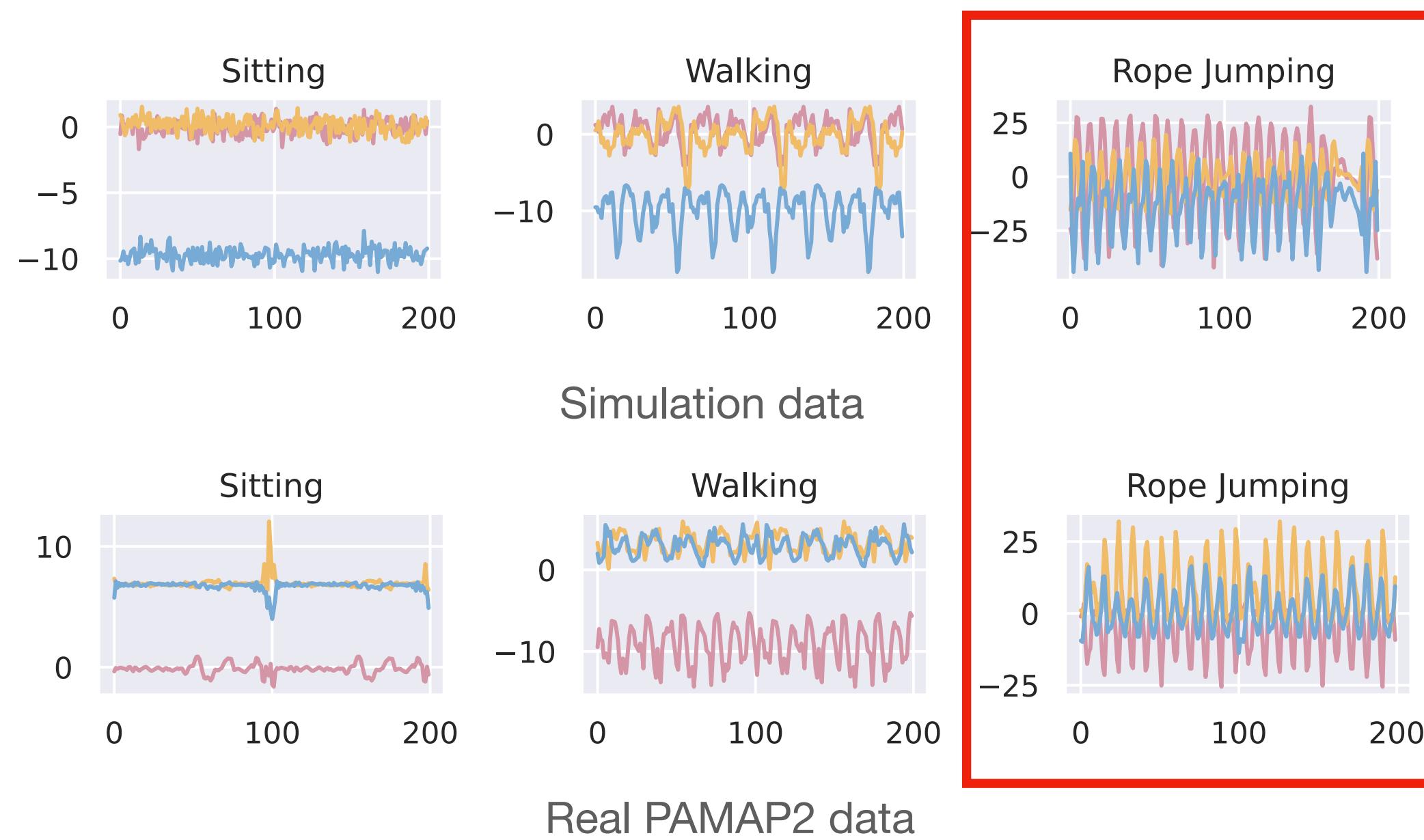
- Simulated data have similar patterns as real data



Experiments

Visualizations

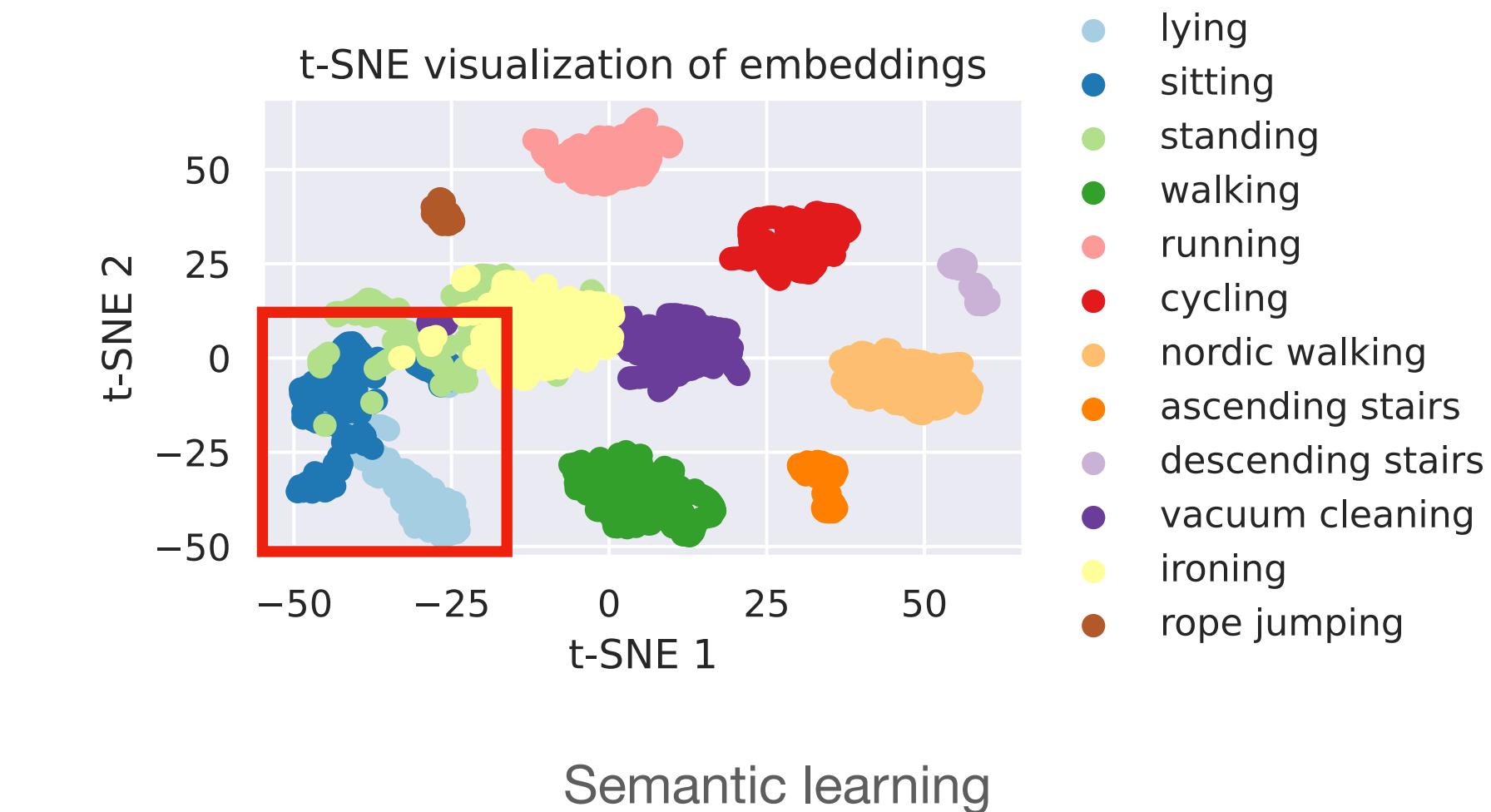
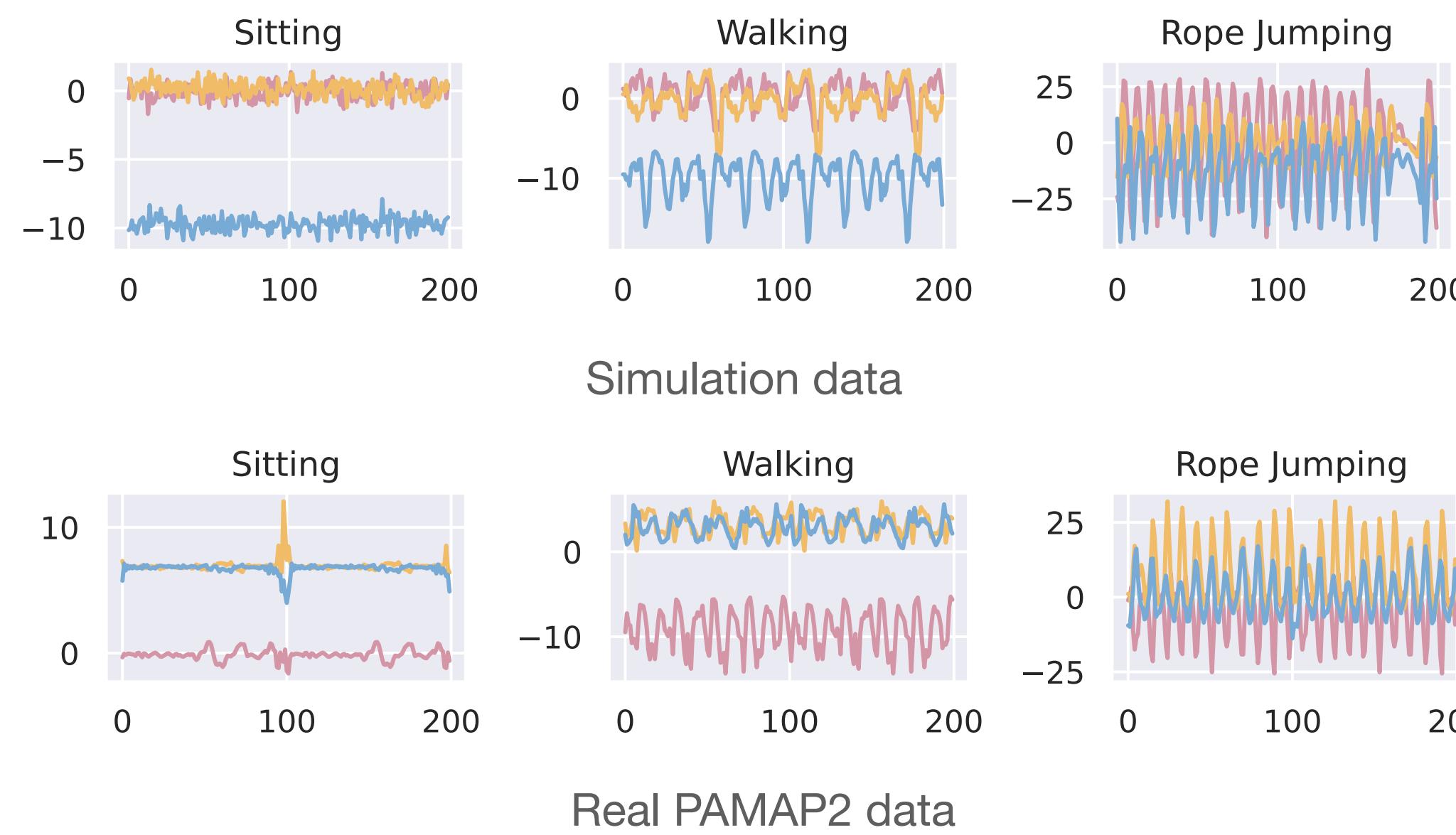
- Simulated data have similar patterns as real data



Experiments

Visualizations

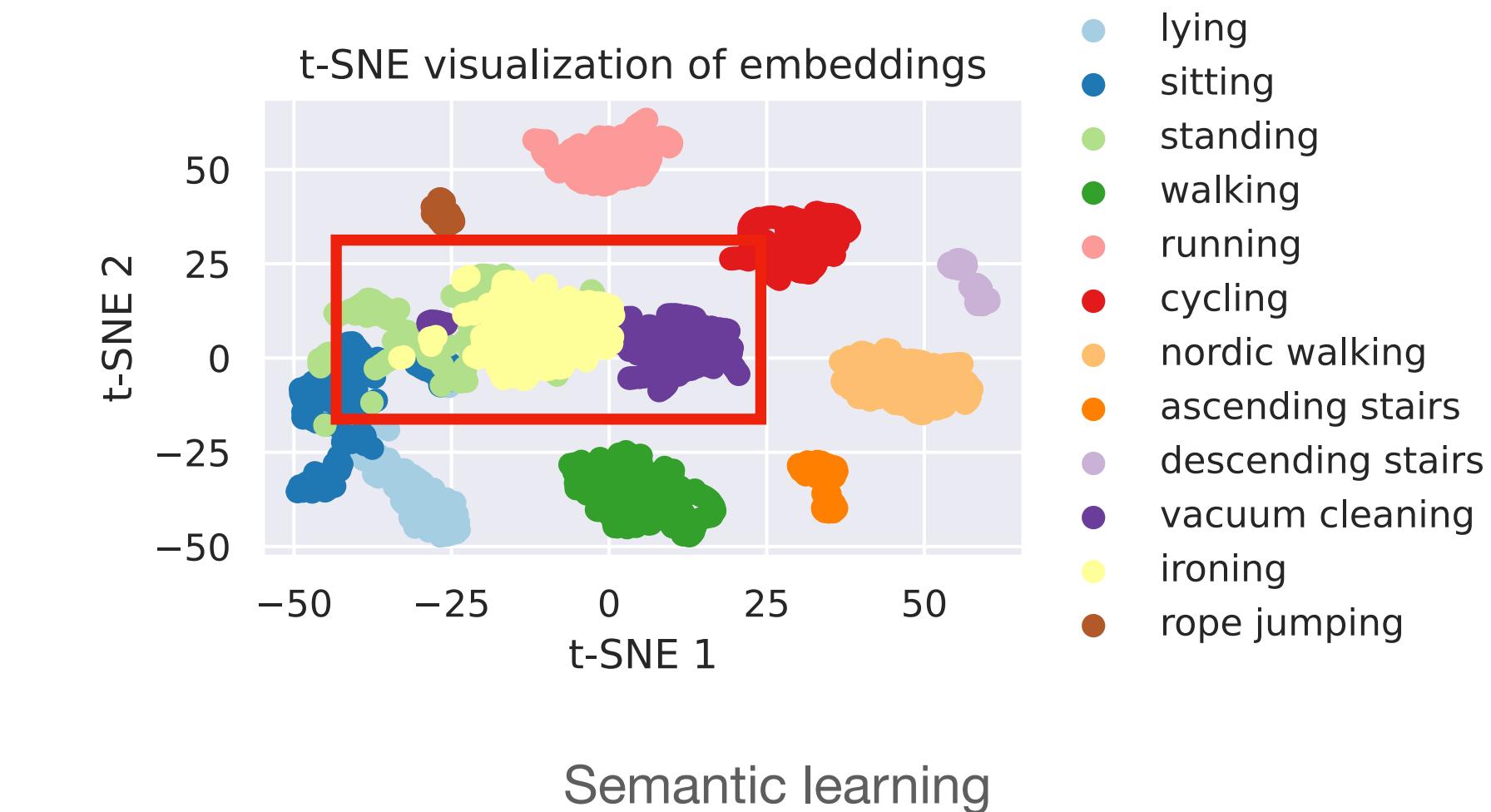
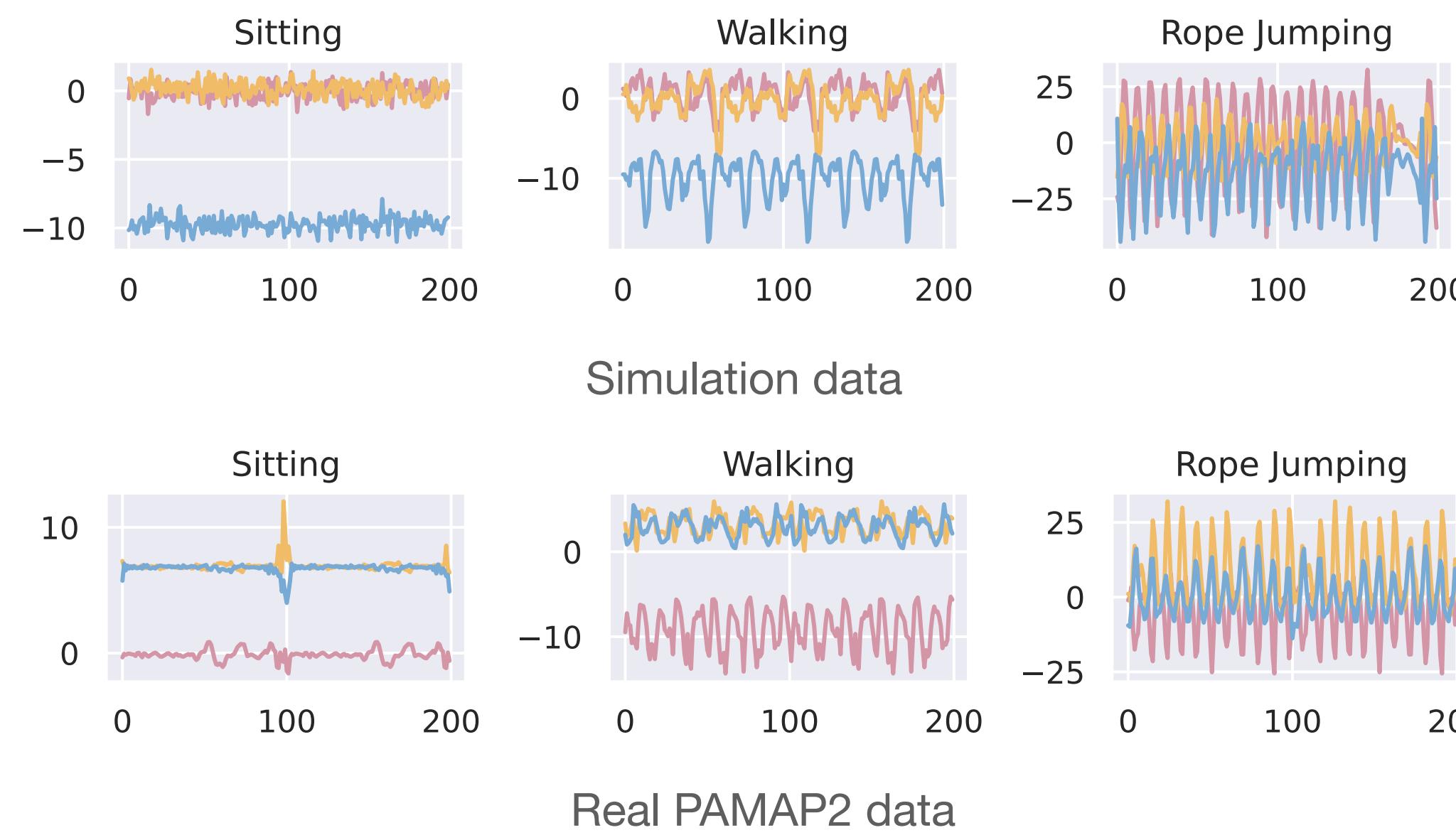
- Simulated data have similar patterns as real data
- Contrastive pre-training learns the semantics of time series data



Experiments

Visualizations

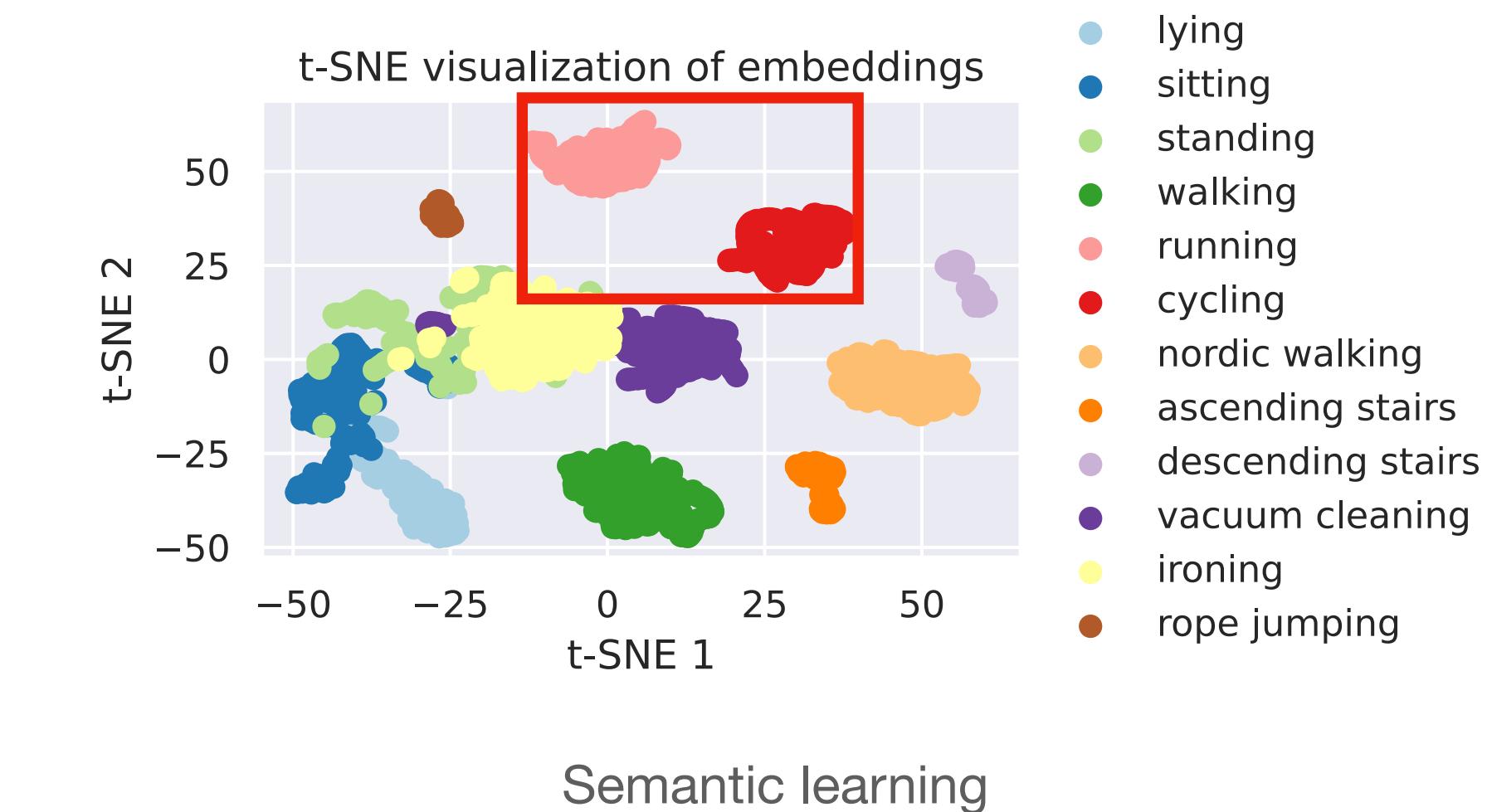
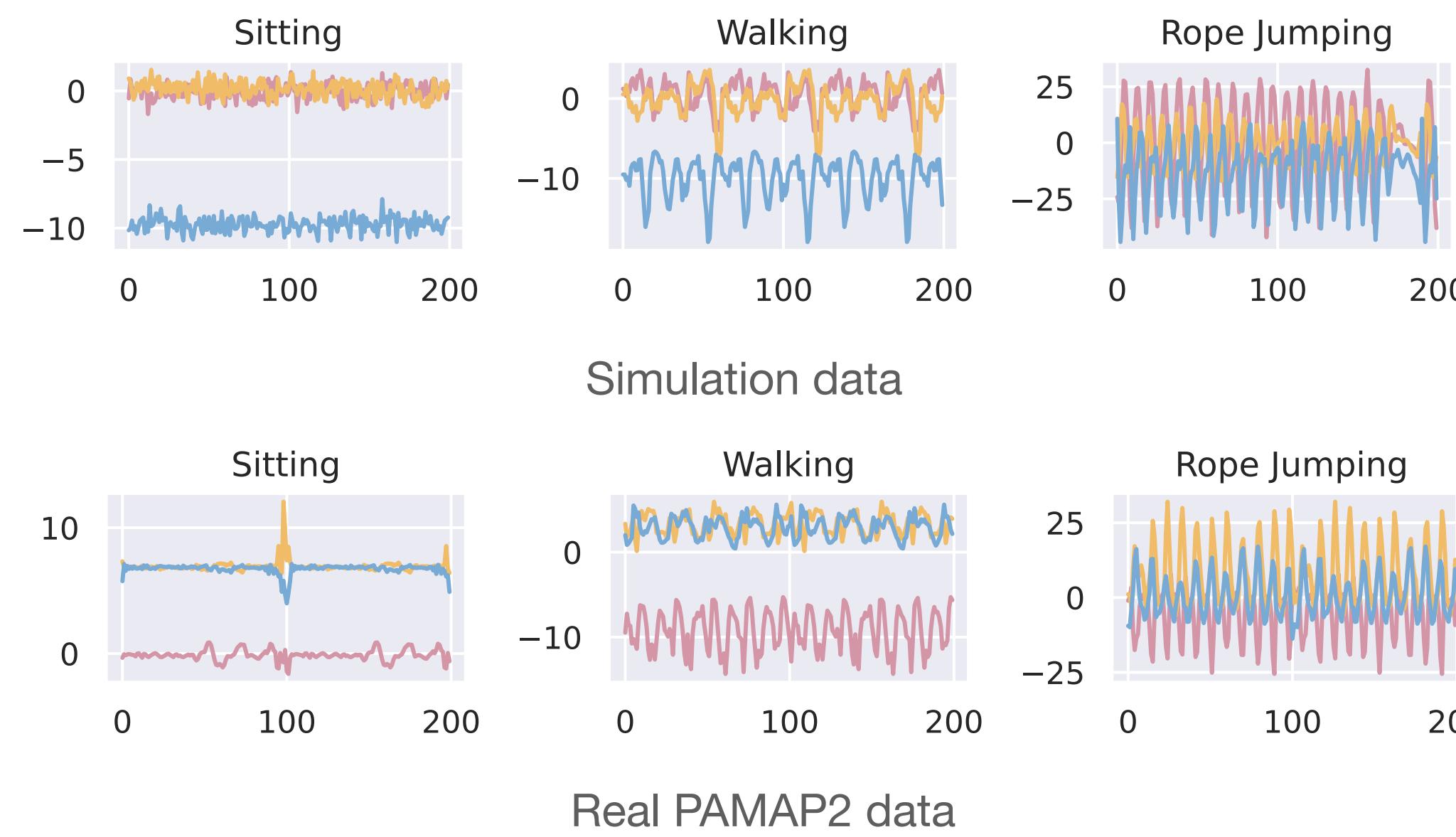
- Simulated data have similar patterns as real data
- Contrastive pre-training learns the semantics of time series data



Experiments

Visualizations

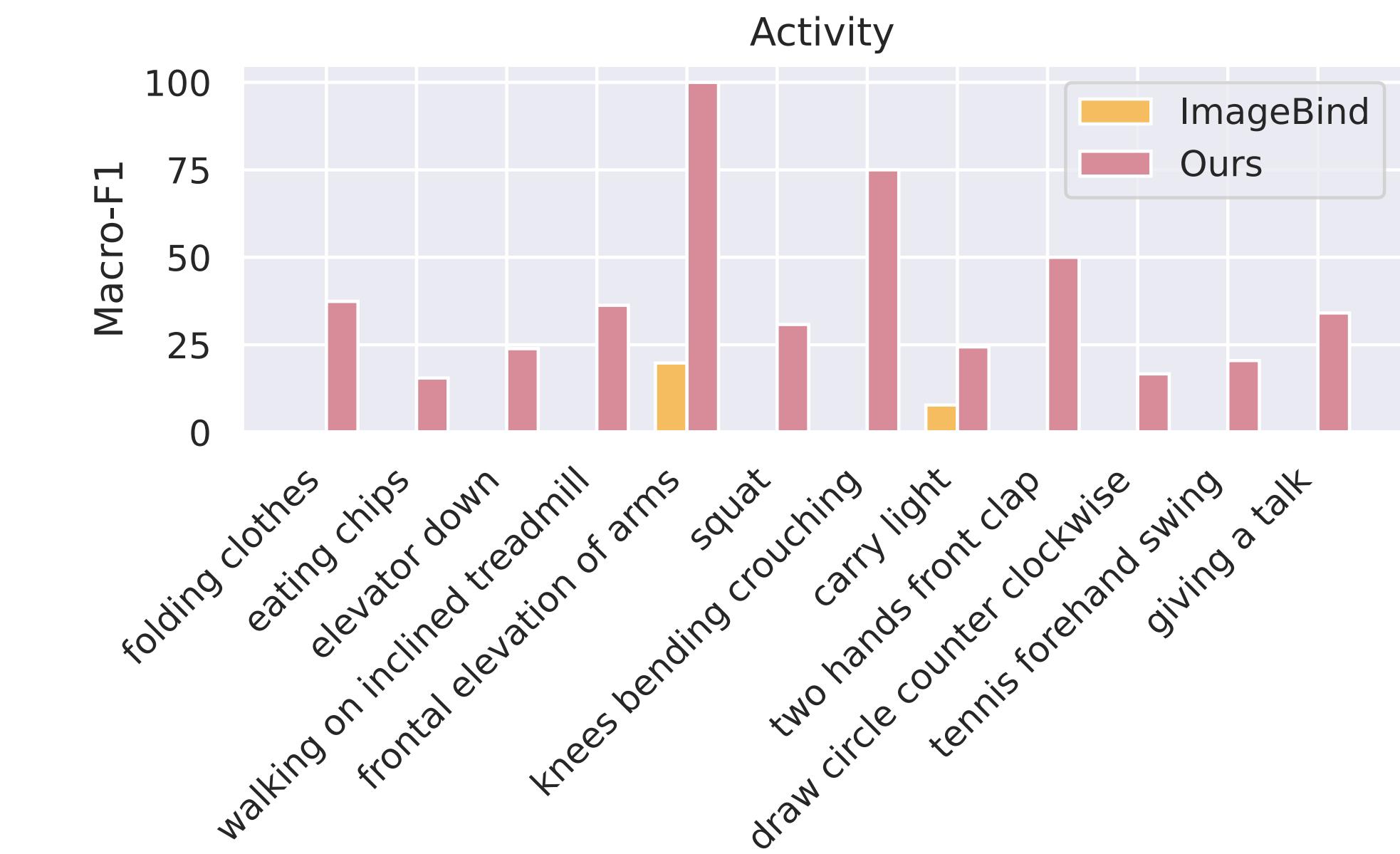
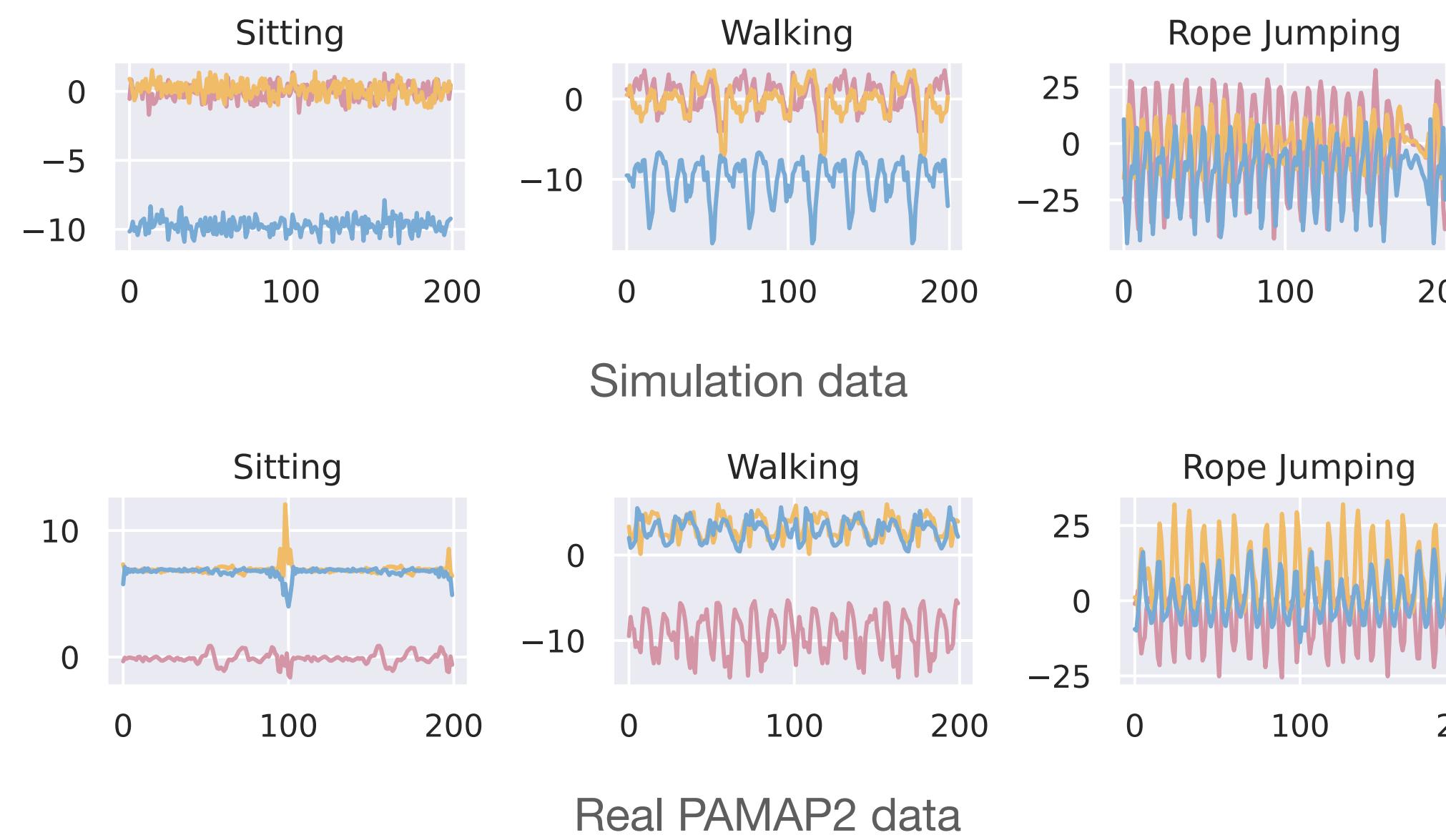
- Simulated data have similar patterns as real data
- Contrastive pre-training learns the semantics of time series data



Experiments

Generalization to New Activities

- Simulated data have similar patterns as real data
- Contrastive pre-training learns the semantics of time series data



Zero-shot generalization to new activities

Thank you!

Contact: xiyuanzh@ucsd.edu

Code Release: <https://github.com/xiyuanzh/UniMTS>

Model Release: <https://huggingface.co/xiyuanz/UniMTS>