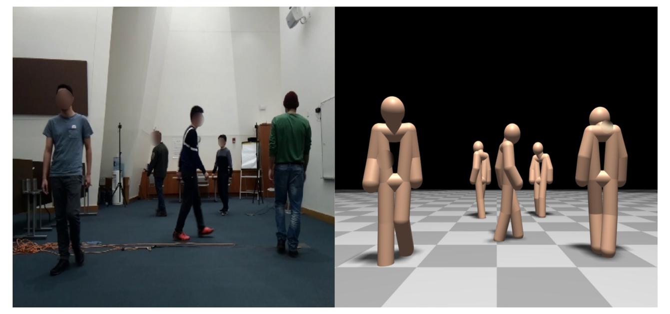
Unsupervised Learning for Human Sensing Using Radio Signals

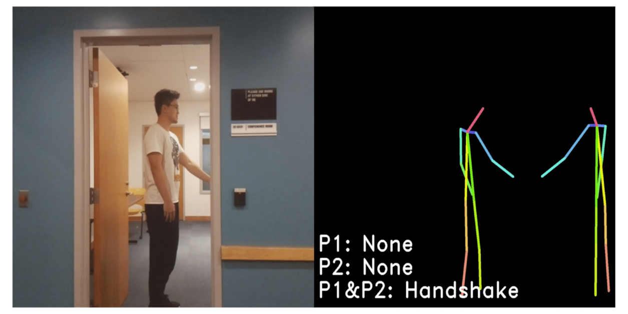
Tianhong Li*, Lijie Fan*, Yuan Yuan*, Dina Katabi
MIT CSAIL

Unsupervised Learning for RF-based Vision

RF-based vision: RF signals traverse walls and occlusions; thus, they can sense humans through walls and occlusions.

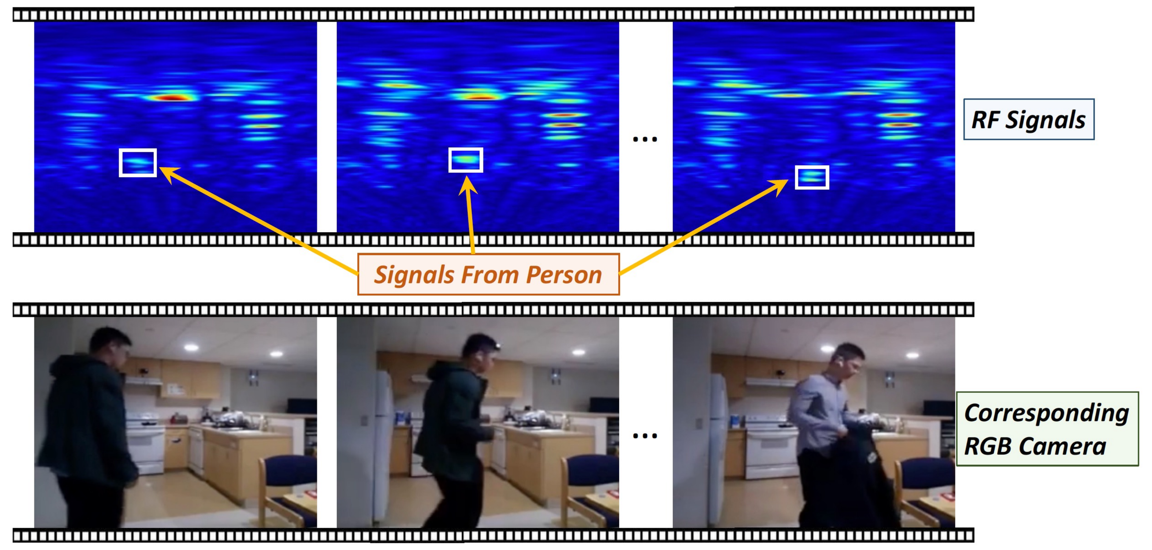
RF 3D Pose Estimation / **RF Action Recognition**





Motivation:
- Labeling RF signals is a daunting task because RF signals are not human interpretable.
- Leveraging large-scale unlabeled radio signals may improve the performance.

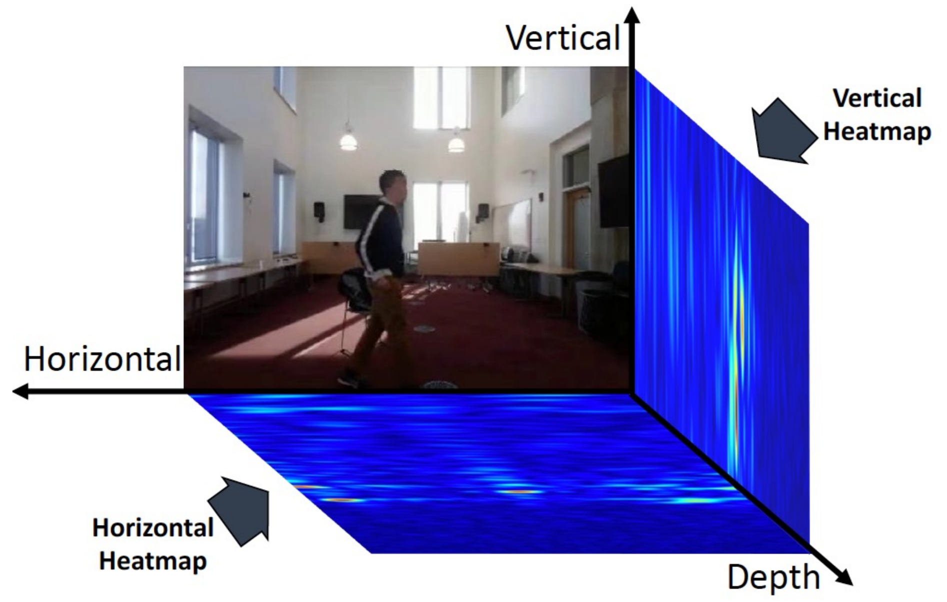
Challenge I: Human-Relevant Information Sparsity



Compared with RGB data:
- The information region in RF signals that corresponds to a person could be extremely small (<1%).
- RF signals carry much information that is irrelevant to the person or task, e.g., some signals that reflect off walls, signals that reflect other objects in the environment.

Solution:
- In most indoor scenarios, people are the only large moving objects. Therefore, we can adapt radar detection algorithms to detect and localize the person.
- Zoom in on radio signals which contain the person by cropping horizontal and vertical heatmaps based on their trajectory.

Challenge II: Augmentation is not applicable to RF signals



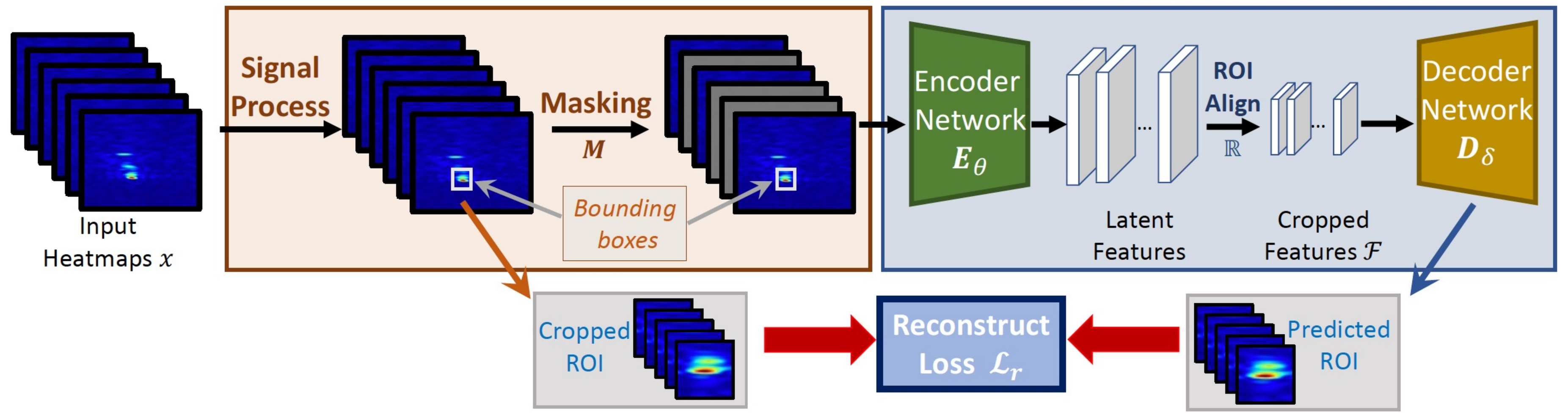
Self-supervised learning typically depends on data augmentation and pretext tasks.

RGB specific augmentations and tasks cannot be directly applied to RF signals, e.g.,
- No color information in RF signals
- RF signal is not invariant to rotation transformation

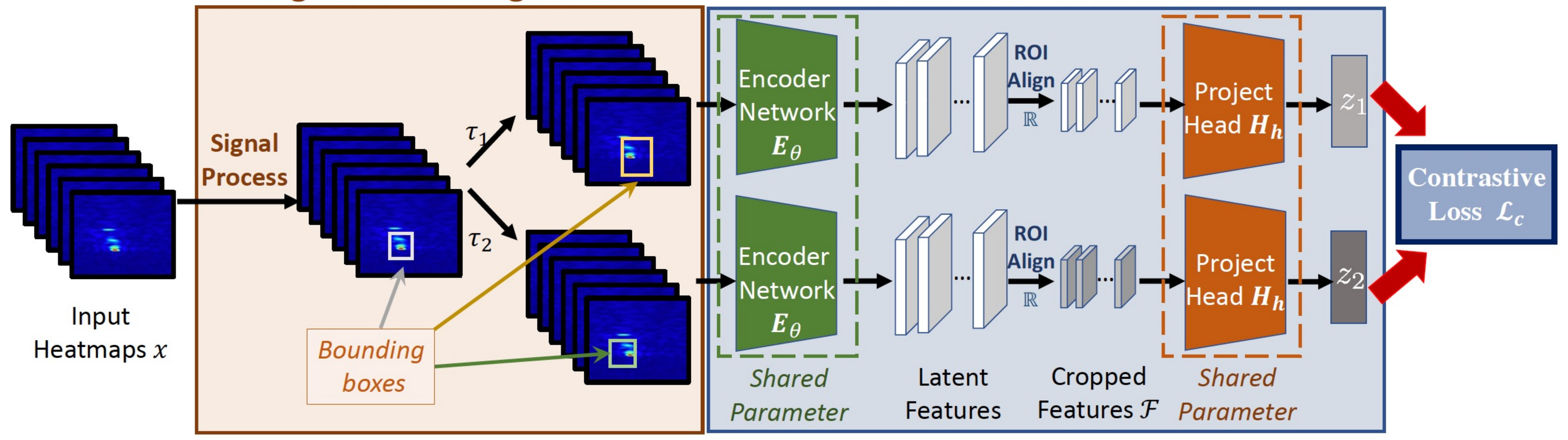
Solution: Predictive unsupervised learning is more suitable than contrastive unsupervised learning. Use adaptive reconstruction loss for RF data unsupervised learning.

Network Structure: Trajectory Guided Unsupervised Learning (TGUL)

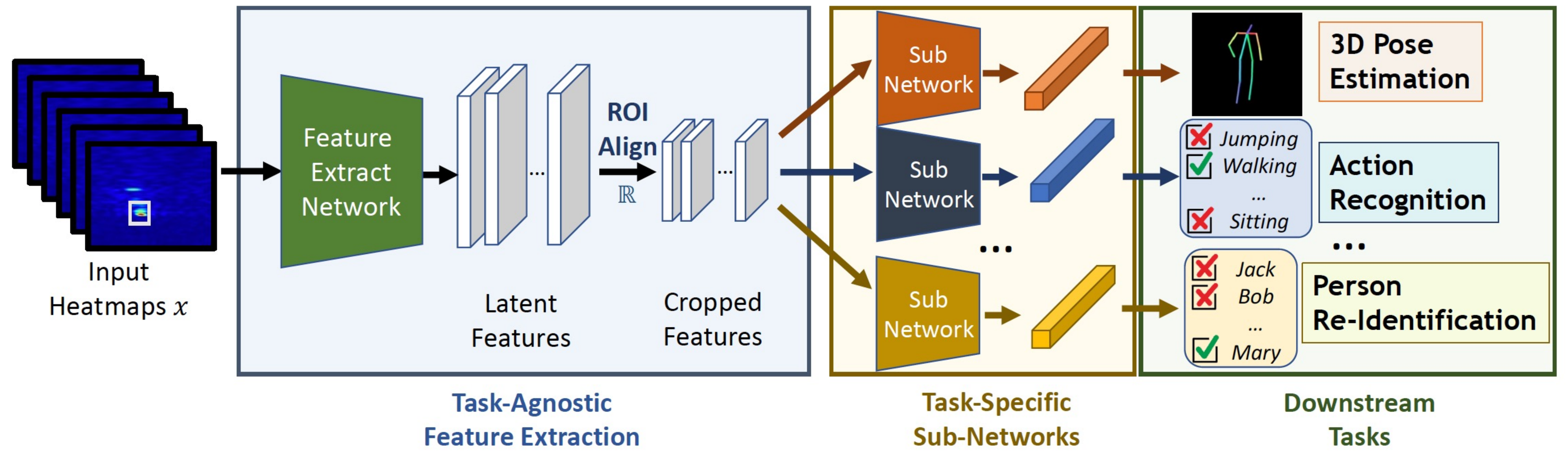
Predictive Unsupervised Learning (PL)



Contrastive Unsupervised Learning (CL)



Experimental Results



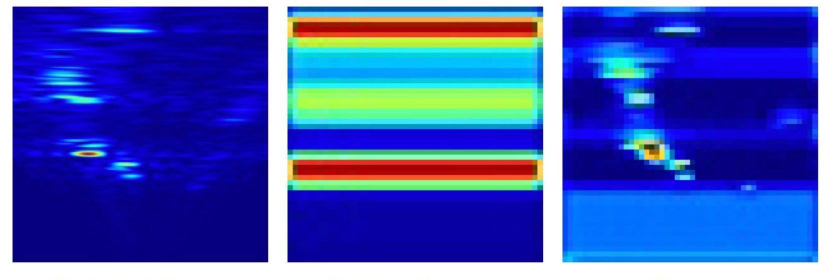
Fixed feature extractor + Fine-tune task-specific parameters:

Tasks	3D Pose Estimation	Action Recognition		Person Re-ID (Campus)			Person Re-ID (Home)		
Metrics	Pose Err.$^\downarrow$ (mm)	mAP$^\uparrow$ $\theta=0.1$	mAP$^\uparrow$ $\theta=0.5$	mAP$^\uparrow$	CMC-1$^\uparrow$	CMC-5$^\uparrow$	mAP$^\uparrow$	CMC-1$^\uparrow$	CMC-5$^\uparrow$
Random init	60.1	60.5	53.3	28.1	43.8	68.8	30.1	54.2	74.6
SimCLR + TGUL	80.5	4.2	0	29.8	44.1	67.5	31.2	55.1	73.8
MoCo + TGUL	77.2	5.1	0.18	29.1	44.7	65.3	30.5	54.5	74.0
CPC + TGUL	78.7	3.6	0	30.0	42.7	69.5	30.7	54.0	75.3
BYOL + TGUL	79.3	4.7	0	29.5	44.4	66.7	30.7	54.6	73.5
Autoencoder	59.4	62.3	54.2	29.0	44.5	67.0	31.1	55.5	75.5
Autoencoder + TGUL	55.7	71.1	63.2	43.8	69.7	87.2	35.2	61.5	81.9
Inpainting	58.0	63.9	55.4	30.2	48.1	70.5	32.8	57.7	76.5
Inpainting + TGUL	**51.1**	**72.3**	**65.5**	**49.8**	**73.1**	**90.5**	**38.5**	**64.2**	**84.7**
Improvement	+15.0%	+19.5%	+22.9%	+77.2%	+66.9%	+31.5%	+27.9%	+18.5%	+13.5%

Fine-tune all parameters:

Tasks	3D Pose Estimation	Action Recognition		Person Re-ID (Campus)			Person Re-ID (Home)		
Metrics	Pose Err.$^\downarrow$ (mm)	mAP$^\uparrow$ $\theta=0.1$	mAP$^\uparrow$ $\theta=0.5$	mAP$^\uparrow$	CMC-1$^\uparrow$	CMC-5$^\uparrow$	mAP$^\uparrow$	CMC-1$^\uparrow$	CMC-5$^\uparrow$
Supervised [23, 11]	38.4	90.1	87.8	59.5	82.1	95.5	46.4	74.6	89.5
SimCLR + TGUL	38.8	89.8	87.4	59.0	81.7	94.1	45.9	73.8	88.5
MoCo + TGUL	38.3	89.7	87.2	59.3	82.0	94.5	46.4	74.3	89.7
CPC + TGUL	38.6	89.9	87.5	59.4	81.5	94.0	46.0	74.5	89.1
BYOL + TGUL	38.5	89.7	87.2	59.4	81.9	94.5	46.6	74.5	89.5
Autoencoder	38.5	90.0	87.7	59.1	81.9	95.5	45.9	74.2	88.6
Autoencoder + TGUL	37.5	91.2	87.9	59.7	82.8	95.5	46.8	74.6	89.8
Inpainting	38.2	90.5	88.0	59.3	82.1	95.7	46.2	74.4	89.2
Inpainting + TGUL	**36.2**	**91.7**	**88.7**	**60.1**	**83.3**	**95.9**	**47.5**	**75.3**	**90.3**
Improvement	+5.7%	+1.8%	+1.0%	+1.0%	+1.5%	+0.4%	+2.4%	+0.9%	+0.9%

Feature Visualization: CL vs. PL



(a) Original Heatmap (b) CL Features (c) PL features

With more unlabeled data on 3D pose estimation

Methods	Pose Err.$^\downarrow$ (mm)
Training from scratch (RF-MMD-S)	48.7
Inpainting on RF-MMD-S+finetune	46.1
Inpainting on RF-MMD+finetune	43.2