GENERATIVE AI-EMPOWERED RF SENSING FOR 3D HUMAN POSE TRACKING, AUGMENTATION AND COMPLETION

Nagoya University, Nagoya, Japan & IEEE ComSoc Tokyo (Joint) Chapter

Shiwen Mao, Auburn University
Joint work with: Ziqi Wang, Chao Yang, and Xuyu Wang

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Outline

- **Human pose tracking: preliminaries and approaches**
- RFID-Pose: 3D human pose monitoring using RFID [1], and its extensions [2,3]
- Generative AI for data augmentation [4-9]
- Generative AI for 3D pose augmentation and completion [10,11]
- Conclusions

[1] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, vol.70, no.3, pp.1218-1231, Sept. 2021.
[2] C. Yang, L. Wang, X. Wang, and S. Mao, "Environment adaptive RFID based 3D human pose tracking with a meta-learning approach," *IEEE Journal of Radio Frequency Identification*, to appear. DOI: 10.1109/JRFID.2022.3140256.
[3] C. Yang, X. Wang, and S. Mao, "TARF: Technology-agnostic RF sensing for human activity recognition," *IEEE Journal of Biomedical and Health Informatics*, vol.27, no.2, pp.636--647, Feb. 2023.
[4] Z. Wang, C. Yang, and S. Mao, "Data augmentation for RFID-based 3D human pose tracking," in *Proc. IEEE VTC-Fall 2022*, London, UK, Sept. 2022.
[5] C. Yang, Z. Wang, and S. Mao, "RFPose-GAN: Data augmentation for RFID based 3D human pose tracking," in *Proc. The 12th IEEE International Conference on RFID Technology and Applications* (*IEEE RFID-TA 2022*), Cagliari, Italy, Sept. 2022, pp.138-141.
[6] Z. Wang and S. Mao, "AIGC for RF sensing: The case of RFID-based human activity recognition," in *Proc. ICNC 2024*, Big Island, HI, Feb. 2024, pp.1092-1097.
[7] Z. Wang and S. Mao, "AIGC for wireless data: The case of RFID-based human activity recognition," in *Proc. IEEE ICC 2024*, Denver, CO, June 2024, pp. 1–6.
[8] Z. Wang, C. Yang, and S. Mao, "AIGC for RF-based human activity sensing," *IEEE Internet of Things Journal*, vol.12, no.4, pp.3991-4005, Feb. 2025.
[9] Z. Wang and S. Mao, "AIGC for Wireless Sensing: Diffusion-empowered Human Activity Recognition," *IEEE Transactions on Cognitive Communications and Networking*, vol.11, no.2, pp.657-671, Apr. 2025
[10] Z. Wang and S. Mao, "Generative AI for 3D human pose completion under RFID sensing constraints," in *Proc. ICNC 2025*, Honolulu, HI, Feb. 2025, pp.485-490.
[11] Z. Wang and S. Mao, "Generative AI-empowered RFID sensing for 3D human pose augmentation and completion," *IEEE Open Journal of the Communications Society*, vol.6, pp.2958-2975, Feb. 2025.

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Human Skeleton Detection and Pose Tracking

Human pose tracking: an important problem of human-computer interaction

Activity recognition

- Full-body sign language interpretation (e.g., hand signals of traffic police, aircraft ground handling)
- Fall detection
- Security/safety surveillance

Motion capture and augmented reality

Somatosensory games

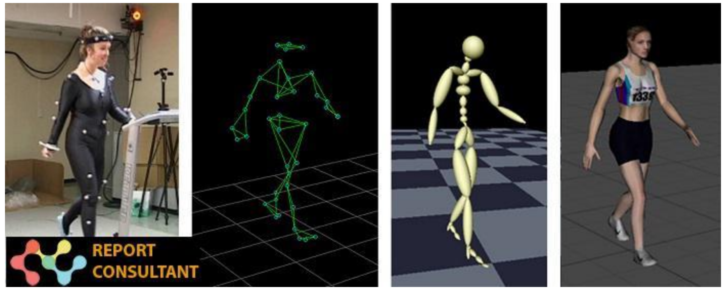


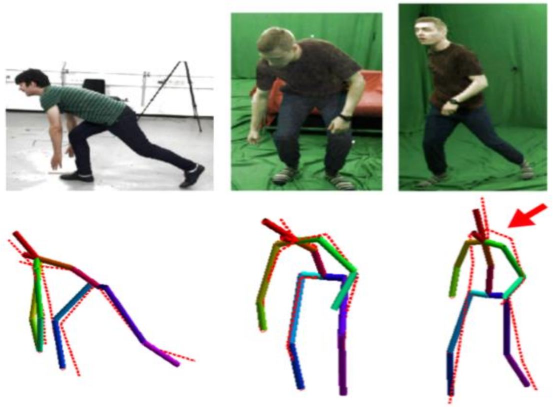
Image Source: *https://medium.com/@victoriamazo/3d-human-pose-estimation-ce1259979306*
Image Source: *https://www.ubisoft.com/en-us/game/just-dance/2021*
Image Source: *https://www.openpr.com/news/1345254/3d-motion-capture-market-witness-a-consistent-growth-in-the-forecast-years-with-the-key-vendors-phoenix-technologies-codamotion-solutions-vicon-motion-analysis-corporation-optitrack.html*

Apple Vision Pro: Spatial Computing

Image Source: https://forums.macrumors.com/threads/how-gestures-work-on-apple-vision-pro.2391964
Image Source: https://www.youtube.com/shorts/QcYK1T1wD6E?feature=share

Traditional Camera based Approaches

- Evolving from (i) 2D to 3D, and (ii) single person to multiple people
- Performance limited by *poor lighting*, *cluttered background*, *occlusion*, or *camera angle*



Security and privacy concerns:

Toronto

Private moments captured on home security cameras being live streamed again on website

Authorities have tried to stop the site, but streaming unsecured cameras isn't illegal

Angelina King, Jason Lo · CBC News · Posted: Jun 29, 2021 4:00 AM ET | Last Updated: June 29

These images were captured on a website that live streams unsecured security cameras from inside homes and businesses across Canada. Clockwise, from top left: an elderly woman is fed in her room, which includes a commode toilet; two women eat lunch in a hair salon; kitchen staff prepare lunch at a restaurant; and a woman leaves her home to take her dog for a walk. (CBC)

Image Source: *https://www.cbc.ca/news/canada/toronto/website-live-streaming-security-cameras-private-1.6083168*

Z. Cao, et al. "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields," *IEEE transactions on pattern analysis and machine intelligence* 43.1 (2019): 172-186.
D. Mehta, et al. "Vnect: Real-time 3d human pose estimation with a single RGB camera," *ACM Transactions on Graphics* (TOG) 36.4 (2017): 1-14.

RF Sensing-based Human Pose Tracking

Strengths:

- No lighting requirements
- Less intrusive and better preserves the privacy of users
- Works through walls and obstacles

Main challenges:

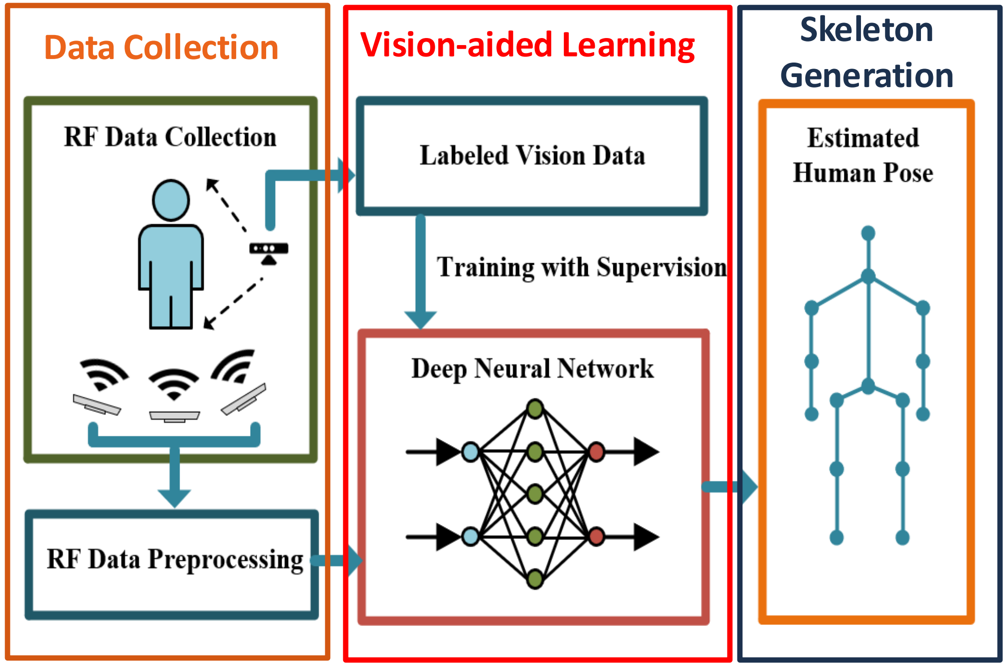
- Motion related feature extraction
- Mapping from RF features to human pose
- Continuously tracking the movements of human limbs: static pose vs. in motion
- Interference from the environment

FMCW Radar, WiFi, mmWave, etc.

A mapping solution:

- Multimodal Learning based approaches

Vision-assisted learning



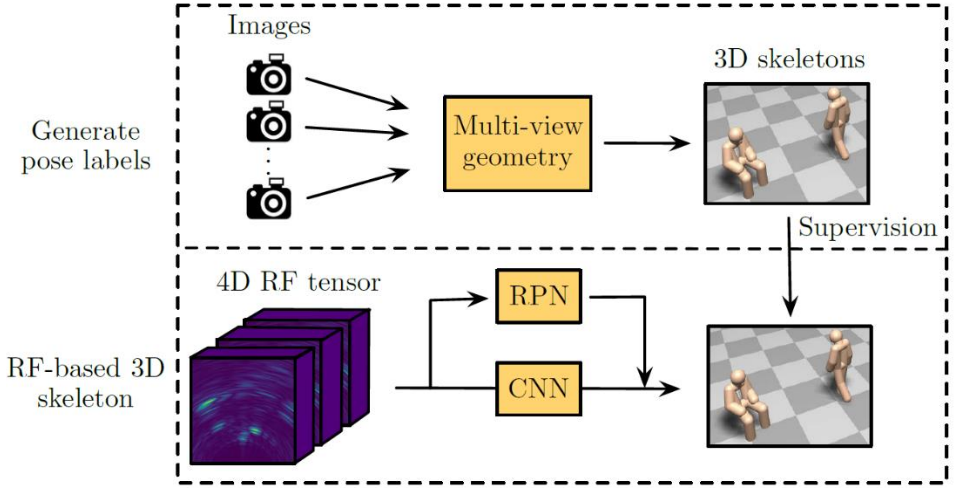
Radar-based Approaches

Strengths:

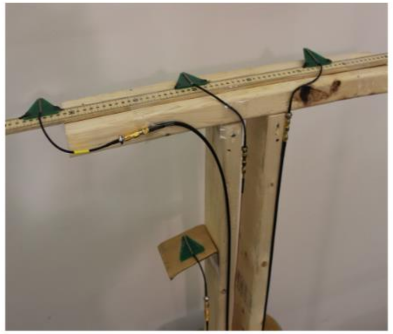
- High accuracy
- Capable of tracking multiple subjects
- Capable of through-wall detection
- More robust to environmental interference than WiFi-based systems

Limitations:

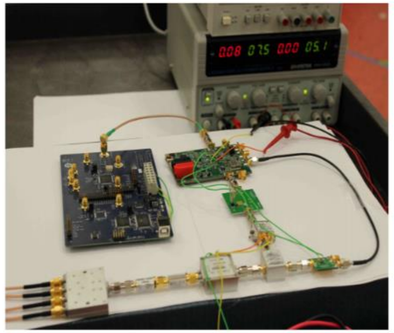
- Implemented with Software-Defined Radio (SDR) and 16 synchronized T-shaped antenna arrays [3]
    - Complicated system and high cost
- Both antenna placement and synchronization need careful calibration



RF-Pose3D system overview (RPN: region proposal network)



(a) **Antenna "T" Setup**



(b) **FMCW Signal Generation**

FMCW radar setup and signal generation

[1] M. Zhao, et al., "Through-wall human pose estimation using radio signals," in *Proc. IEEE CVPR 2018*, Salt Lake City, UT, June 2018, pp. 7356–7365.

[2] M. Zhao, et al., "RF-based 3D skeletons," in *Proc. ACM SIGCOM 2018,* Budapest, Hungary, Aug. 2018, pp. 267–281.

[3] F. Adib, et al., "3D tracking via body radio reflections," in *Proc. 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI'14),* Seattle, WA, Apr. 2014.

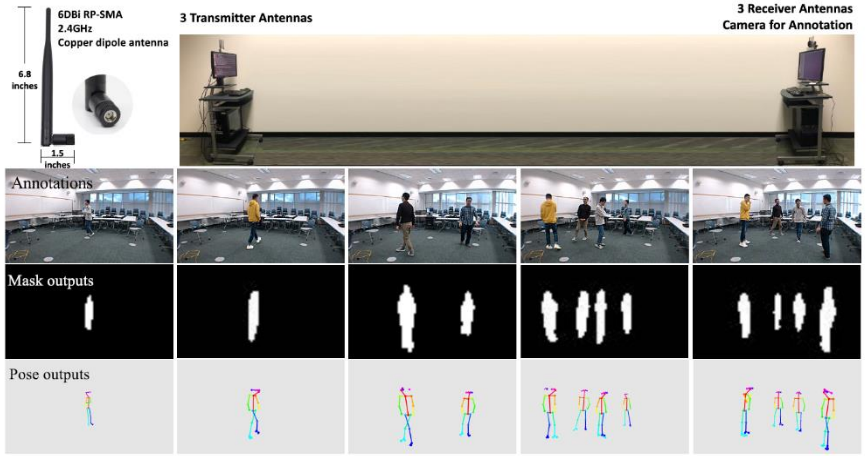
WiFi-based Techniques

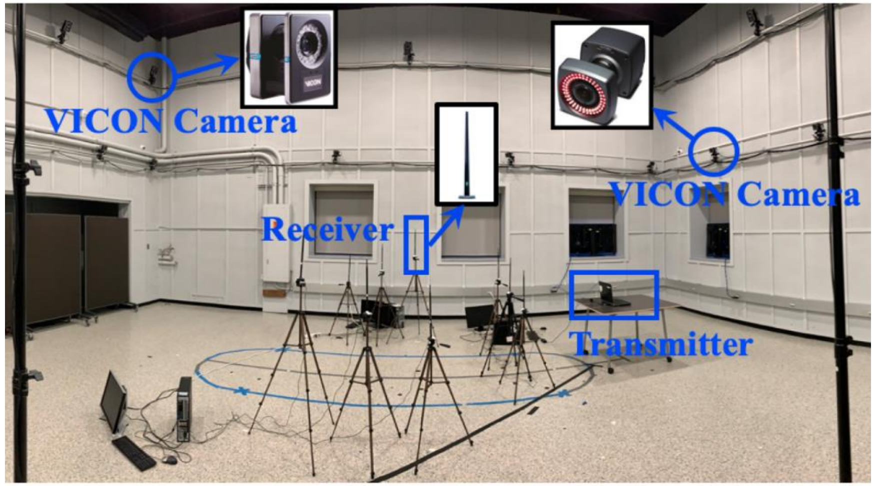
Strengths:

- Less intrusive, and a wide range of detection
- 2D pose estimation for multiple subjects [1] and 3D pose generation for a single subject [2]
- Commodity devices, low-cost hardware

Limitations:

- Sensitive to interference from the testing environment (e.g., moving people or objects, obstacles, etc.)
- Expensive VICON system





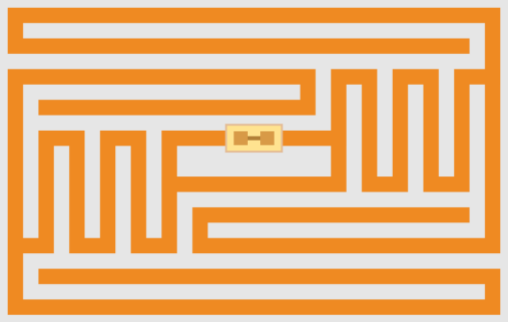
[1] F. Wang, et al., "Person-in-WiFi: Fine-grained person perception using WiFi," in *Proc. IEEE ICCV 2019*, Seoul, Republic of Korea, Oct. 2019, pp. 5452–5461.
[2] W. Jiang, et al., "Towards 3D human pose construction using WiFi," in *Proc. ACM MobiCom'20*, London, UK, Sept. 2020, pp. 1–14.

RFID: Communication Based Applications

Electronic Product Code (EPC): a universal identifier providing a unique identity for every physical object anywhere in the world (96 to 496 bits)

- Person identification
- Vehicle parking monitoring
- Fast-lane and E-Zpass road toll system
- Secure entry cards
- Supply chain management
- Food distribution control

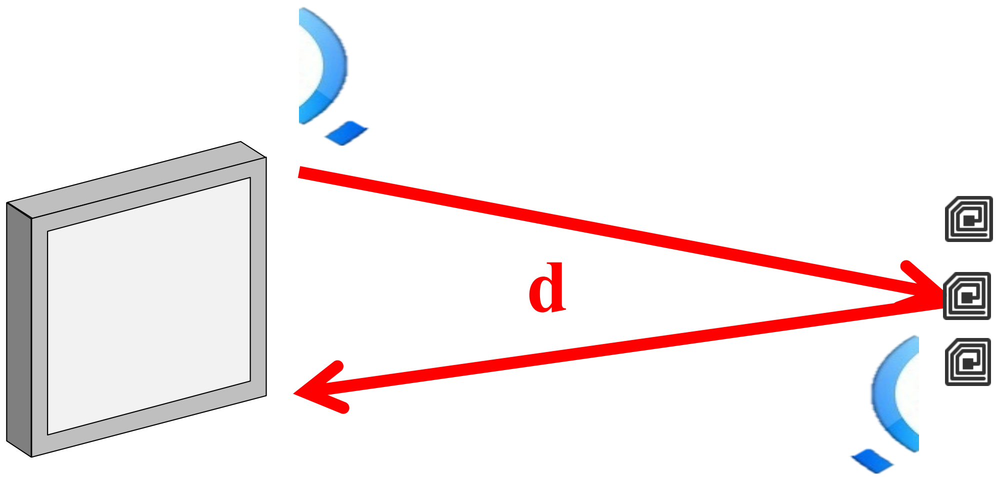
Communication ➔ deliver stored data when being queried

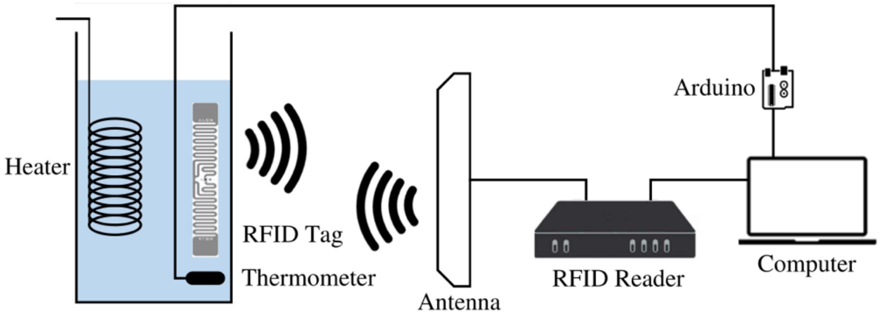


An EPC RFID tag used by Wal-Mart

Image Source: *https://medicalfuturist.com/rfid-implant-chip/*
Image Source: *https://www.wikiwand.com/en/Electronic_Product_Code*
Image Source: *https://www.atlasrfidstore.com/marathon-uhf-rfid-shoe-tag/*
Image Source: *https://pilotonline.com/news/local/transportation/article_62a3b00e-64fb-11e8-88d9-5fbb5a27dbe8.html*

RFID: RF Sensing Based Applications





Phase of the received signal:

$$\varphi = \mathrm{mod}\left(\frac{2\pi d}{\lambda} + \alpha_T + \alpha_R + \alpha_{Tag},\ 2\pi\right)$$

Wireless Channel ➔ RF phase angle, Doppler frequency, and Peak RSSI

RFID based sensing applications:

- Indoor localization
- Temperature measurement
- Gesture recognition
- Vital signal monitoring
- Driving fatigue detection

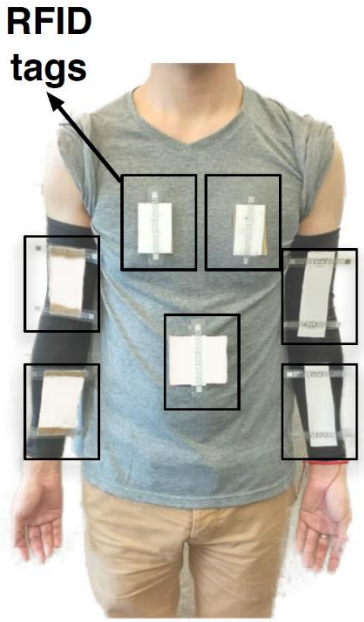


Image Source: https://senion.com/indoor-positioning-for-retail/
Image Source: https://www.psychologytoday.com/us/blog/the-fifth-vital-sign/201711/why-cognition-should-be-the-fifth-vital-sign

Existing RFID based Pose Tracking Systems

Angle-of-arrival (AoA)-based limb orientation monitoring [4,5]:

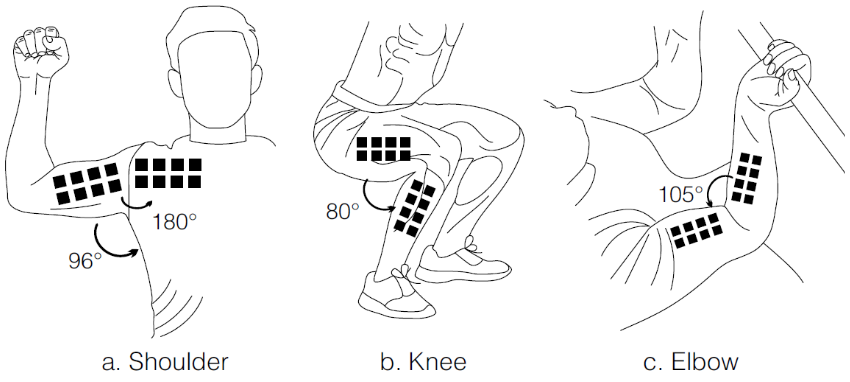
- Utilizing RFID tag arrays
- Angle estimation with the RF hologram technique

Limitations:

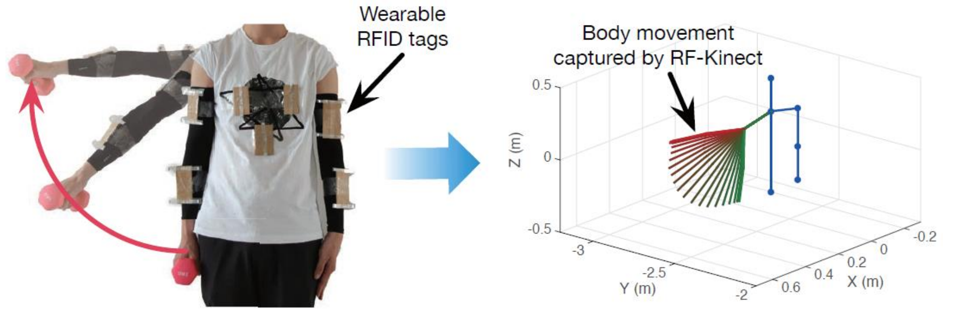
- Many tags are needed for monitoring the entire body
- Generating real-time RF hologram map is challenging

Observations:

- Using AoA to localize multiple tags in realtime is very challenging (not ML based)
- Multimodal learning shall be helpful



RF-Wear tracks the user's skeleton using passive RFID tags [4]



RF-Kinect: Tracking the body movement based on wearable RFID tags [5]

[4] H. Jin, Z. Yang, S. Kumar, and J. I. Hong, "*Towards wearable everyday body-frame tracking using passive RFIDs," Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–23, Dec. 2018.

[5] C. Wang, J. Liu, Y. Chen, L. Xie, H. B. Liu, and S. Lu, "RF-Kinect: A wearable RFID-based approach towards 3D body movement tracking," *Proc. ACM Int., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, Mar. 2018.

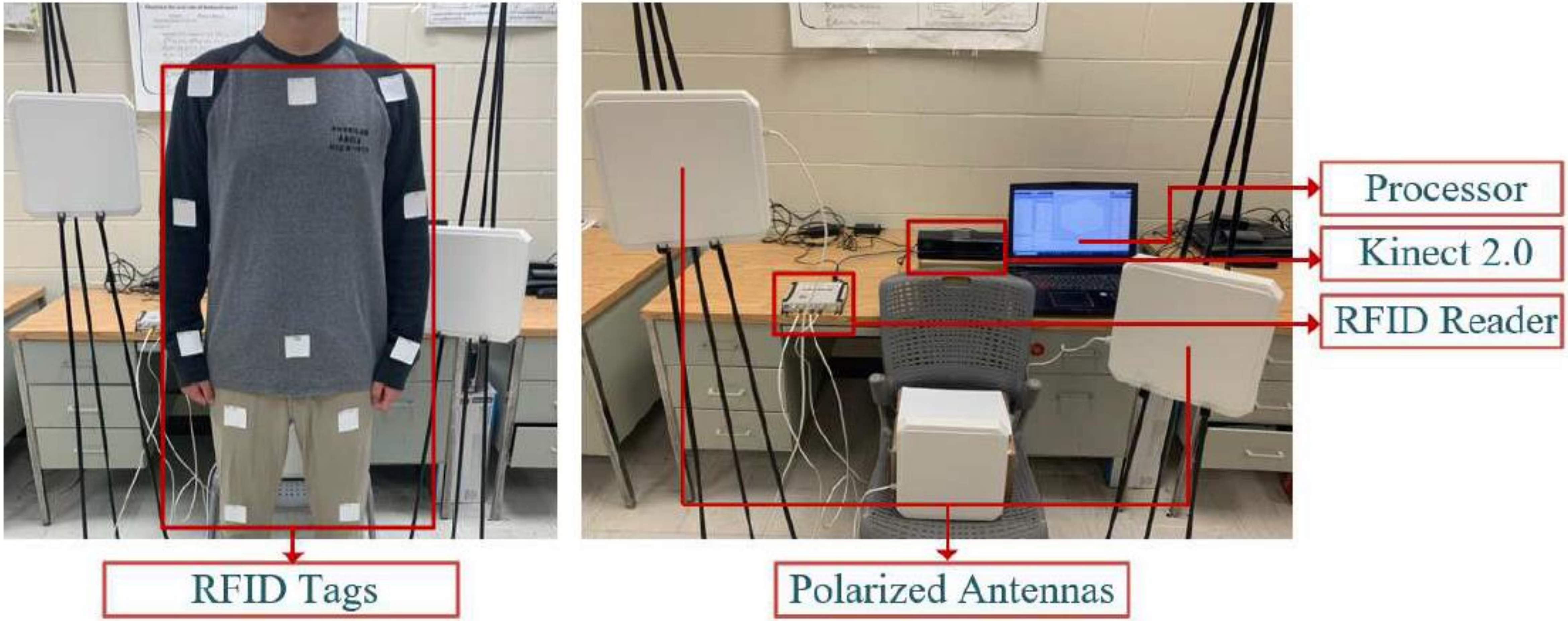
Outline

- Human pose tracking: preliminaries and approaches
- **RFID-Pose: 3D human pose monitoring using RFID [1], and its extensions** [2,3]
- Generative AI for data augmentation [4-9]
- Generative AI for 3D pose augmentation and completion [10,11]
- Conclusions

[1] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, vol.70, no.3, pp.1218-1231, Sept. 2021.
[2] C. Yang, L. Wang, X. Wang, and S. Mao, "Environment adaptive RFID based 3D human pose tracking with a meta-learning approach," *IEEE Journal of Radio Frequency Identification*, to appear. DOI: 10.1109/JRFID.2022.3140256.
[3] C. Yang, X. Wang, and S. Mao, "TARF: Technology-agnostic RF sensing for human activity recognition," *IEEE Journal of Biomedical and Health Informatics*, vol.27, no.2, pp.636--647, Feb. 2023.
[4] Z. Wang, C. Yang, and S. Mao, "Data augmentation for RFID-based 3D human pose tracking," in *Proc. IEEE VTC-Fall 2022*, London, UK, Sept. 2022.
[5] C. Yang, Z. Wang, and S. Mao, "RFPose-GAN: Data augmentation for RFID based 3D human pose tracking," in *Proc. The 12th IEEE International Conference on RFID Technology and Applications* (*IEEE RFID-TA 2022*), Cagliari, Italy, Sept. 2022, pp.138-141.
[6] Z. Wang and S. Mao, "AIGC for RF sensing: The case of RFID-based human activity recognition," in *Proc. ICNC 2024*, Big Island, HI, Feb. 2024, pp.1092-1097.
[7] Z. Wang and S. Mao, "AIGC for wireless data: The case of RFID-based human activity recognition," in *Proc. IEEE ICC 2024*, Denver, CO, June 2024, pp. 1–6.
[8] Z. Wang, C. Yang, and S. Mao, "AIGC for RF-based human activity sensing," *IEEE Internet of Things Journal*, vol.12, no.4, pp.3991-4005, Feb. 2025.
[9] Z. Wang and S. Mao, "AIGC for Wireless Sensing: Diffusion-empowered Human Activity Recognition," *IEEE Transactions on Cognitive Communications and Networking*, vol.11, no.2, pp.657-671, Apr. 2025
[10] Z. Wang and S. Mao, "Generative AI for 3D human pose completion under RFID sensing constraints," in *Proc. ICNC 2025*, Honolulu, HI, Feb. 2025, pp.485-490.
[11] Z. Wang and S. Mao, "Generative AI-empowered RFID sensing for 3D human pose augmentation and completion," *IEEE Open Journal of the Communications Society*, vol.6, pp.2958-2975, Feb. 2025.

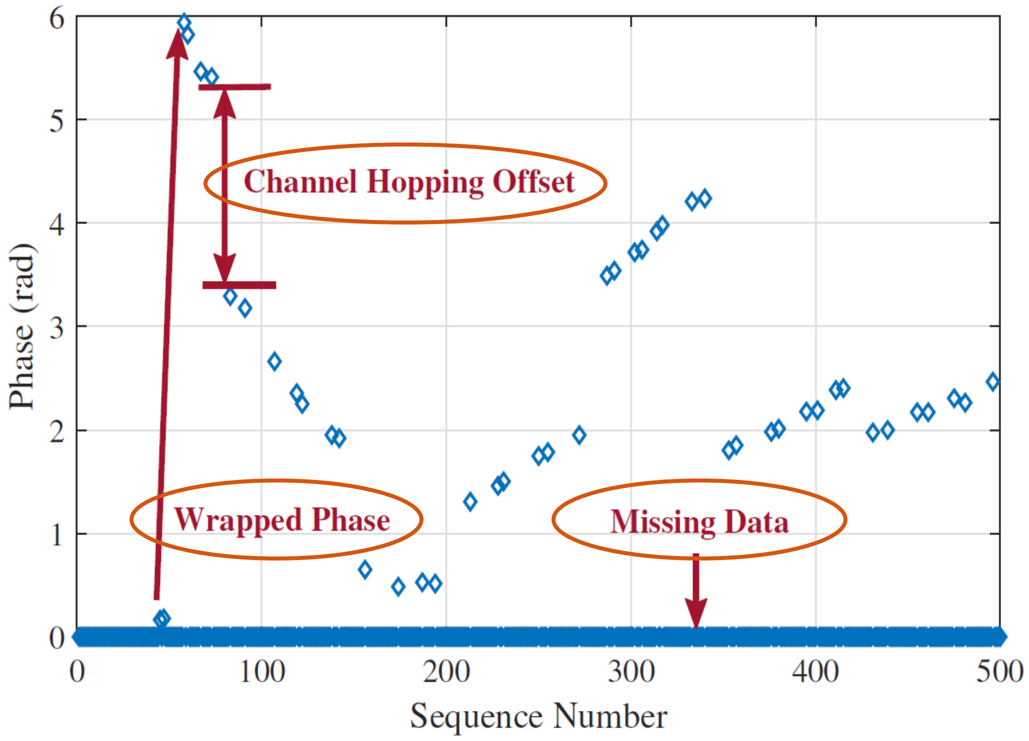
AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

RFID-Pose: Vision-aided 3D Human Pose Estimation



AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Challenges: Noisy and Sparse RFID Data



Raw phase sampled from one of the RFID tags by a single reader antenna

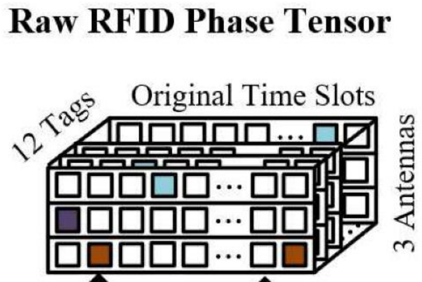
Collected phase for each channel:

$$\phi_s = \boxed{\text{mod}}\left(\frac{2\pi 2 L f_s}{c} + \boxed{\phi_s^0}, 2\pi\right), \quad s = 1, 2, \ldots, 50$$

Channel hopping phase offset of channel *s*

Missing samples in tensor of the RFID data:

Only one tag is sampled by an antenna at a time



Extremely high sparsity: with 12 tags and 3 antennas: 35/36 ≈ 97.22%

Skeleton Generation from RFID Data

Most existing systems are based on the confidence map, which is not suitable for RFID systems (with a 110Hz sampling rate)

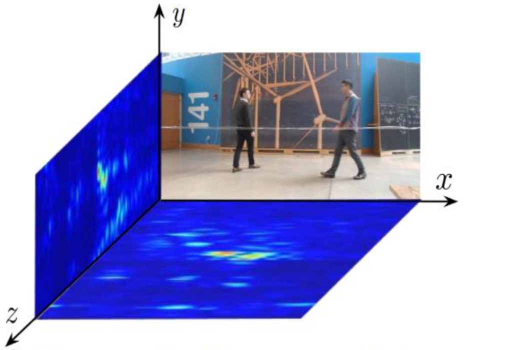
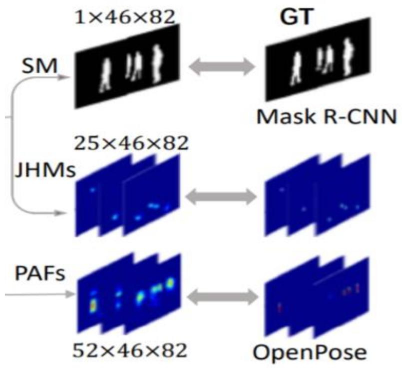


Figure 2: RF heatmaps and an RGB image recorded at the same time.

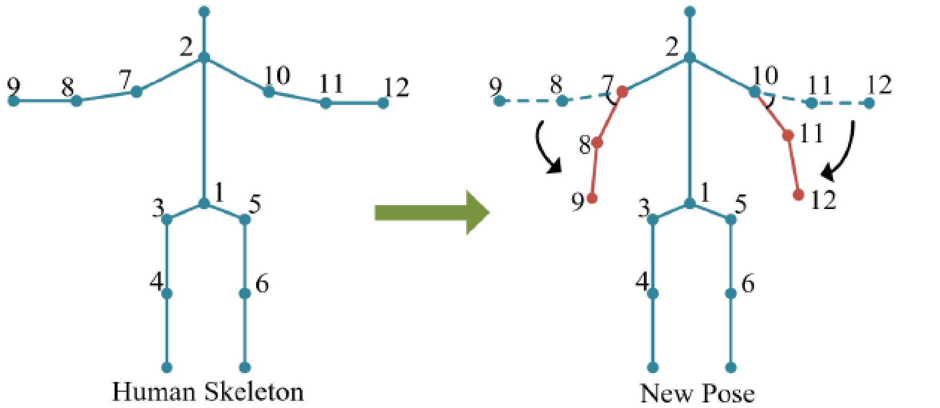


To generate a confidence map video at 10 fps, only 11 phase samples could be used for map generation

Even if we reduce the map resolution to 100×100, transforming the 11 samples to 10,000 pixels in a map is a severely ill-posed problem

The *forward kinematic* technique:

- New location derived from (i) the parent joint location, and (ii) the 3D rotation



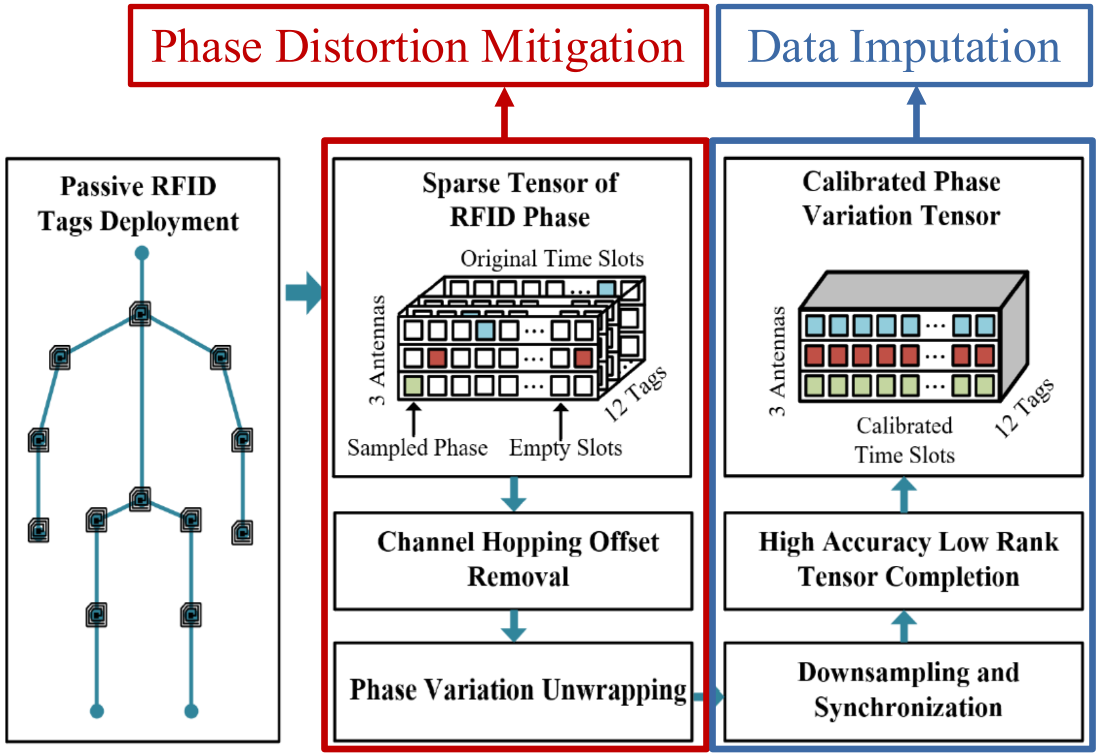
The 3D rotation of each joint for each frame can be represented with 4 parameters (i.e., as a *unit quaternion*)

$$r + x\vec{\alpha} + y\vec{\beta} + z\vec{\gamma}$$

Thus only 48 parameters are needed to estimate the 3D positions of the 12 human joints

M. Zhao, T. Li, M. Abu Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *Proc. IEEE CVPR 2018*, Salt Lake City, UT, June 2018, pp. 7356–7365.
F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang, "Person-in-WiFi: Fine-grained person perception using WiFi," in *Proc. IEEE ICCV 2019,* Seoul, Republic of Korea, Oct. 2019, pp. 5452–5461.

RFID Phase Distortion Mitigation and Data Imputation



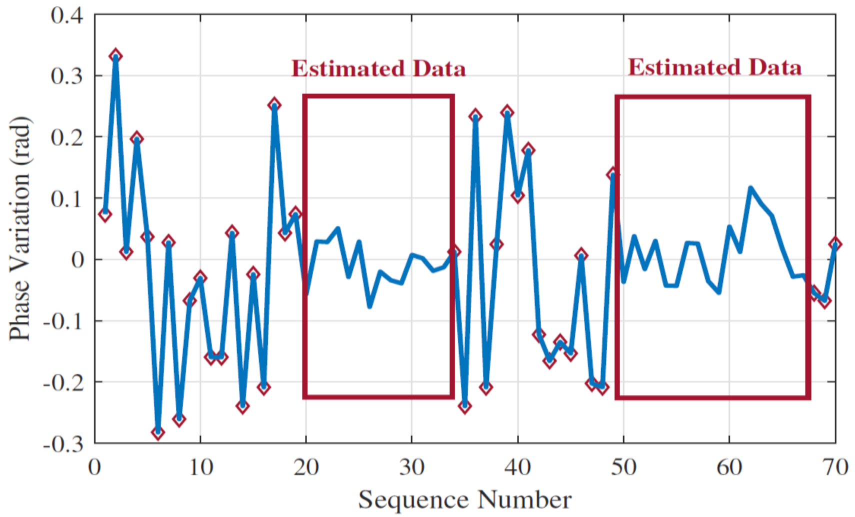
RFID data preprocessing

Phase Distortion Mitigation:

- Tensor construction
- Channel hopping offset mitigation
- Phase variation unwrapping

Data Imputation:

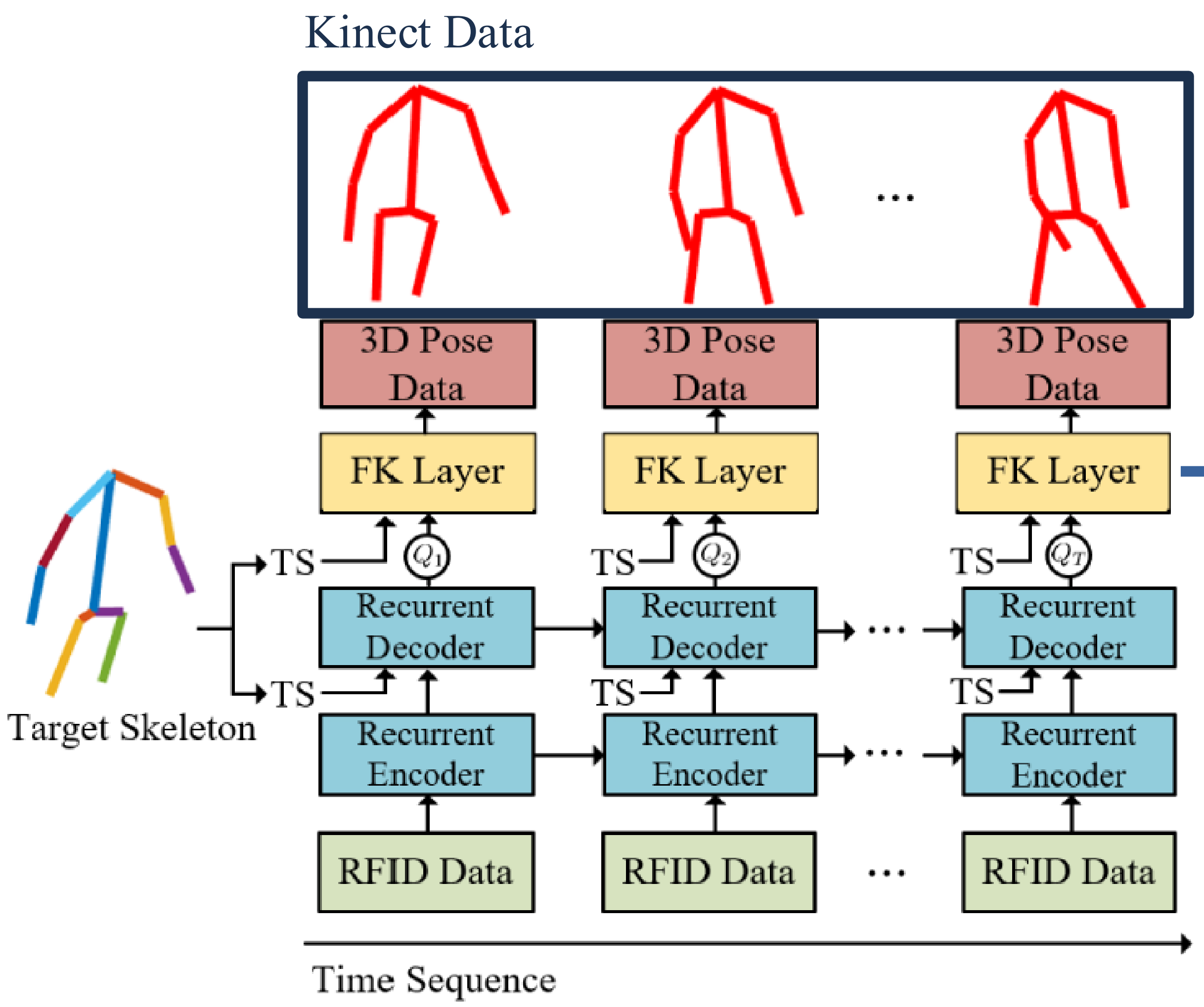
- Downsampling and synchronization
- High Accuracy Low Rank Tensor Completion (HaLRTC)

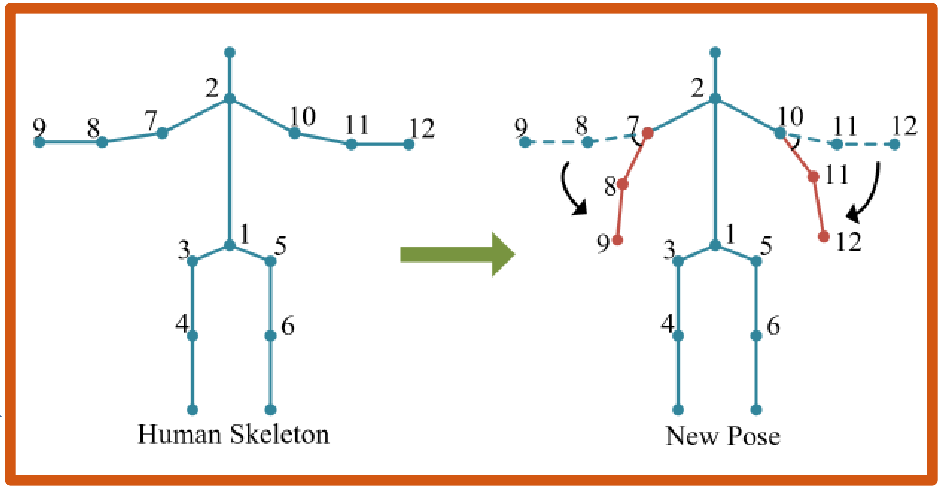


The missing data are estimated by HaLRTC

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

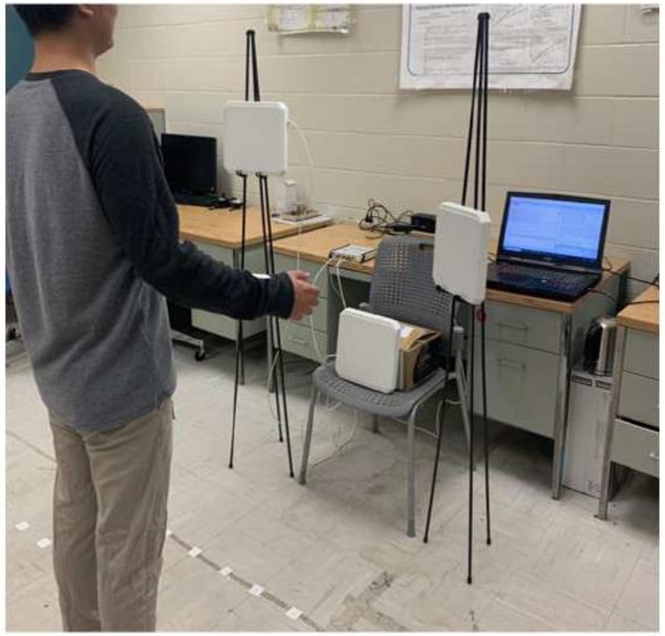
The Deep Kinematic Neural Network Model





- Recurrent Autoencoder (256 gated recurrent units (GRU)):
  RF data ➔ unit quaternion
- Forward kinematic layer:
  Rotation matrix ➔ 3D pose
- Kinect data:
  - labels, for training and performance evaluation
  - Not needed after training the model

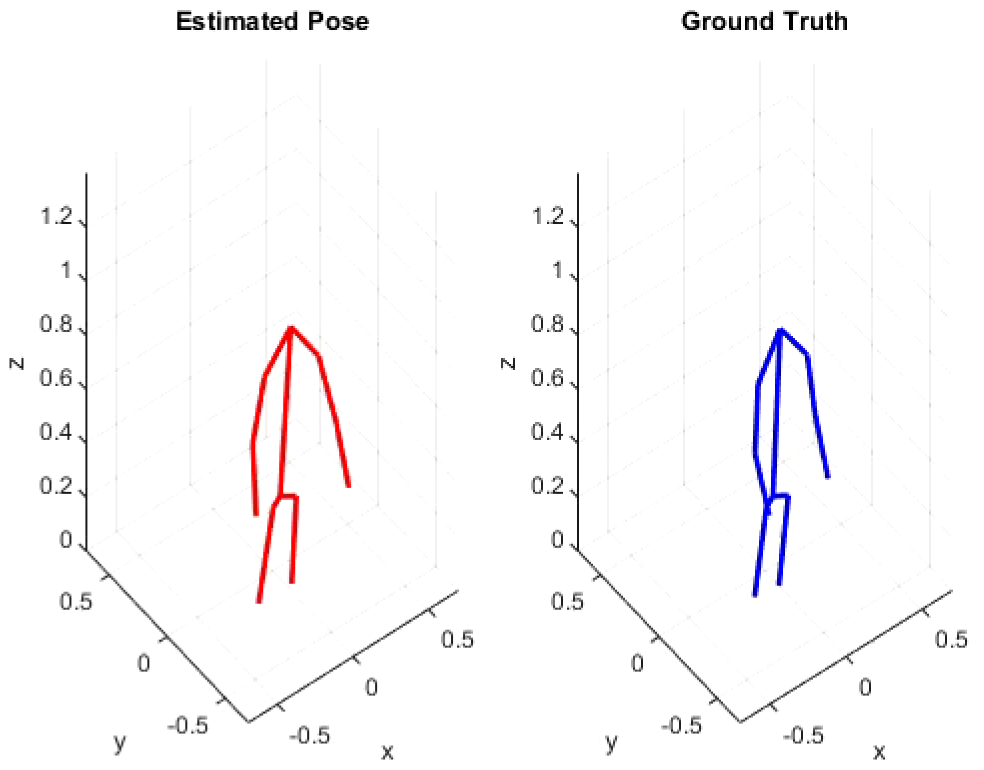
Implementation and Evaluation



Standing Still

Walking

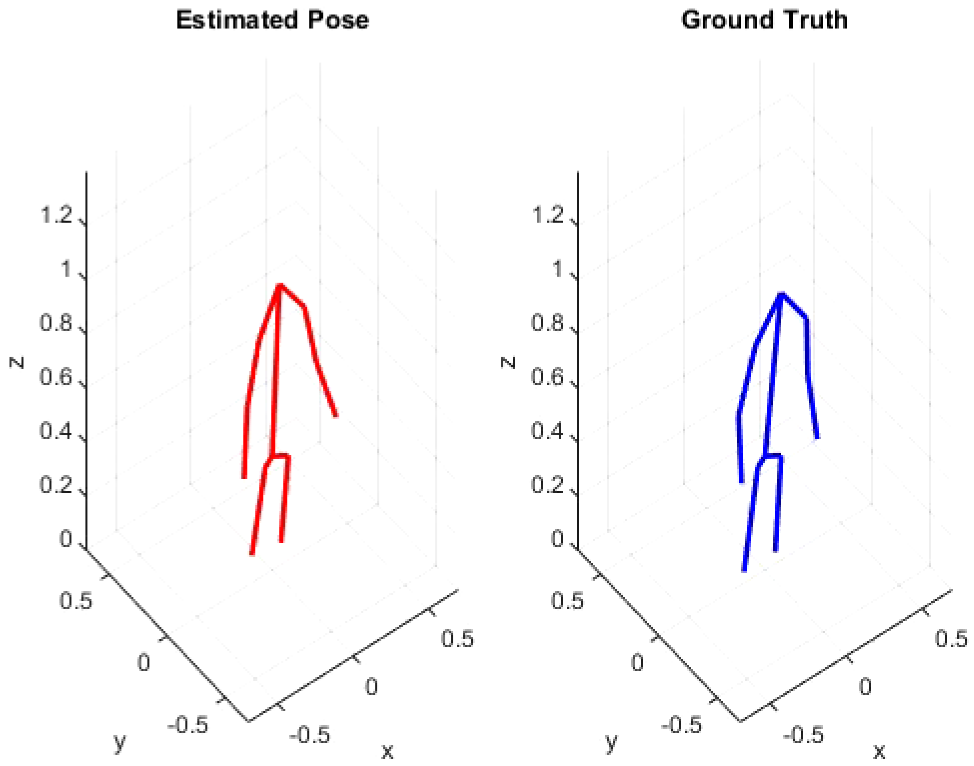
Pose tracking experiments



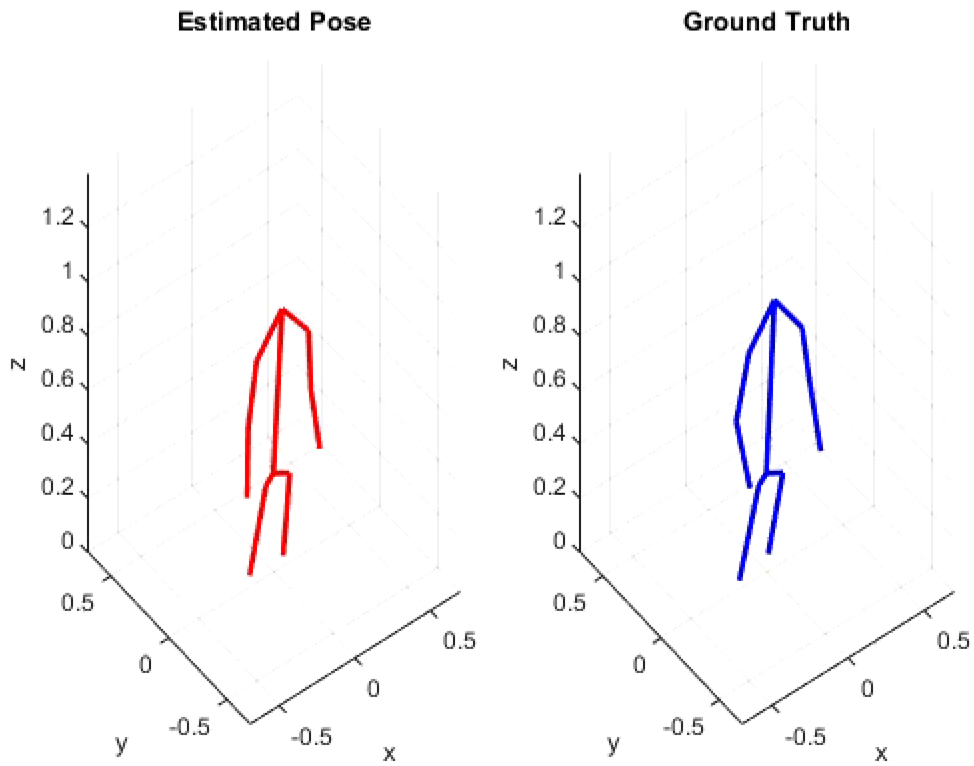
Pose tracking: walking

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Experiment Results: Pose Tracking

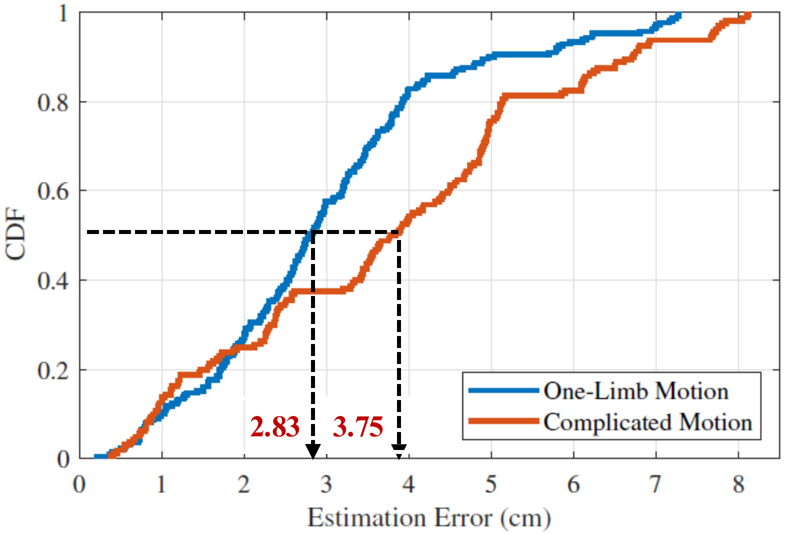


Pose tracking: squatting

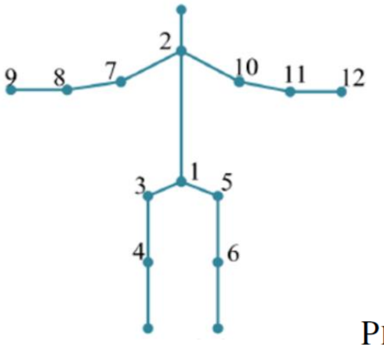


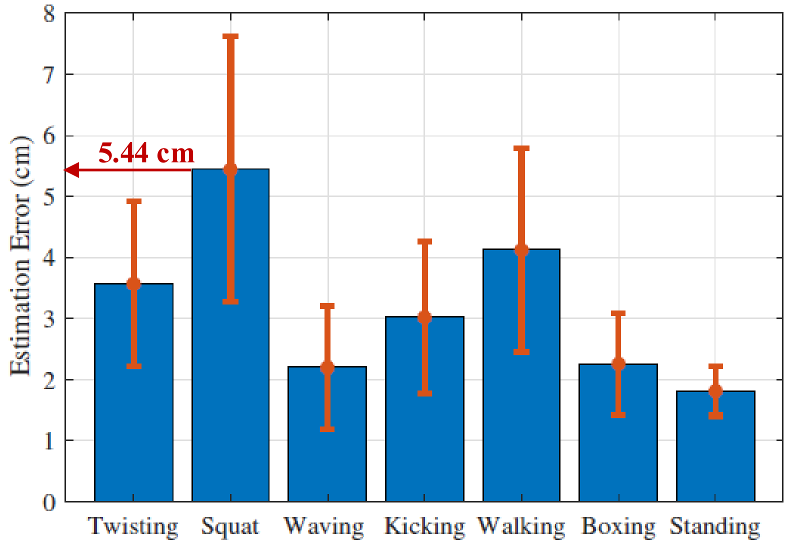
Pose tracking: twisting

Experimental Results: Estimation Error



Overall pose estimation accuracy

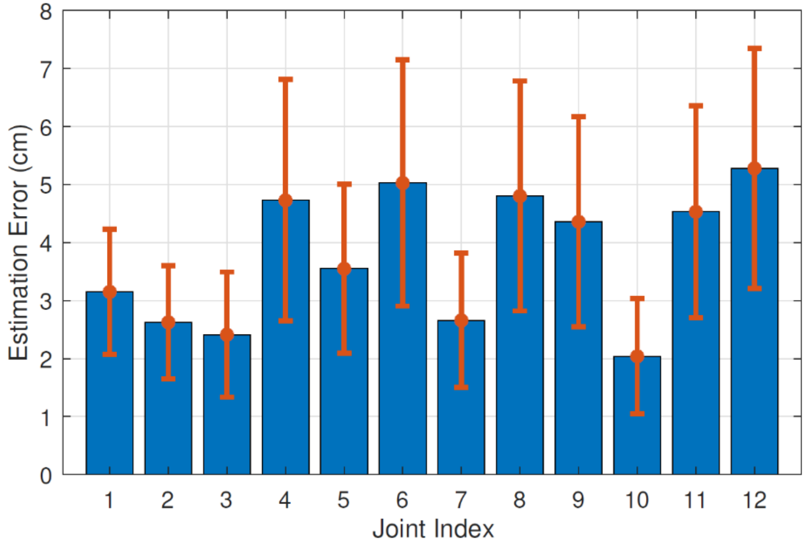




Estimation errors for different types of motions

TABLE III
PERFORMANCE EVALUATION UNDER DIFFERENT ENVIRONMENTS

Testing Environments	*Estimation Error*
Computer Lab-1	3.83cm
Computer Lab-2	3.90cm
Corridor	4.03cm
Living Room	3.75cm



Estimation errors for different joints

Diversity in Different Data Domains

The same activity could generate very different RF data when sampled in different environments

Developing a human pose estimation techniques that are generalizable to different environments ➔ a great challenge for RF sensing

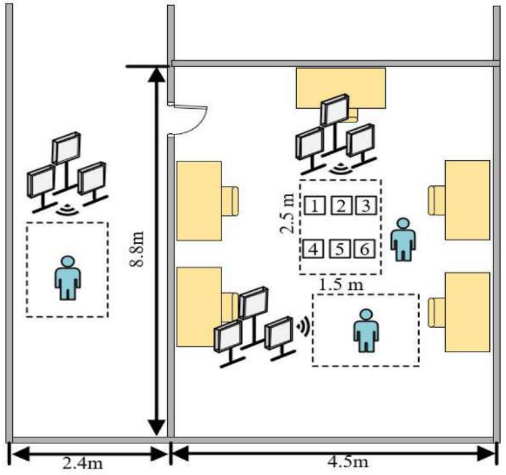
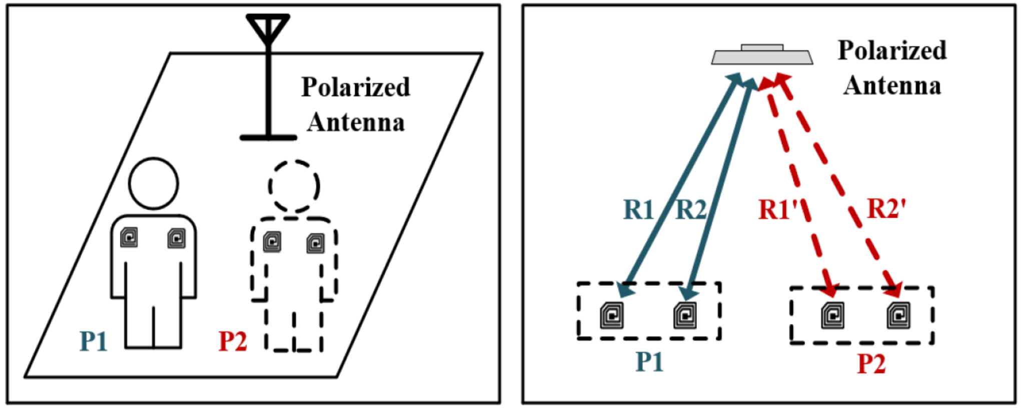
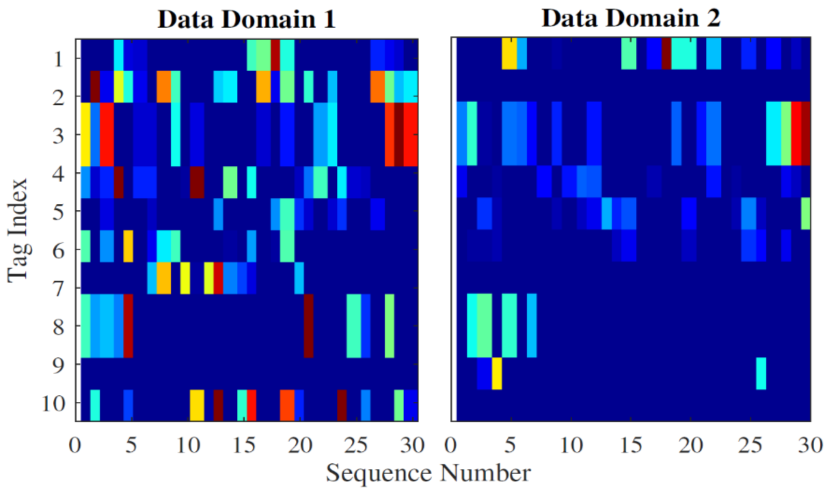


TABLE IV
PERFORMANCE EVALUATION FOR DIFFERENT STANDING POSITIONS

Position Index	*Estimation Error*
Position 1 (Trained)	4.53cm
Position 2 (Trained)	3.82cm
Position 3 (Trained)	4.75cm
Position 4 (Untrained)	8.38cm
Position 5 (Untrained)	5.71cm
Position 6 (Untrained)	9.14cm

Different deployment environments and standing positions

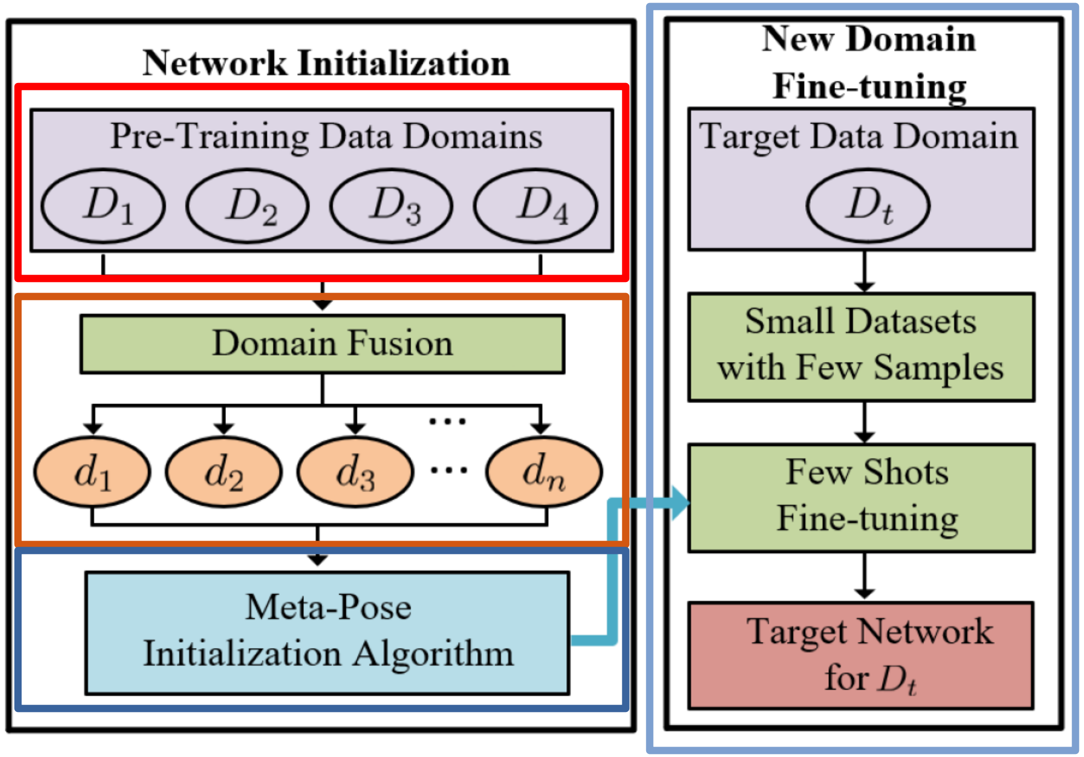




RFID Phase collected in two different environments for the same activity

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Meta-Pose Can Be Helpful



Training framework of Meta-Pose

The deep learning model is pretrained with data from four known data domains

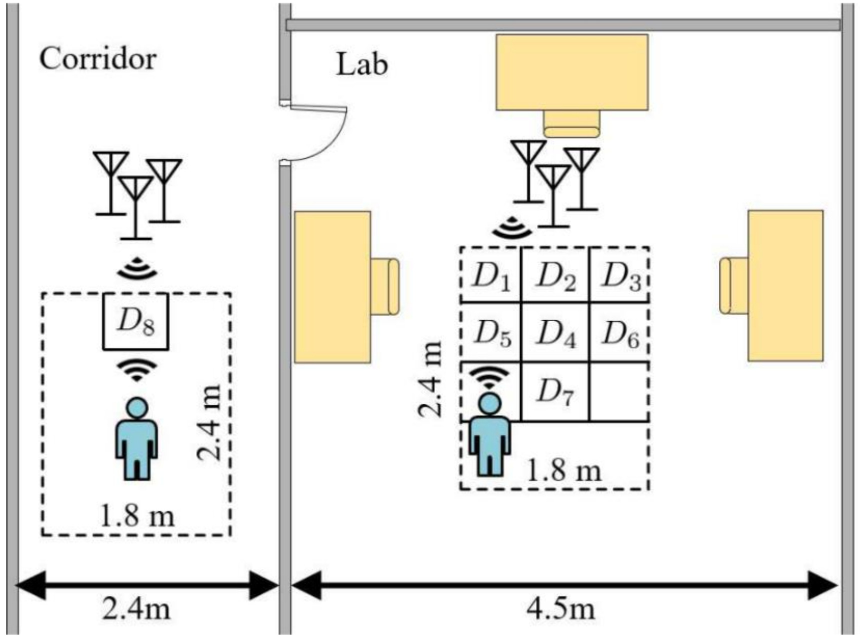
Domain fusion algorithm: to produce more data domains

The training variables are updated recursively by the Reptile and model-agnostic meta-learning (MAML) meta-learning algorithm

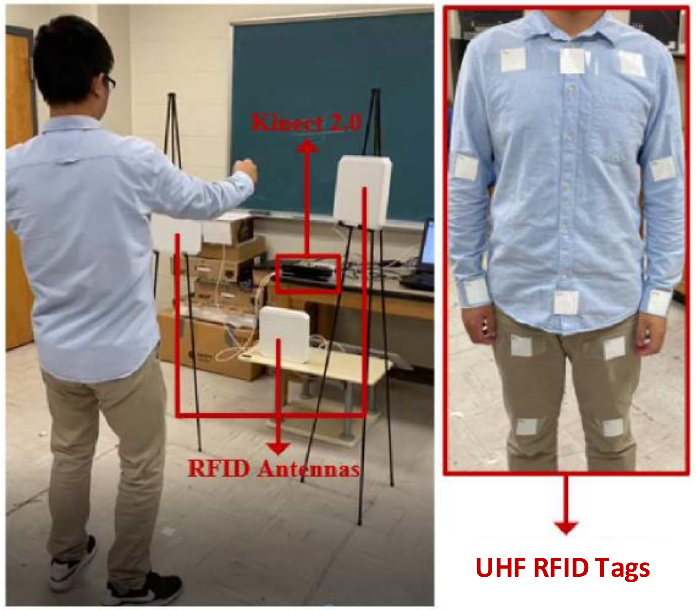
When transferring to a new data domain, we only need to collect a few examples to fine-tune the generalized network

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Implementation and Evaluation



Data domains used in the experiments



Hardware configuration of Meta-Pose

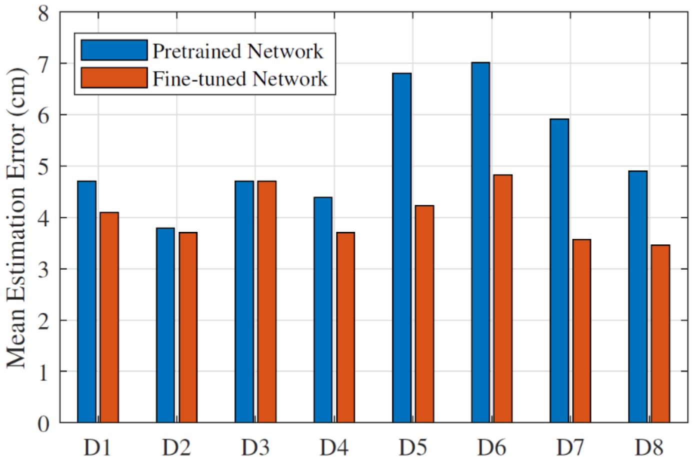
Seven data domains are sampled in the computer lab, and the 8th domain is sampled in an empty corridor

- D1 to D4 are used for pretraining
- D5 to D8 are considered as new data domains for validation

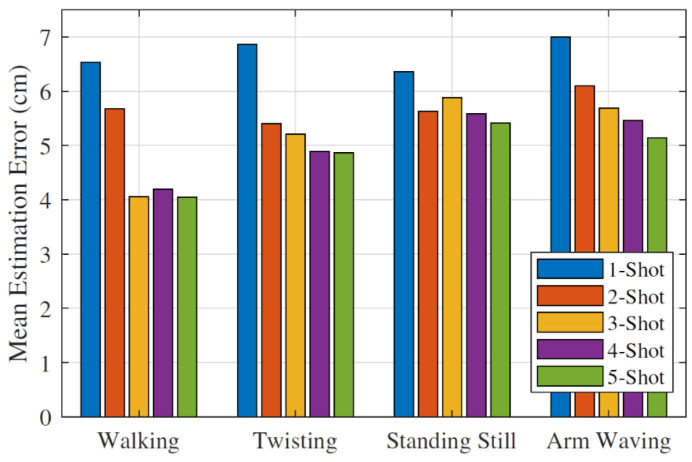
Five subjects participate in the experiments

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

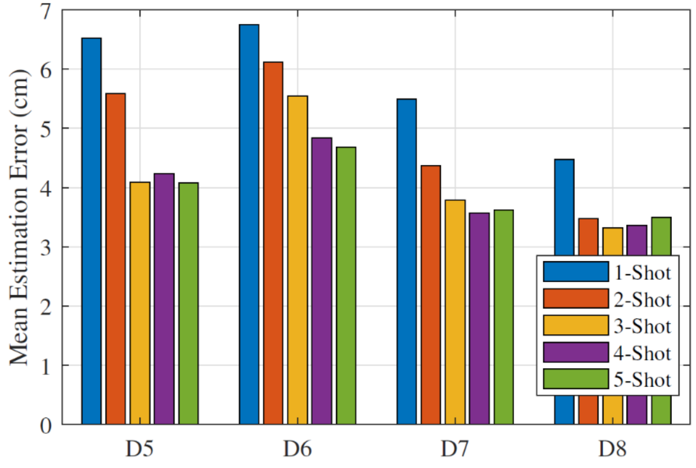
Experimental Results and Analysis



Overall performance in terms of mean estimation error in the eight different data domains



Fine-tuning performance of different activities with different shots of new data in new data domain D5



Fine-tuning performance of different new data domains with different shots of new data

Average error comparison with the baseline method RFID-Pose

Domain Index	*RFID-Pose*	*Meta-Pose*
D_5	6.72cm	3.72cm
D_6	7.62cm	4.32cm
D_7	5.46cm	3.51cm
D_8	4.62cm	4.11cm
D_{all}	6.27cm	3.97cm

One shot of data is defined as consecutive samples for 6 seconds

With few-shot fine-tuning, the mean error of all the new data domains is 3.97cm, which is very similar to that of the pretrained data domains

4-shot fine-tuning is sufficient; the minimum error is achieved when walking

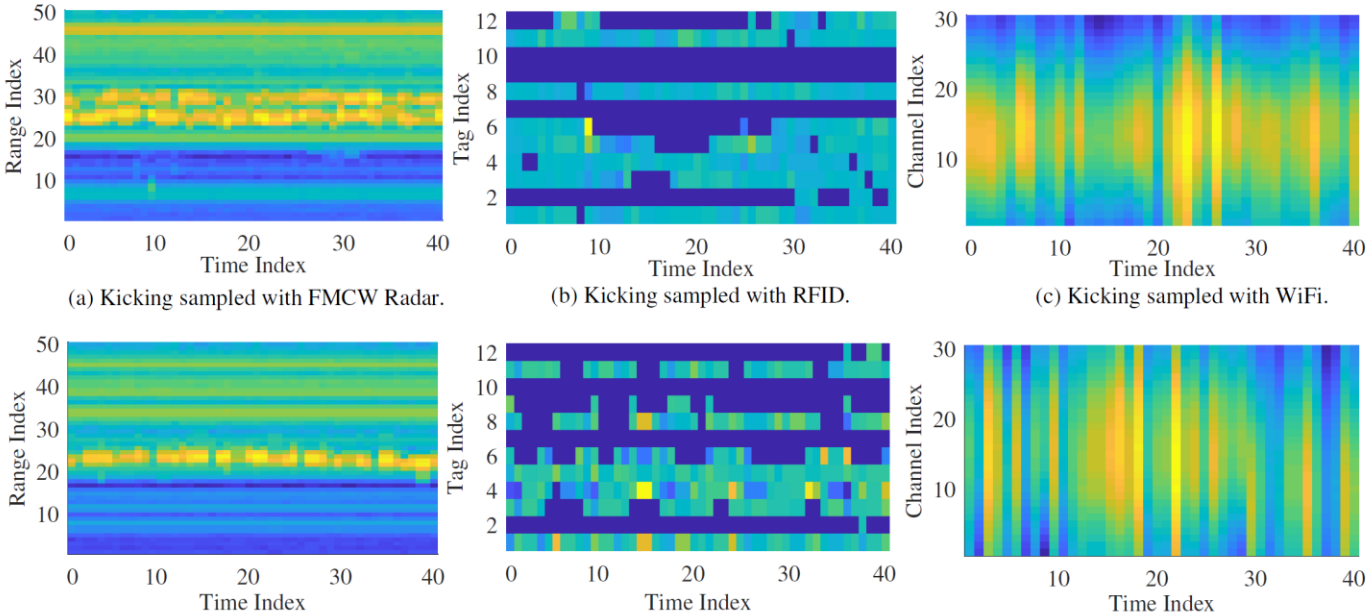
Mean error of RFID-Pose for all the new data domains is 6.27cm, while that for Meta-Pose is only 3.97cm

➔ a 36.68% reduction

Generalization to Different RF Technologies

Goal: a human activity recognition (HAR) system that works with many different RF technologies

- To reduce the cost and overcome the *barrier of wide deployment*
- To exploit *complementary* various RF technologies for robust systems



(a) Kicking sampled with FMCW Radar. (b) Kicking sampled with RFID. (c) Kicking sampled with WiFi.
(d) Running sampled with FMCW Radar. (e) Running sampled with RFID. (f) Running sampled with WiFi.

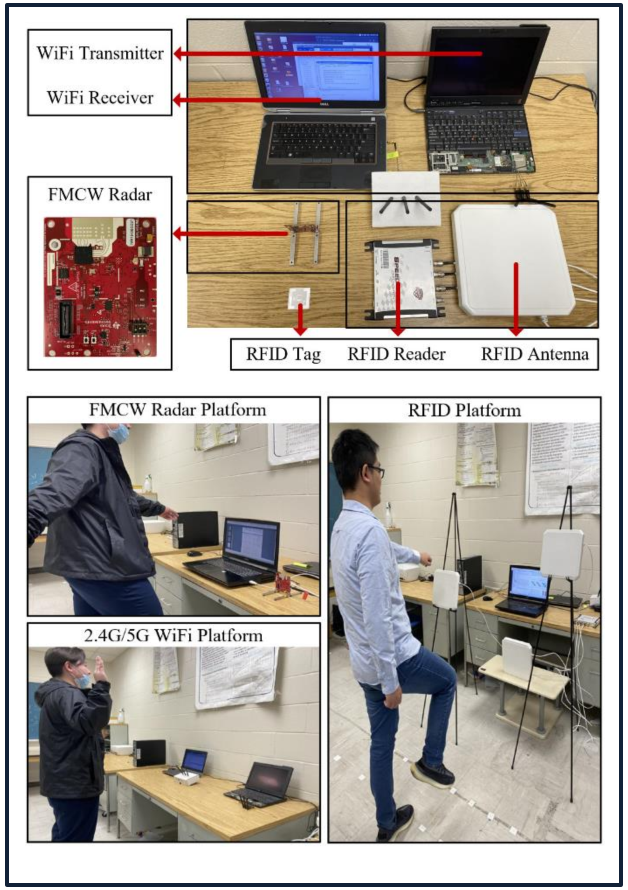
Raw data sampled by different RF technologies for the same activity over a 4-second period (FMCW Radar: range profile, RFID: phase, WiFi: Phase difference)

Challenges: With different RF platforms, the same human activity will be captured in very different forms of RF data: frequency bands, network protocols, device drivers, and hardware

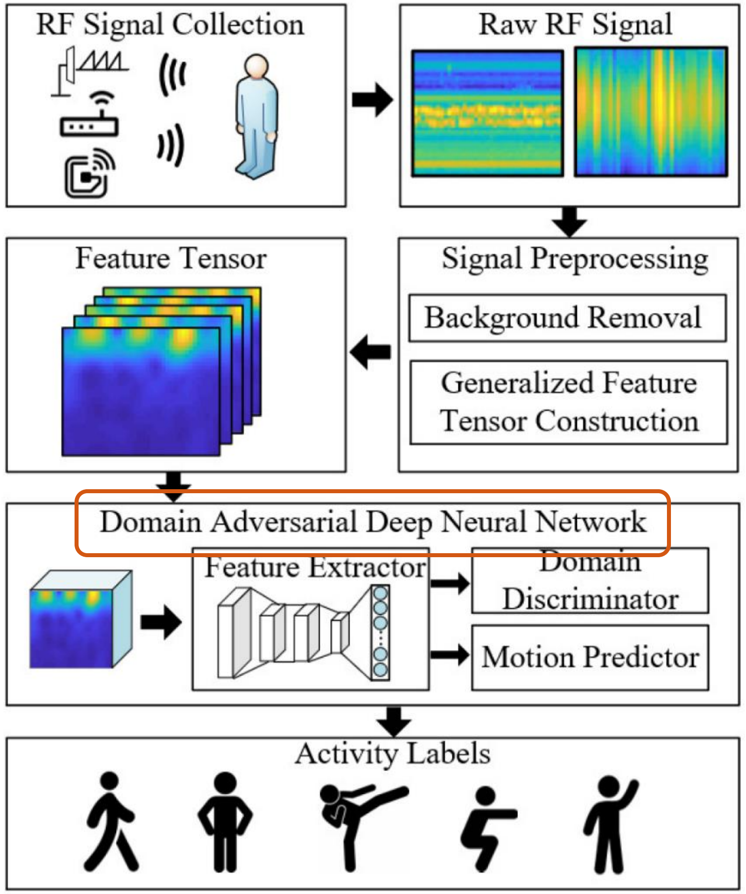
- Diversity in sampled data format
- Diversity in sensitivity
- Diversity in the translation of motion feature to RF data

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

TARF: Technology-agnostic RF HAR Solution



Human activity data sampling using different RF platforms

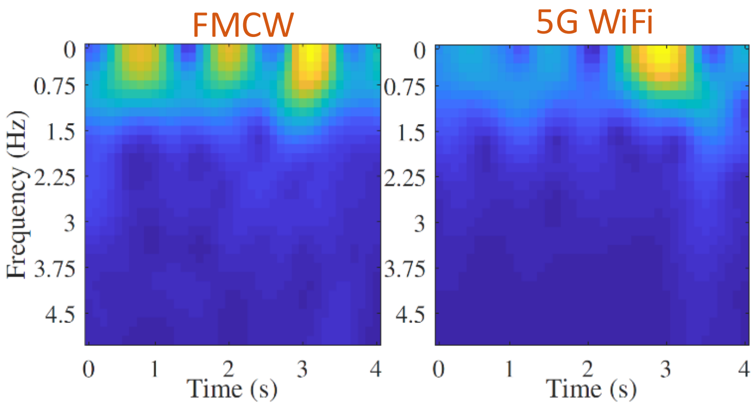


Architecture of TARF

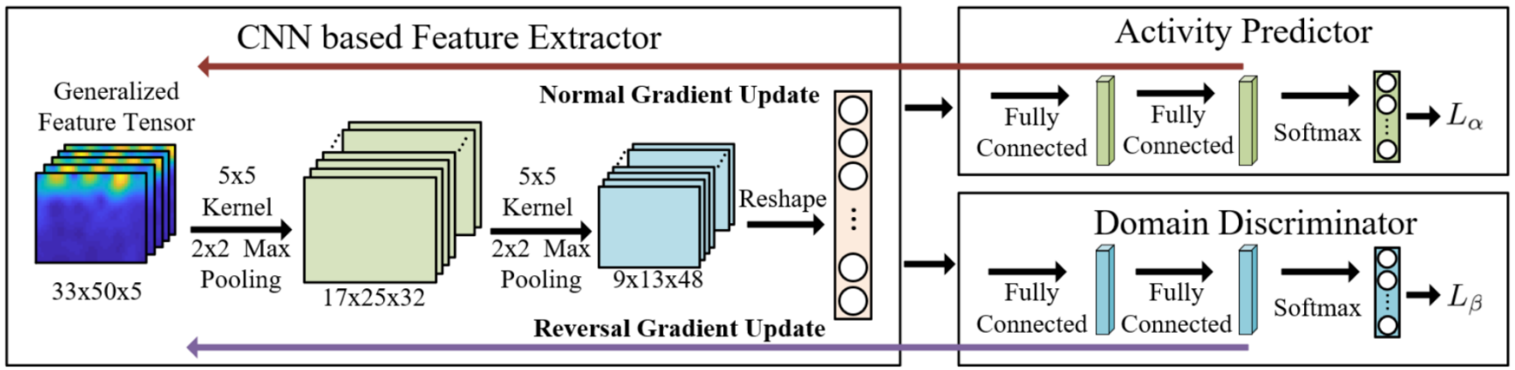
- RF Signal Collection
- Generalized RF Signal Preprocessing
  - Background removal
  - Generalized feature tensor construction
- Domain Adversarial Deep Neural Network (DANN) for Activity Recognition
  - CNN based feature extractor
  - Motion identifier
  - Domain discriminator

Activity Recognition with Domain Adversarial Neural Network

Challenge: motion feature translation



Examples of one slice of the generalized feature tensor for the kicking activity



Structure of the domain adversarial deep neural network used in the TARF system.

- Time-frequency domain transformation and tensorization
  - Short Time Fourier Transform
- Feature extraction with CNN
- Motion predictor
- Domain discriminator

- Loss of the activity predictor

$$L_\alpha = \frac{1}{N_b}\sum_{b=1}^{N_b}\sum_{k=1}^{N_a} \hat{y}_k^b \log\left(y_k^b\right)$$

N_a: Number of activity classes

- Loss of the domain discriminator

$$L_\beta = \frac{1}{N_b}\sum_{b=1}^{N_b}\sum_{q=1}^{N_d} \hat{y}_q^b \log\left(y_q^b\right)$$

N_d: Number of RF technologies

- Weight updates:

$$\hat{X_\gamma} = X_\gamma - \xi\left(\frac{\partial L_\alpha}{\partial X_\gamma} - \boxed{C_r}\frac{\partial L_\beta}{\partial X_\gamma}\right)$$

Combating rate

$$\hat{X_\alpha} = X_\alpha - \xi\frac{\partial L_\alpha}{\partial X_\alpha}$$

$$\hat{X_\beta} = X_\beta - \boxed{\xi} C_r\frac{\partial L_\alpha}{\partial X_\beta},$$

Learning rate

Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *J. Machine Learning Research*, vol. 17, no. 1, pp. 2096–2030, Apr. 2016.

Experiment Results

- Seven activities:
  - Standing still–ST, walking–WA, running-RU, squatting–SQ, body twisting–BT, kicking–KI, and hand waving–WH
- Baseline scheme: CNN (i.e., without the domain discriminator)

Accuracy: 90.86%

Output Class \ Target Class	ST	WA	RU	SQ	BT	KI	HW
ST	98.2%	0.3%	0.5%	0.7%	1.5%	0.1%	1.4%
WA	0.4%	93.2%	1.9%	1.8%	4.8%	1.3%	0.4%
RU	0.3%	3.6%	94.5%	1.1%	3.3%	1.8%	3.6%
SQ	0.0%	0.2%	1.7%	91.0%	2.2%	1.3%	2.0%
BT	0.1%	1.1%	0.6%	1.6%	83.5%	9.1%	1.4%
KI	0.7%	0.3%	0.5%	3.6%	4.3%	85.0%	0.7%
HW	0.2%	1.2%	0.2%	0.2%	0.5%	1.6%	90.4%

Accuracy: 91.00%

Output Class \ Target Class	ST	WA	RU	SQ	BT	KI	HW
ST	95.9%	0.3%	0.5%	0.6%	1.4%	0.1%	1.4%
WA	1.1%	94.3%	2.5%	1.6%	4.7%	1.2%	0.4%
RU	0.3%	2.5%	94.9%	0.9%	3.8%	1.7%	4.3%
SQ	0.0%	0.2%	0.8%	91.2%	1.4%	1.2%	1.0%
BT	1.8%	1.1%	0.6%	1.9%	84.0%	7.7%	1.4%
KI	0.7%	0.3%	0.5%	3.6%	4.2%	85.7%	0.7%
HW	0.2%	1.2%	0.2%	0.2%	0.5%	2.5%	90.8%

Confusion matrix of human activity recognition: FMCW Radar only
Left: CNN baseline; Right: TARF

Accuracy: 60.40%

Output Class \ Target Class	ST	WA	RU	SQ	BT	KI	HW
ST	83.1%	1.7%	1.6%	4.5%	5.8%	0.4%	7.3%
WA	1.6%	64.5%	5.6%	11.2%	19.0%	5.0%	2.2%
RU	5.1%	18.6%	83.9%	6.7%	13.2%	7.1%	18.0%
SQ	2.0%	1.2%	5.1%	42.5%	8.7%	5.0%	10.1%
BT	4.7%	5.8%	1.9%	10.4%	34.3%	36.6%	7.3%
KI	2.7%	1.7%	1.3%	23.1%	16.9%	39.5%	3.4%
HW	0.8%	6.4%	0.5%	1.5%	2.1%	6.3%	51.7%

Accuracy: 81.11%

Output Class \ Target Class	ST	WA	RU	SQ	BT	KI	HW
ST	89.1%	0.7%	1.4%	1.2%	2.7%	0.2%	2.9%
WA	3.0%	88.2%	6.3%	3.0%	8.8%	2.1%	0.9%
RU	0.8%	5.1%	87.0%	1.8%	7.1%	3.0%	8.6%
SQ	0.0%	0.5%	2.1%	83.1%	2.7%	2.1%	2.0%
BT	4.8%	2.3%	1.6%	3.6%	69.8%	14.1%	2.9%
KI	1.8%	0.7%	1.2%	7.0%	7.9%	73.9%	1.3%
HW	0.5%	2.5%	0.5%	0.4%	1.0%	4.5%	81.4%

Confusion matrix of human activity recognition: All four technologies
Left: CNN baseline; Right: TARF

ACCURACY COMPARISON WITH DIFFERENT TESTING SCENARIOS

Testing Environment	*WiFi 5GHz*	*WiFi 2.4GHz*	*FMCW*	*RFID*	*CNN Baseline*	*TARF*
LOS	91.86%	89.37%	91.22%	90.73%	63.41%	82.73%
NLOS	90.76%	88.71%	81.77%	74.22%	61.29%	*81.24%*
Dynamic Environment	75.05%	71.44%	79.29%	89.38%	62.54%	*80.18%*

Outline

- Human pose tracking: preliminaries and approaches
- RFID-Pose: 3D human pose monitoring using RFID [1], and its extensions [2,3]
- **Generative AI for data augmentation** [4-9]
- Generative AI for 3D pose augmentation and completion [10,11]
- Conclusions

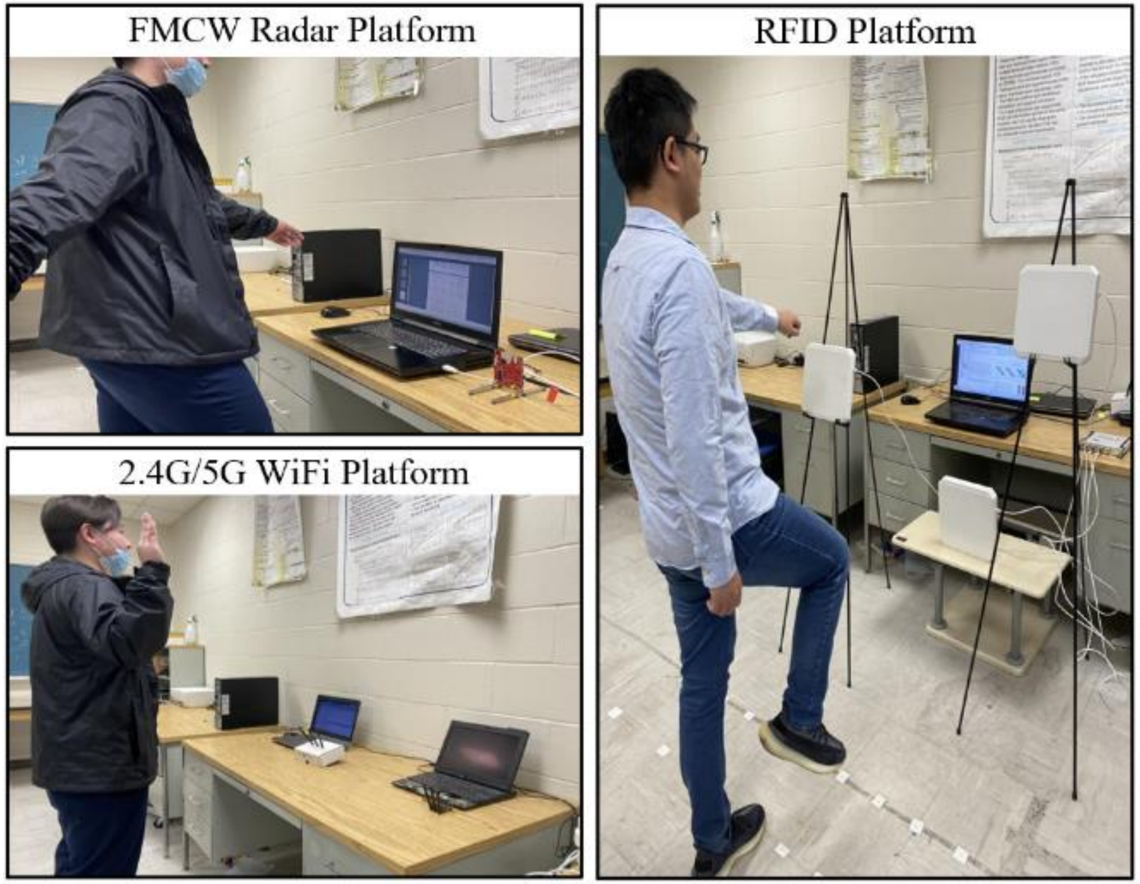
[1] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, vol.70, no.3, pp.1218-1231, Sept. 2021.
[2] C. Yang, L. Wang, X. Wang, and S. Mao, "Environment adaptive RFID based 3D human pose tracking with a meta-learning approach," *IEEE Journal of Radio Frequency Identification*, to appear. DOI: 10.1109/JRFID.2022.3140256.
[3] C. Yang, X. Wang, and S. Mao, "TARF: Technology-agnostic RF sensing for human activity recognition," *IEEE Journal of Biomedical and Health Informatics*, vol.27, no.2, pp.636--647, Feb. 2023.
[4] Z. Wang, C. Yang, and S. Mao, "Data augmentation for RFID-based 3D human pose tracking," in *Proc. IEEE VTC-Fall 2022*, London, UK, Sept. 2022.
[5] C. Yang, Z. Wang, and S. Mao, "RFPose-GAN: Data augmentation for RFID based 3D human pose tracking," in *Proc. The 12th IEEE International Conference on RFID Technology and Applications* (*IEEE RFID-TA 2022*), Cagliari, Italy, Sept. 2022, pp.138-141.
[6] Z. Wang and S. Mao, "AIGC for RF sensing: The case of RFID-based human activity recognition," in *Proc. ICNC 2024*, Big Island, HI, Feb. 2024, pp.1092-1097.
[7] Z. Wang and S. Mao, "AIGC for wireless data: The case of RFID-based human activity recognition," in *Proc. IEEE ICC 2024*, Denver, CO, June 2024, pp. 1–6.
[8] Z. Wang, C. Yang, and S. Mao, "AIGC for RF-based human activity sensing," *IEEE Internet of Things Journal*, vol.12, no.4, pp.3991-4005, Feb. 2025.
[9] Z. Wang and S. Mao, "AIGC for Wireless Sensing: Diffusion-empowered Human Activity Recognition," *IEEE Transactions on Cognitive Communications and Networking*, vol.11, no.2, pp.657-671, Apr. 2025
[10] Z. Wang and S. Mao, "Generative AI for 3D human pose completion under RFID sensing constraints," in *Proc. ICNC 2025*, Honolulu, HI, Feb. 2025, pp.485-490.
[11] Z. Wang and S. Mao, "Generative AI-empowered RFID sensing for 3D human pose augmentation and completion," *IEEE Open Journal of the Communications Society*, vol.6, pp.2958-2975, Feb. 2025.

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Data Collection in Learning-based RF Sensing

Training data collection is challenging:

- RF sensing data collection is time-consuming
  - Hours of data
  - Camera and RF data should be synchronized
- Diversity of training subjects
- Diversity in the RF signal representations from different RF devices



AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Solution: Data Augmentation

- **Data Augmentation**: techniques used to increase the amount of data by adding slightly modified copies of the existing data or newly created synthetic data from existing data
- Images: resize, crop, rotate, flip, etc.

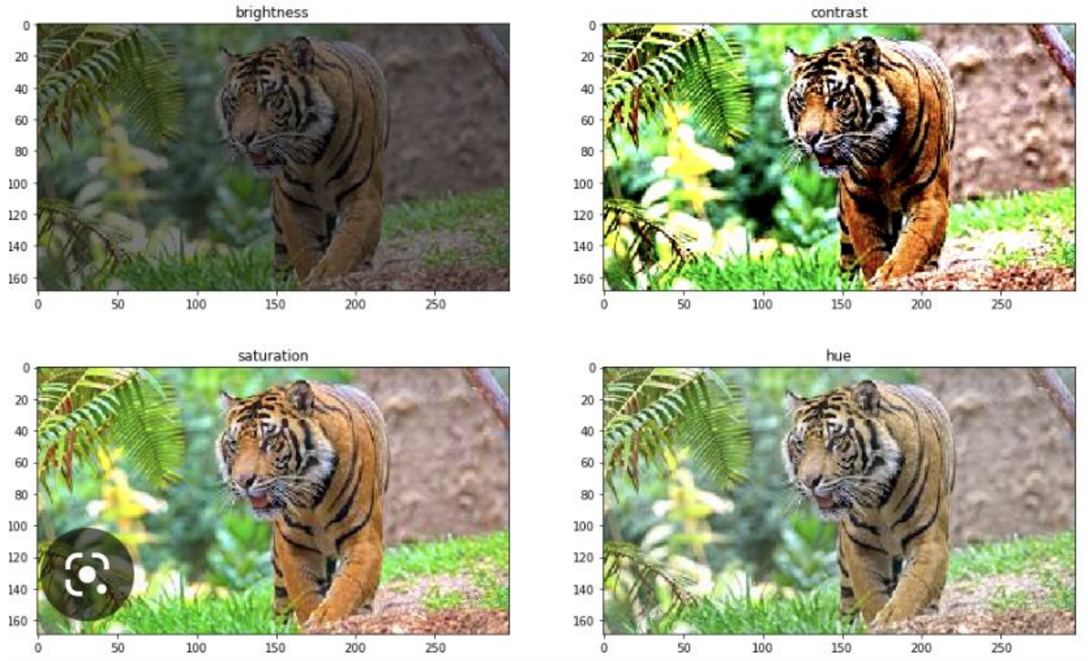


Image source: https://www.v7labs.com/blog/data-augmentation-guide
https://en.wikipedia.org/wiki/Data; https://www.simplilearn.com/dat

Augmentation of RF data:

- To greatly reduce the data collection efforts
- RF data: random and hard to manipulate
  - A more challenging problem

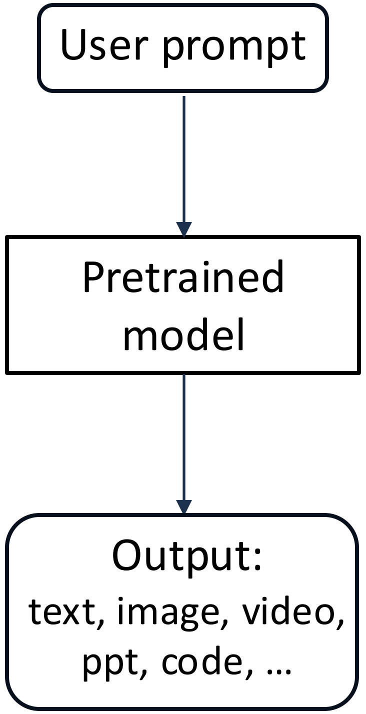
Observation: Pose, on the other hand, can be more easily manipulated in term of movement variations, body forms, camera angles, and locations

Question: how to map the 3D human pose data to RF features?

By enhancing the diversity of pose data, we can, in turn, augment RF data by transforming the augmented pose data into high quality RF data

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

AIGC: GAN vs. Diffusion



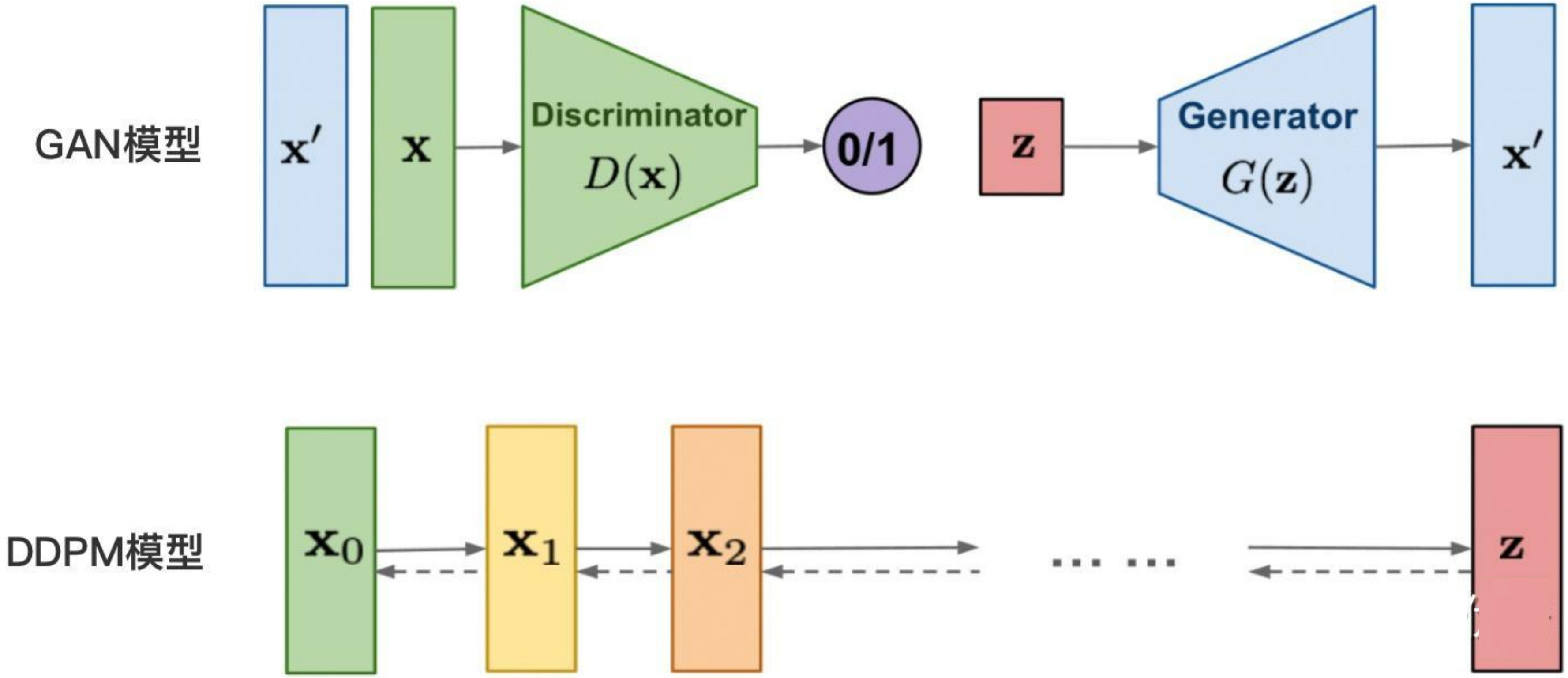
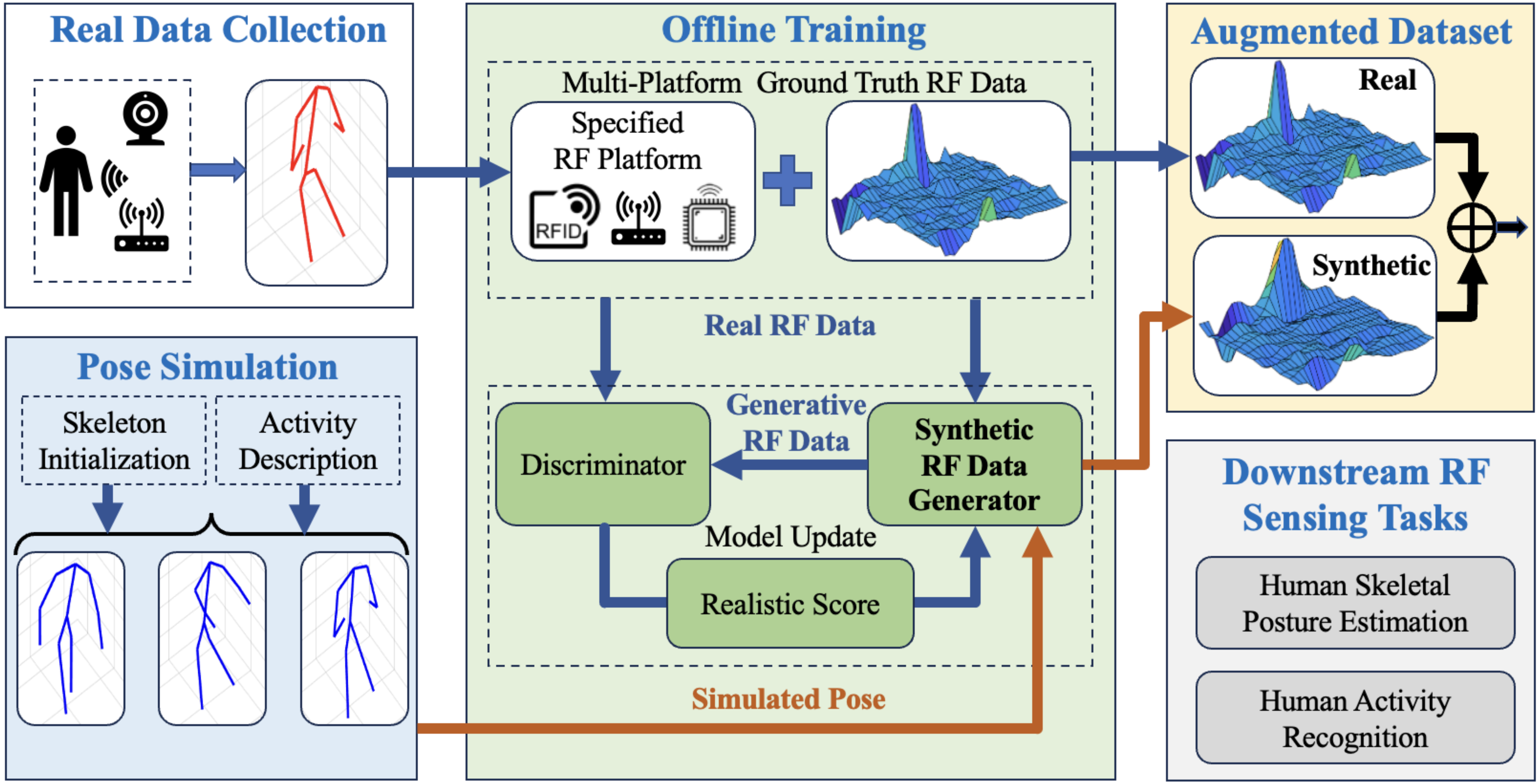


Image source: https://zhuanlan.zhihu.com/p/590840909

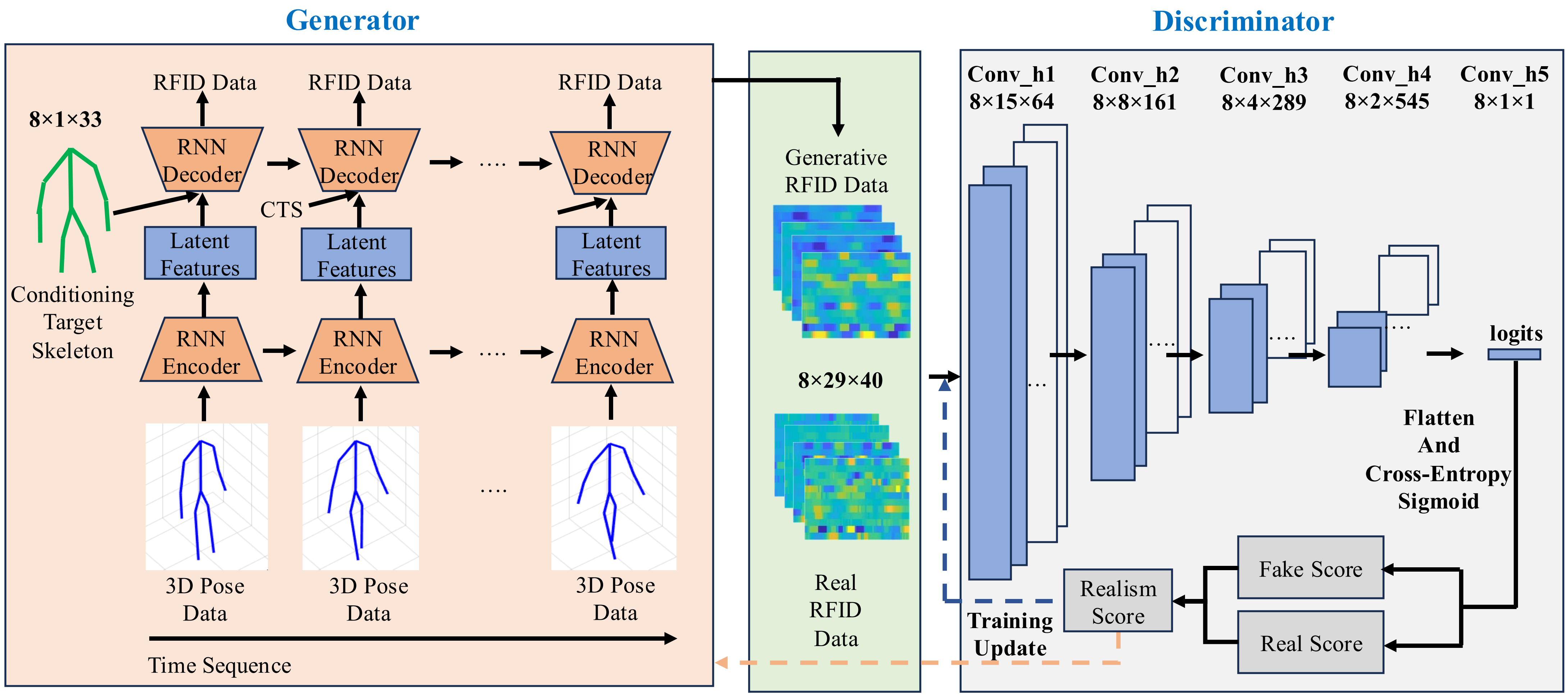
J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *arXiv:2006.11239*, June 2020. [Online] Available: https://arxiv.org/abs/2006.11239

Proposed Solution: Data Augmentation with R-GAN



AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Recurrent Generative Adversarial Network (R-GAN)



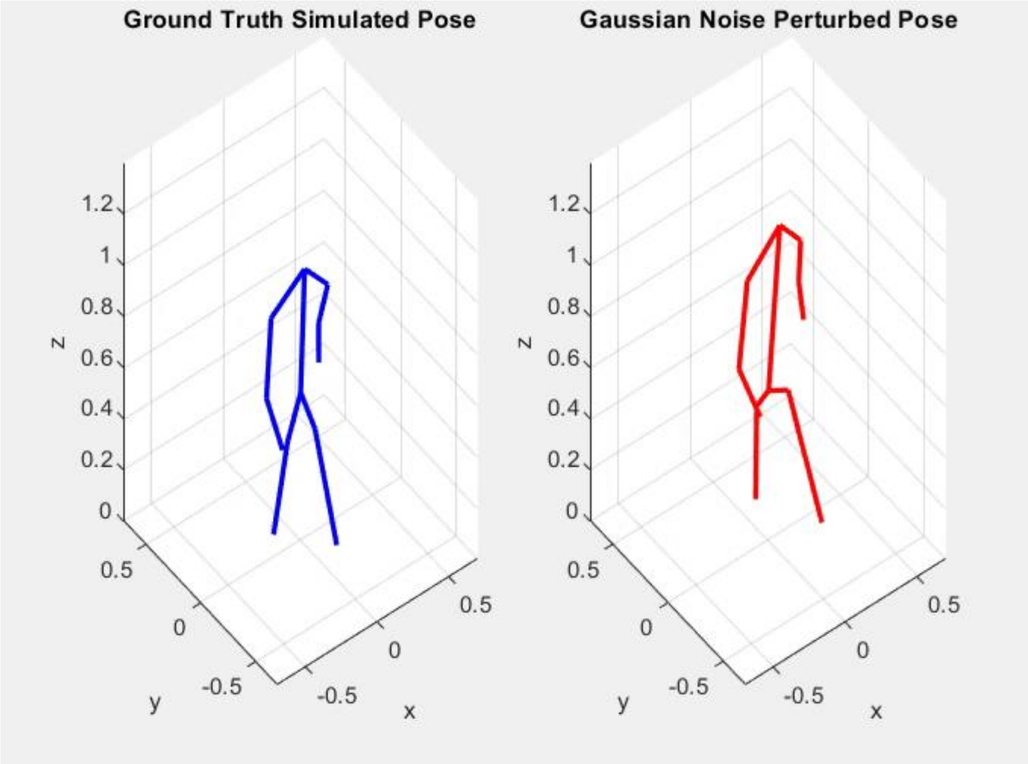
An RNN Autoencoder serves as the Generator of the GAN, and a 1D CNN serves as the Discriminator

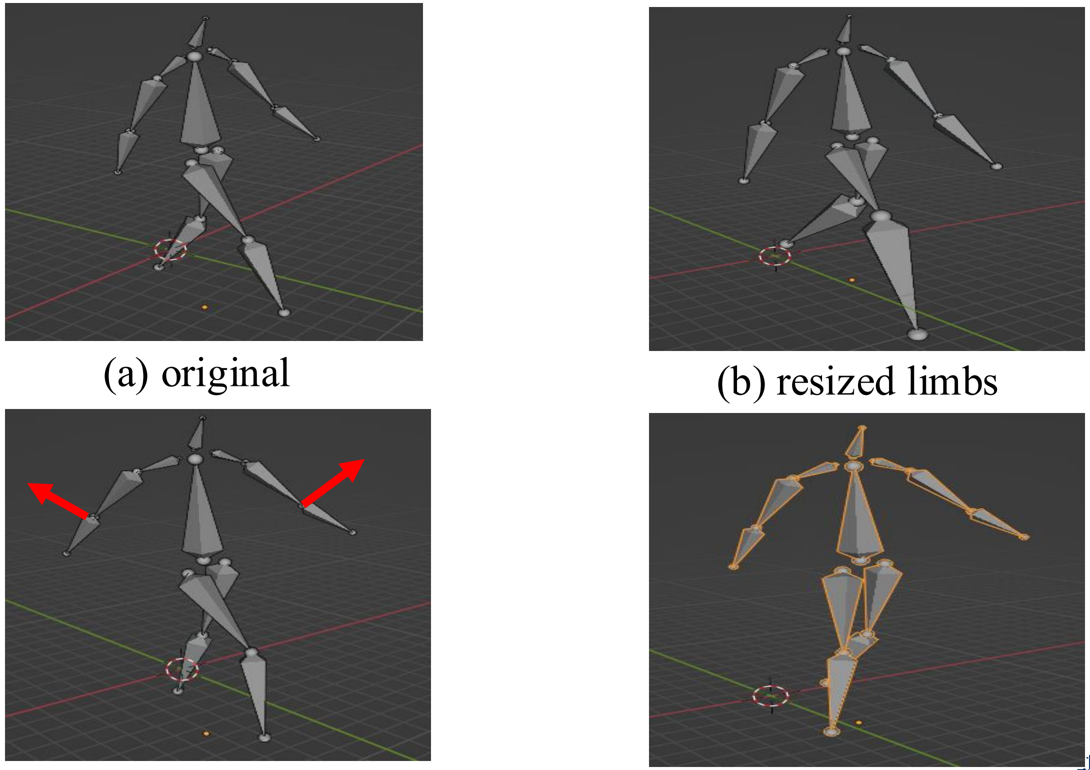
The final layer of the discriminator is a 1D CNN layer with 1 kernel for dimensionality reduction, to be flattened to a logits vector for computing a realistic score

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Simulated Human Pose Data

- Training data collection: performing activities in front of both Kinect camera and RF platforms
- Pose data generated using a simulation tool *Blender* [1]
- Two ways to enhance diversity: (i) TGNP: introduce *independent Gaussian noise* to the joints (0-mean, small variance); (ii) PoseMod: introduce variations in poses movements, skeletons, and camera viewpoints and locations [2]





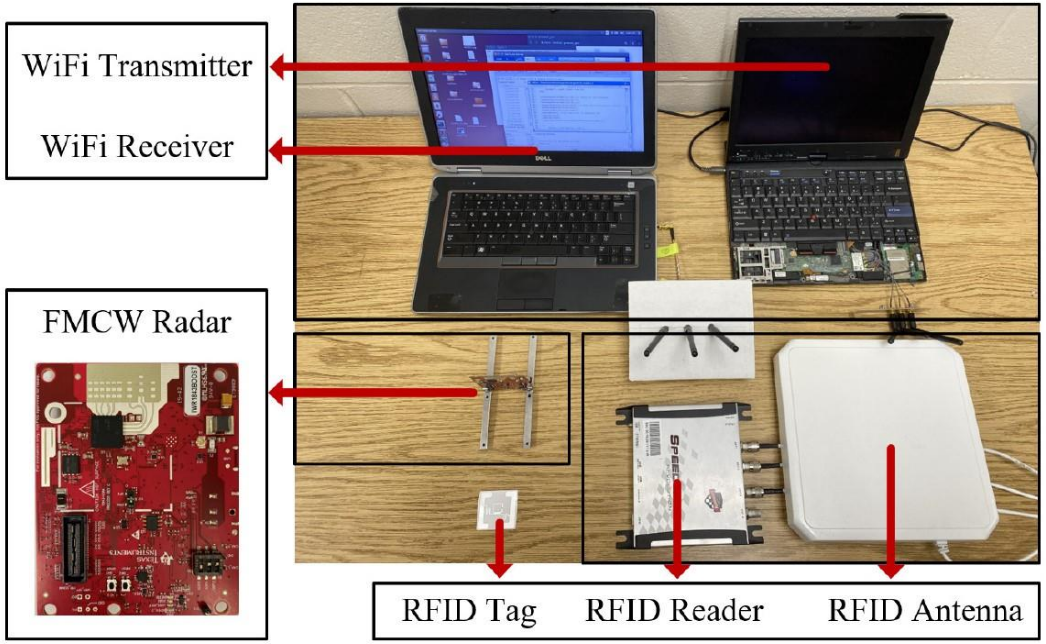
(a) original (b) resized limbs

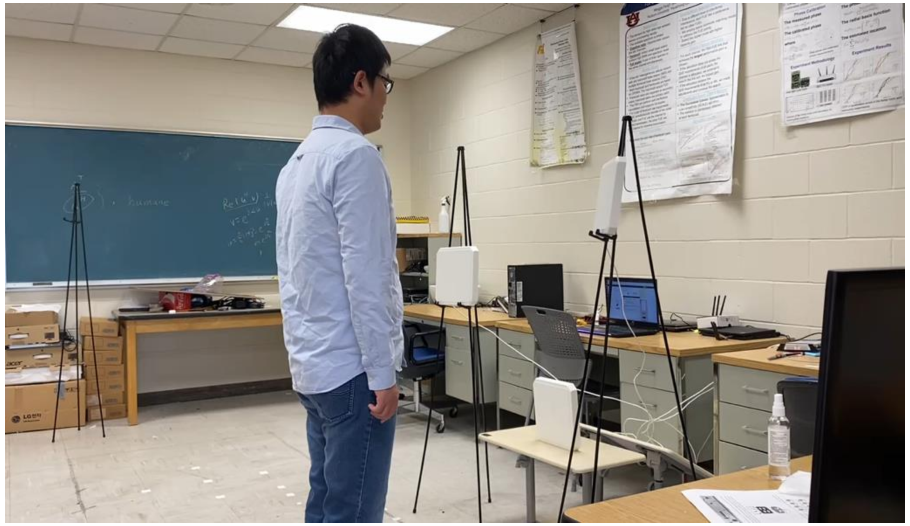
(c) Extended movement variations (d) Different locations and viewpoints

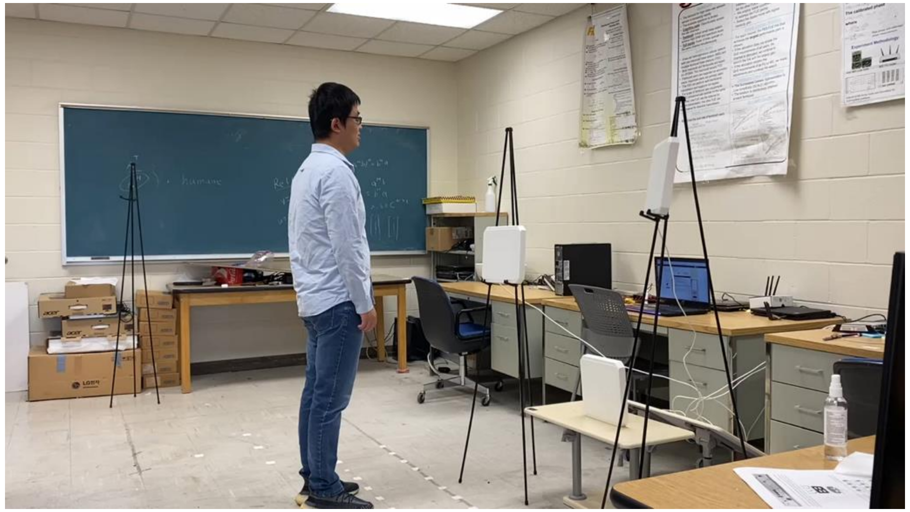
[1] Blender - a 3D modelling and rendering package: http://www.blender.org

[2] K. Gong, J. Zhang, and J. Feng, "PoseAug: A Differentiable Pose Augmentation Framework for 3D Human Pose Estimation," in *Proc. IEEE/CVF CVPR'21*, Virtual Conference, Sept. 2021

Implementation and Evaluation







RFID: an off-the-shelf Impinj R420 reader, passive ALN-9634 (HIGG-3) tags, and three S9028PCR polarized antennas

mmWave Radar: IWR1843 BOOST single-chip FMCW sensor

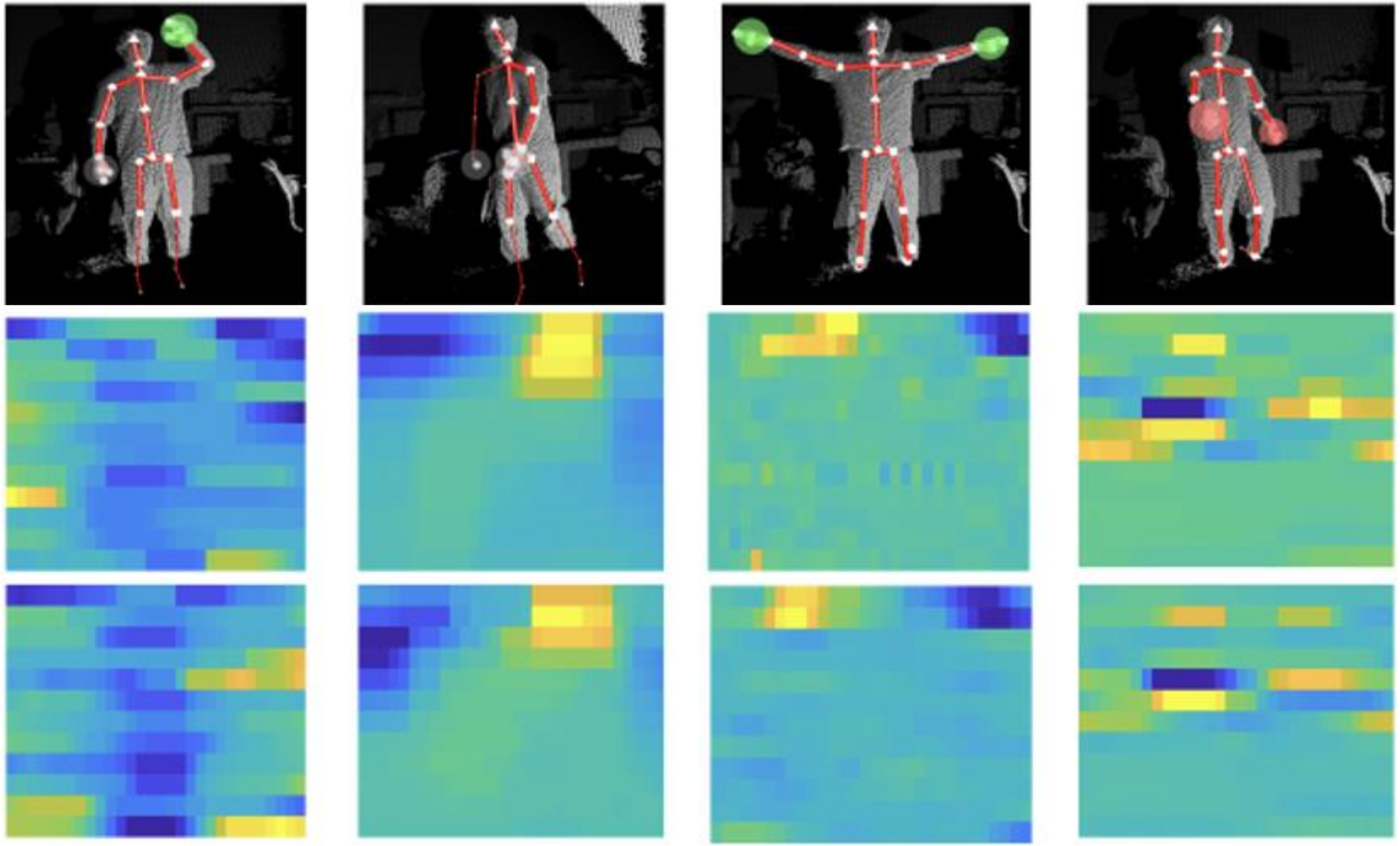
WiFi: 5300 network interface card (NIC): 2.4 GHz or 5 GHz

Training with a GTX 1660 Ti Graphics card

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Example of Synthesized RF Data

	2.4 GHz WiFi (CSI): Drinking	5 GHz WiFi (CSI): Twisting	RFID (Phase): Waving arms	mmWave (Range Profile): Walking
Kinect				
Ground truth				
Synthesized				



Quality of Synthesized RF Data

Structural Similarity Index (SSIM)

$$SSIM(x,x') \triangleq \frac{(2\mu_x\mu_{x'} + C_1)(2\sigma_{xx'} + C_2)}{(\mu_x^2 + \mu_{x'}^2 + C_1)(\sigma_x^2 + \sigma_{x'}^2 + C_2)}$$

luminance, contrast, and structure

Frechet Inception Distance (FID)

$$\mathbf{FID} = \|\mu - \mu'\|_2^2 + \mathrm{Tr}(\Sigma + \Sigma' - 2\sqrt{\Sigma \times \Sigma'})$$

$$\text{Diversity} = \frac{1}{S_{div}} \sum_{i=1}^{S_{div}} \|f_i - f_i'\|_2$$

$$\text{Multimodality} = \frac{1}{Z \times S_{mul}} \sum_{z=1}^{Z} \sum_{i=1}^{S_{mul}} \left\|f_{z,i} - f_{z,i}'\right\|_2$$

Table I
SSIM Scores Achieved by RF-AIGC for the Four RF Platforms

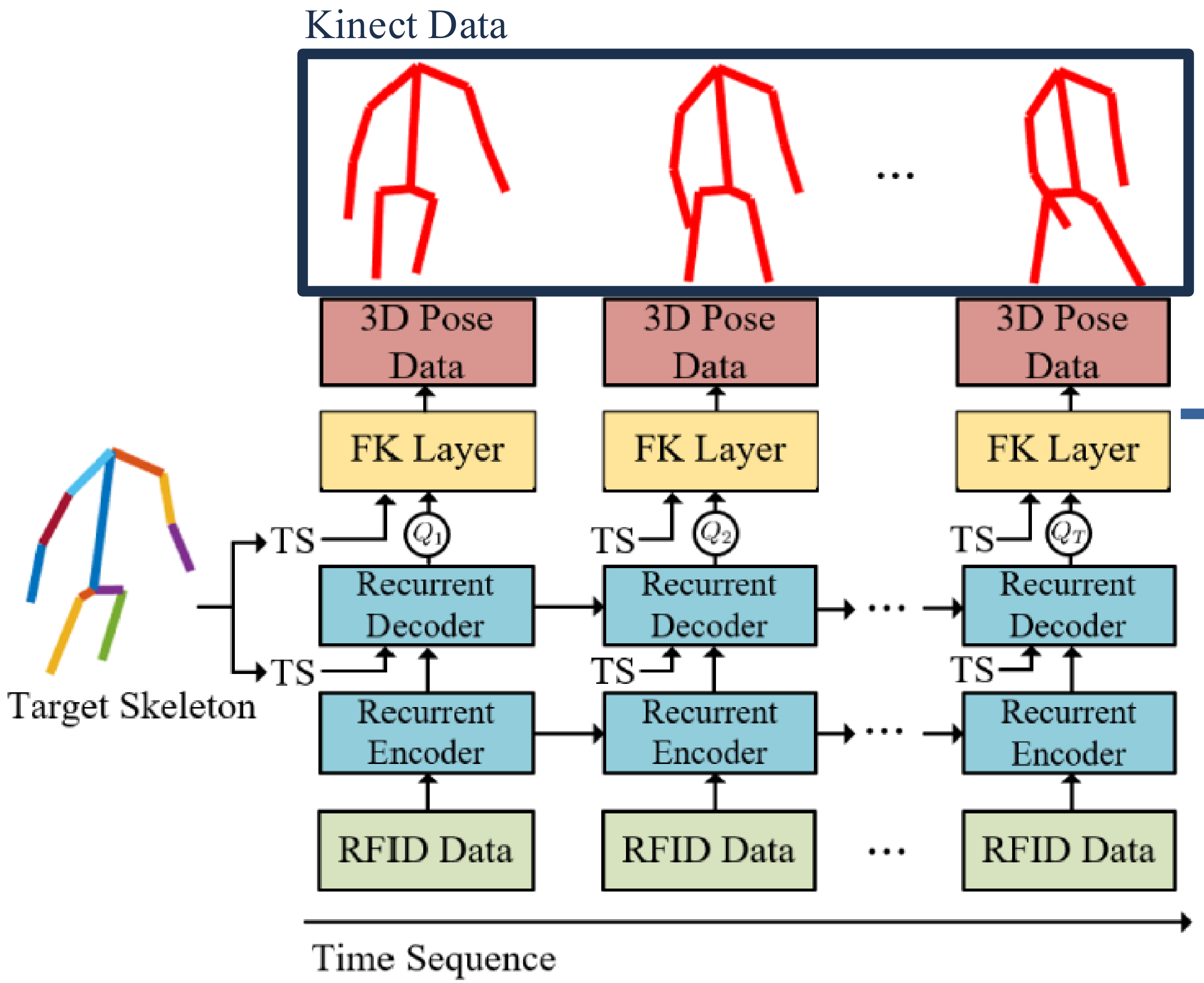
RF Platforms	SSIM Score ↓	SSIM Structure Score ↓
RFID	0.8995	0.9310
5G WiFi	0.8363	0.8675
FMCW Radar	0.8282	0.8563
2.4G WiFi	0.7473	0.7718

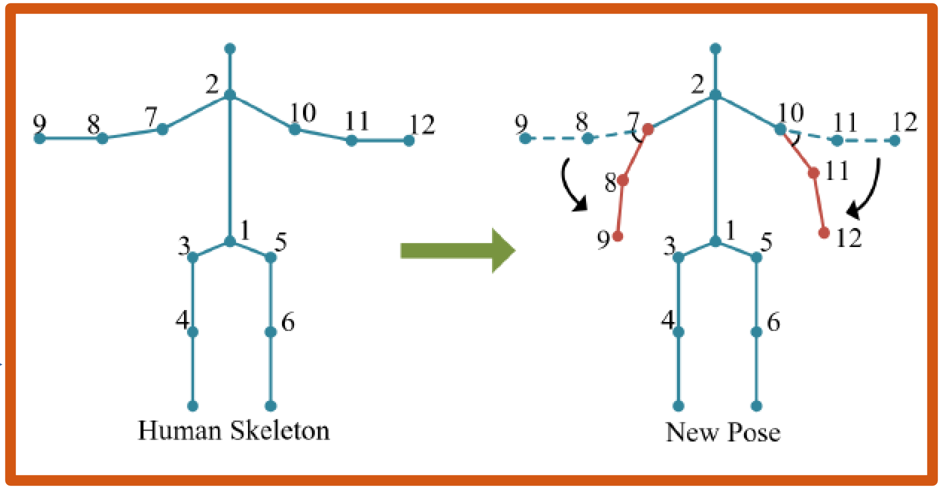
Table II
Comparison of FID, Diversity, and Mutlimodality Scores for Generated and Real RF Data

	FID ↓	Diversity ↓	Multimodality ↓
PoseMod Synth.	$58.128^{\pm 0.103}$	$10.843^{\pm 0.266}$	$9.008^{\pm 0.317}$
TGNP Synth.	$50.500^{\pm 0.091}$	$9.594^{\pm 0.287}$	$8.058^{\pm 0.414}$
Sufficient Real	$6.216^{\pm 0.025}$	$9.329^{\pm 0.230}$	$8.392^{\pm 0.391}$
Limited Real	$4.548^{\pm 0.008}$	$8.584^{\pm 0.243}$	$7.353^{\pm 0.409}$

Downstream Task I: 3D Pose Tracking

The Recurrent Autoencoder based Deep Kinematic Neural Network Model



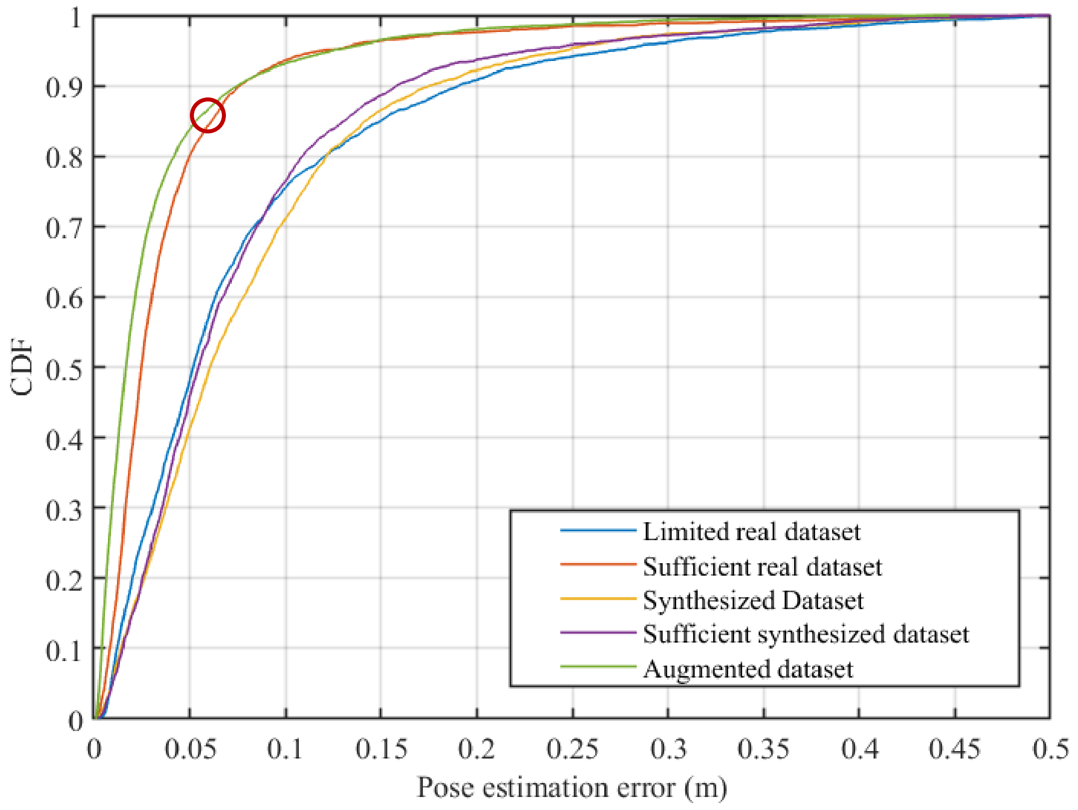


- Recurrent Autoencoder (256 gated recurrent units (GRU)):
  RF data ➔ unit quaternion
- Forward kinematic layer:
  Rotation matrix ➔ 3D pose
- Kinect data: labels, for training and performance evaluation

Downstream Task: Pose Estimation (RFID)

Dataset	Amounts	Composition	Mean Error
Limited-Real	17.6 min	Limited real data	8.06 cm
Sufficient-Real	105.6 min	Sufficient real data	**3.54 cm**
Synth	316.8 min	3 batches of synthesized data	7.24 cm
Sufficient-Synth	422.4 min	4 batches of synthesized data	7.03 cm
Aug	440 min	Sufficient Augmented Dataset (limited real data + sufficient synthesized data)	**2.97 cm**

Test Dataset	3 datasets per activity of 3 subjects (26.4 min) for each platform (RFID, 2.4GHz/5GHz WiFi, mmWave radar)

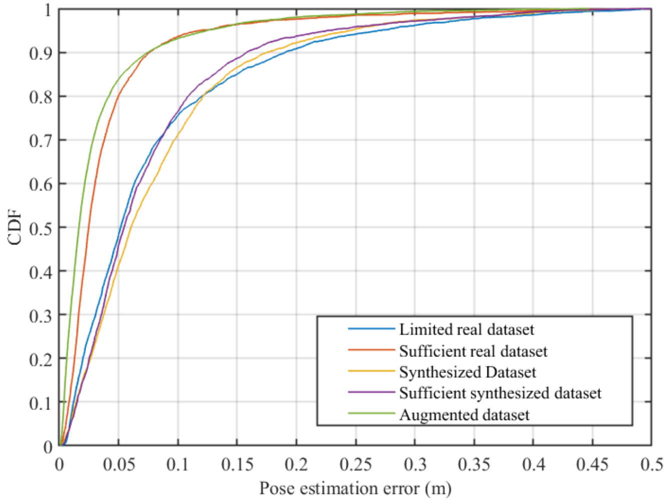


CDF curves for estimation errors of 5 models trained with the 5 different datasets, respectively

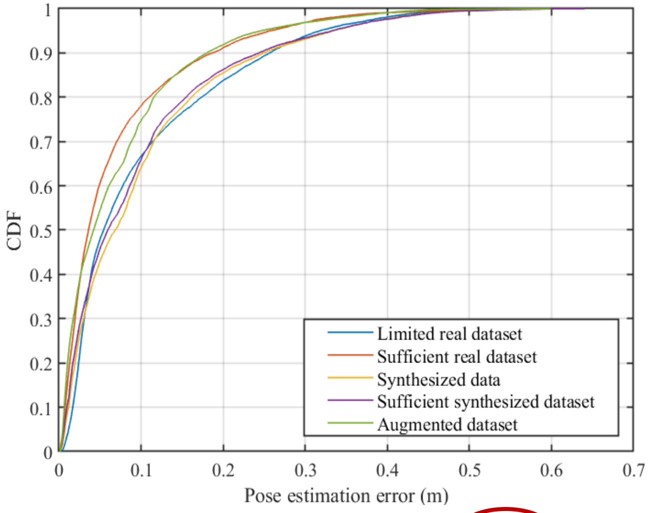
AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Pose Estimation Error – Four Technologies

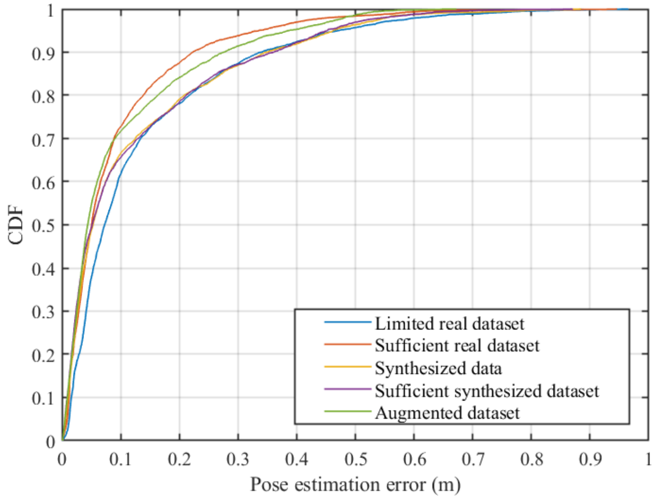
- RFID achieves the best performance with augmented data
- 5G WiFi has an adequate performance, while 2.4G WiFi and FMCW platforms has the poorest performance among the four platforms
- Nevertheless, data augmentation boosts the pose estimation performance to a level that is on par or better than the case with sufficient real data



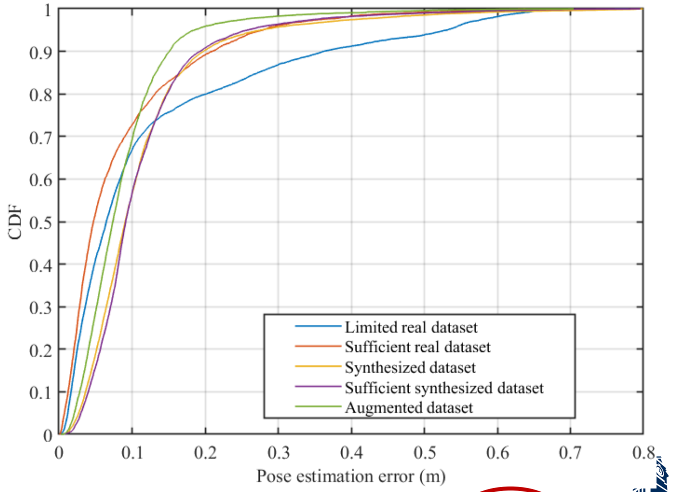
(a) CDF of all estimated joints by RFID-Pose



(b) CDF of all estimated joints by CSI5G-Pose



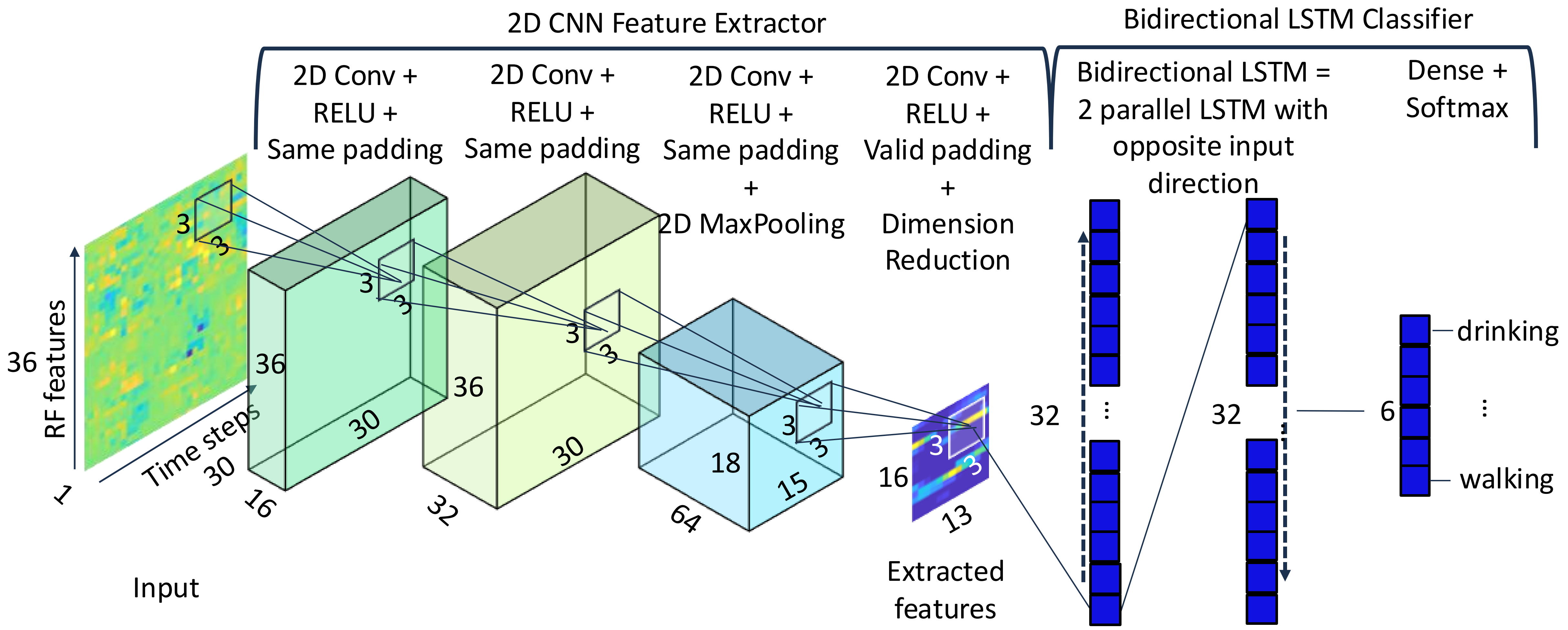
(c) CDF of all estimated joints by CSI2.4G-Pose



(d) CDF of all estimated joints by FMCW-Pose

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

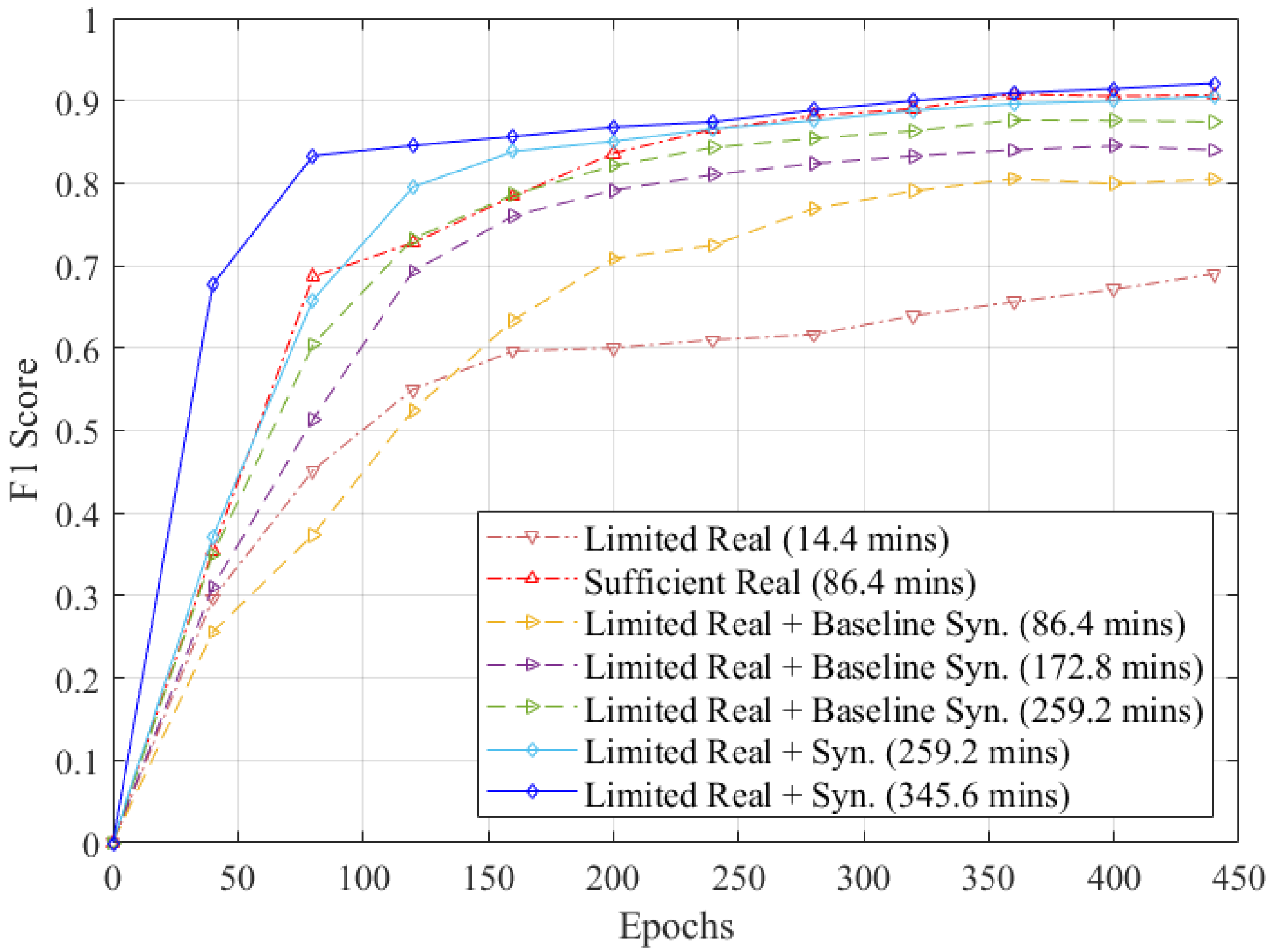
Downstream Task: Human Activity Recognition



AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

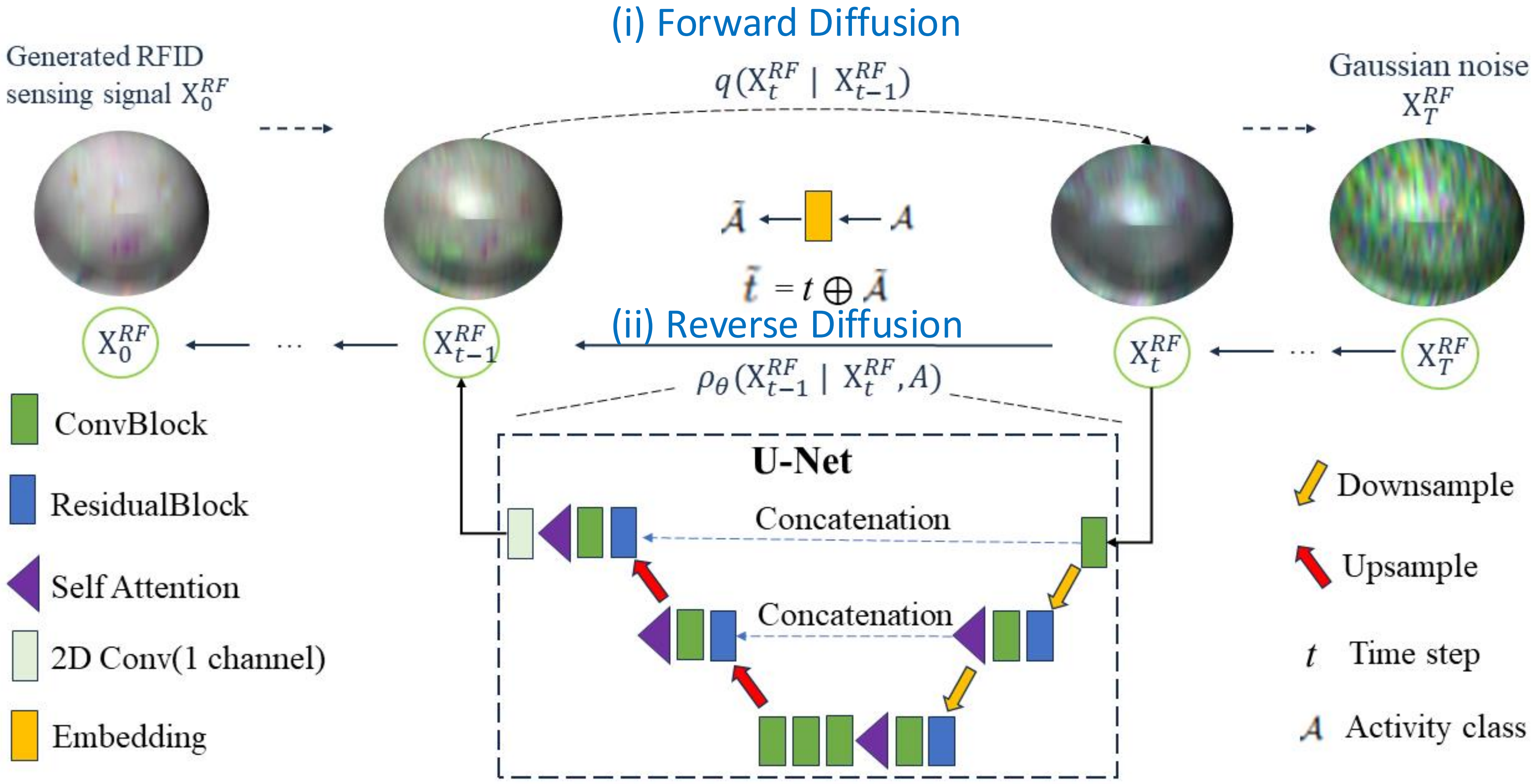
Improvements through Augmentation (RFID)

- Positive correlation between the amount of synthesized data and model performance
- Gaussian noise approach:
  - F1 score of 87%
  - Lower than the sufficient real model
- Pose perturbation approach:
  - F1 score of 92.09%
  - Outperform the case with sufficient real data
  - Costs around almost 4 times the amount of real data
- Diversity and amount both improved by data augmentation



AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

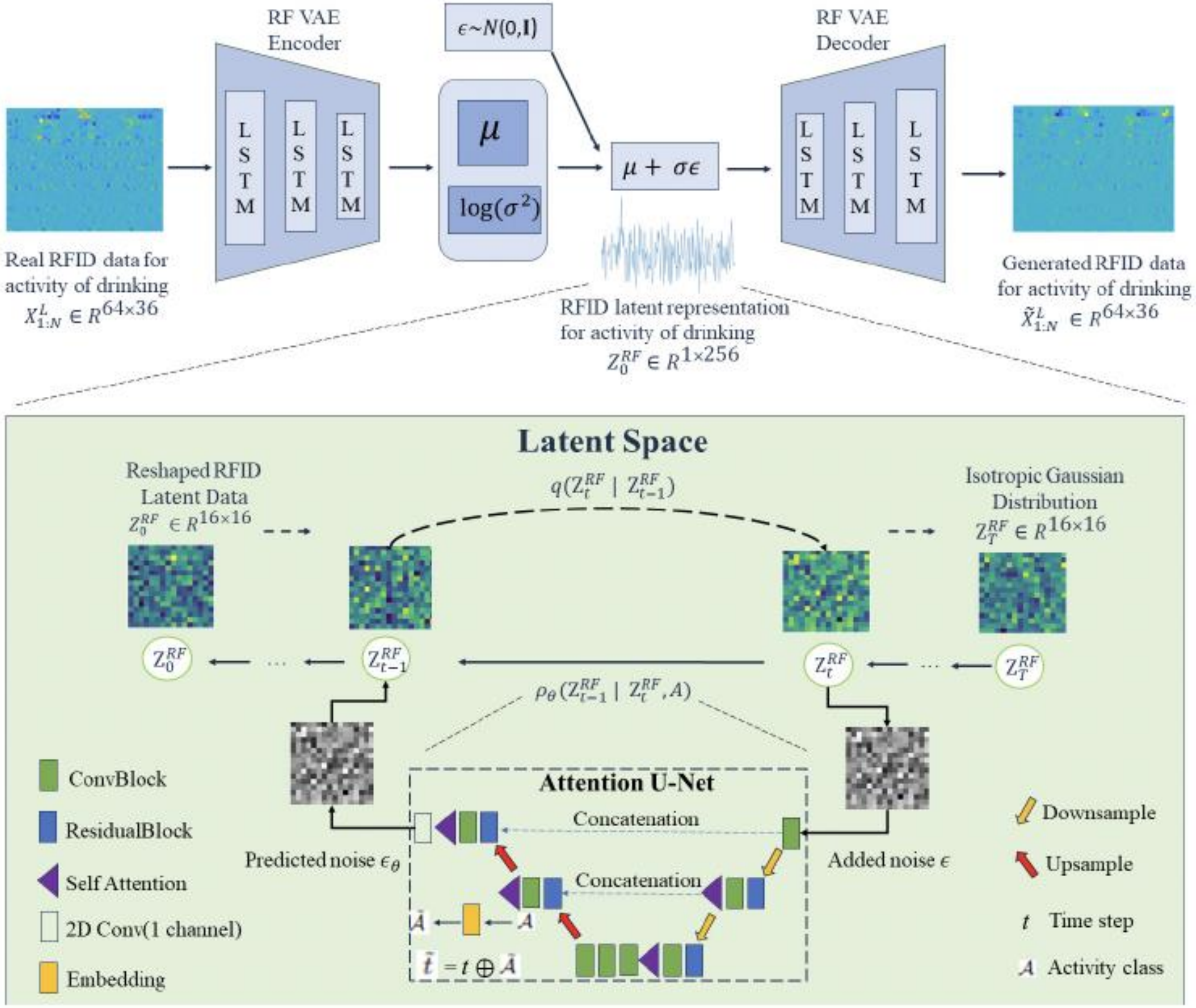
Class Conditional Diffusion for Generating RF Data



Class conditioning enabled through: Activity Class *A* is first embedded through MLP layers, then incorporated into U-Net through simple concatenation with time step *t*

Stable Diffusion-based Approach

- Diffusion on the latent representations of raw RF data
- The procedure of conditional RF data generation with RFIDACCLDM
  - The reverse process *p* progressively transforms random Gaussian noises into plausible time series data, conditioned on embedded class labels
  - The structure of the denoiser, the U-Net model, is also illustrated

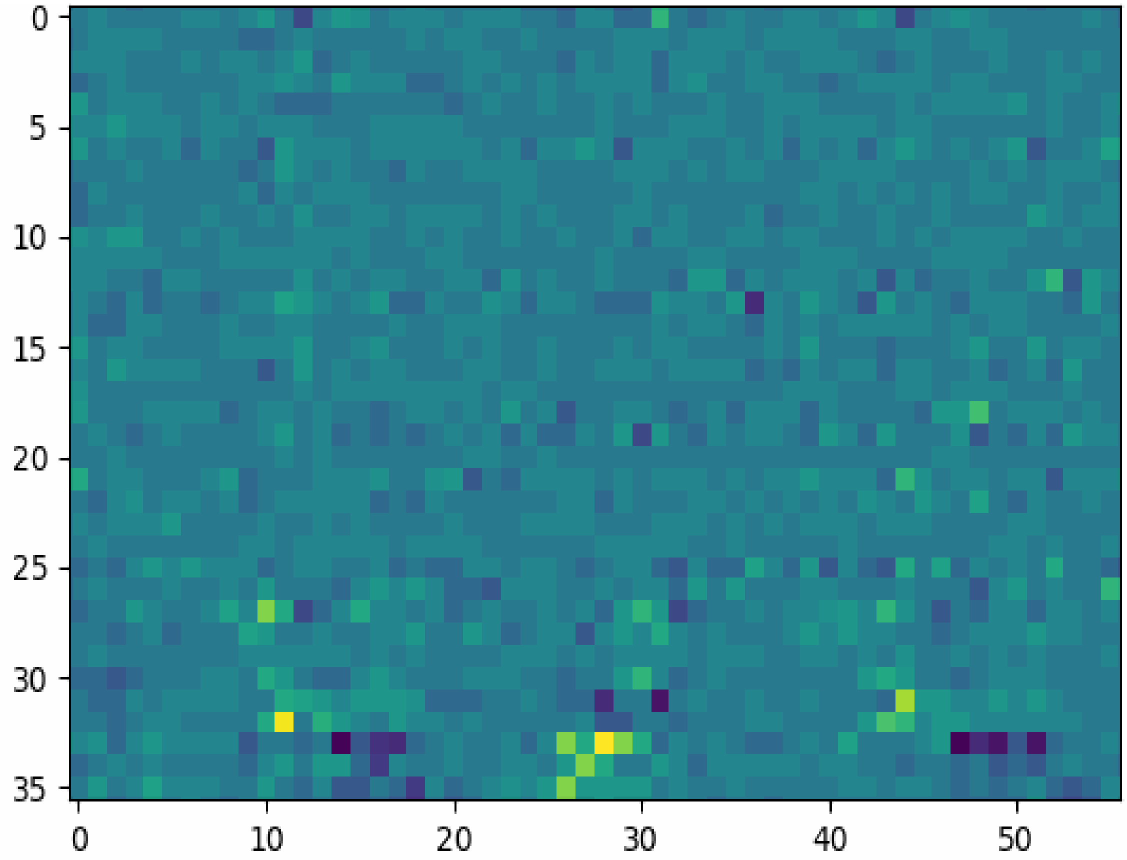


AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Diffusion Examples

Reverse Diffusion (Generation)

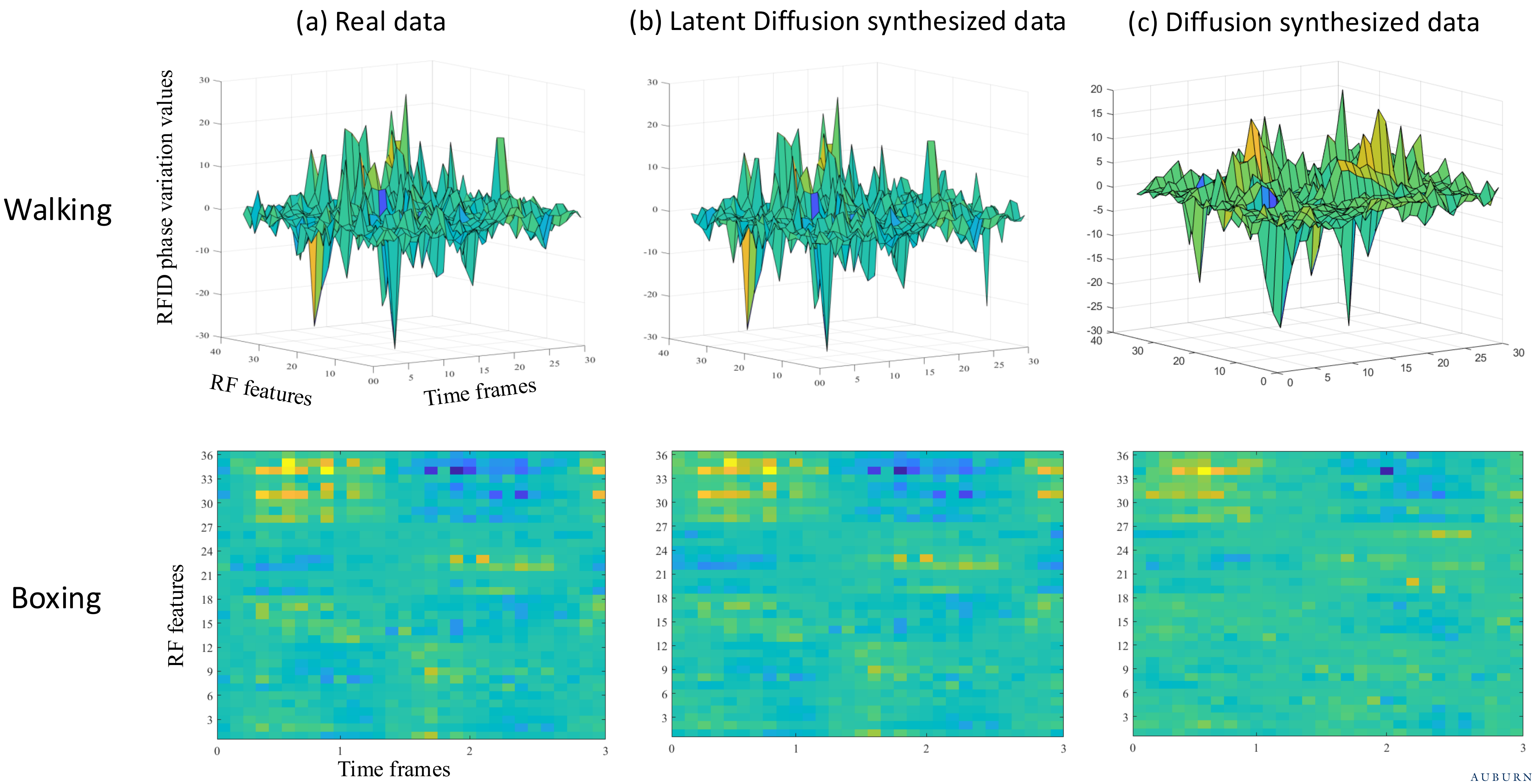
Reverse Diffusion (Generation): RFID data



The starting point (appearing as random Gaussian noise) is the final step of the forward diffusion process which progressively adds Gaussian noise to eventually result in an Isotropic Gaussian distribution

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Fidelity of Diffusion Generated RF Data



AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Fidelity of Diffusion Generated RF Data (cont'd)

COMPARISON OF DIVERSITY SCORES

Model	Diversity score
RFPose-GAN [23]	$9.48^{\pm 0.25}$
RFID-ACCDM	$11.10^{\pm 0.21}$
RF-ACCLDM	$9.16^{\pm 0.31}$
Real	$9.33^{\pm 0.25}$

- RFID-ACCDM (Activity Class Conditional Diffusion Model)
- RF-ACCLDM (Activity Class Conditional Latent Diffusion Model)

OUR LATENT DIFFUSION GENERATED SAMPLE QUALITY COMPARISON IN FID WITH PLAIN DIFFUSION MODEL, AUTOENCODER-BASED RFPOSE-GAN MODELS, AND REAL DATA FOR SELECTED HUMAN ACTIVITIES AND ALL ACTIVITIES.

Model	Standing	Waving	Walking	Boxing	Overall
RFPose-GAN	36.18	33.01	44.97	69.56	48.89
RFID-ACCDM	8.79	8.25	20.68	40.54	25.64
RF-ACCLDM	4.56	7.01	3.64	4.84	10.45
Real	5.17	7.36	4.78	4.49	6.22

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Downstream Task: Human Activity Recognition

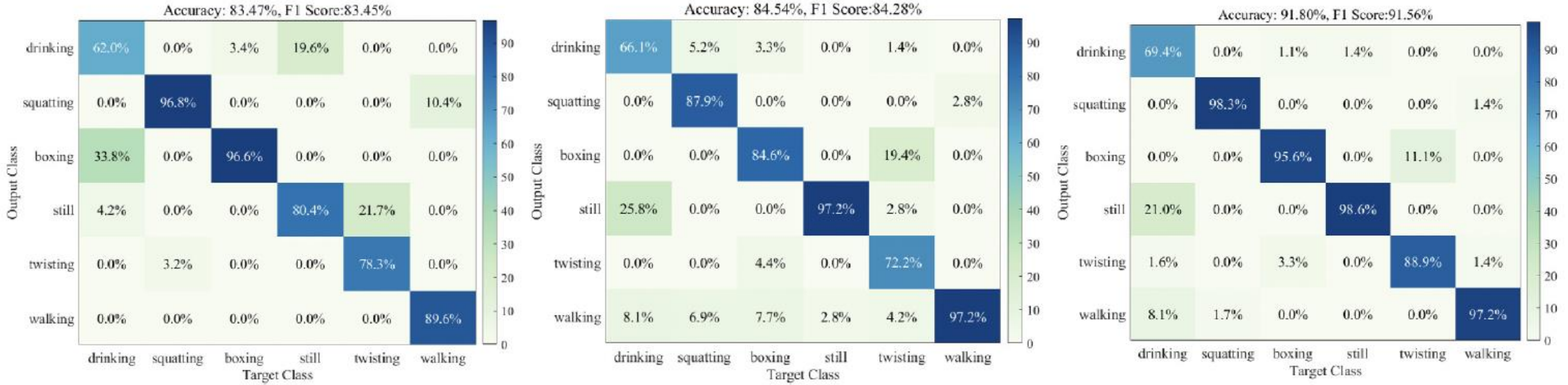
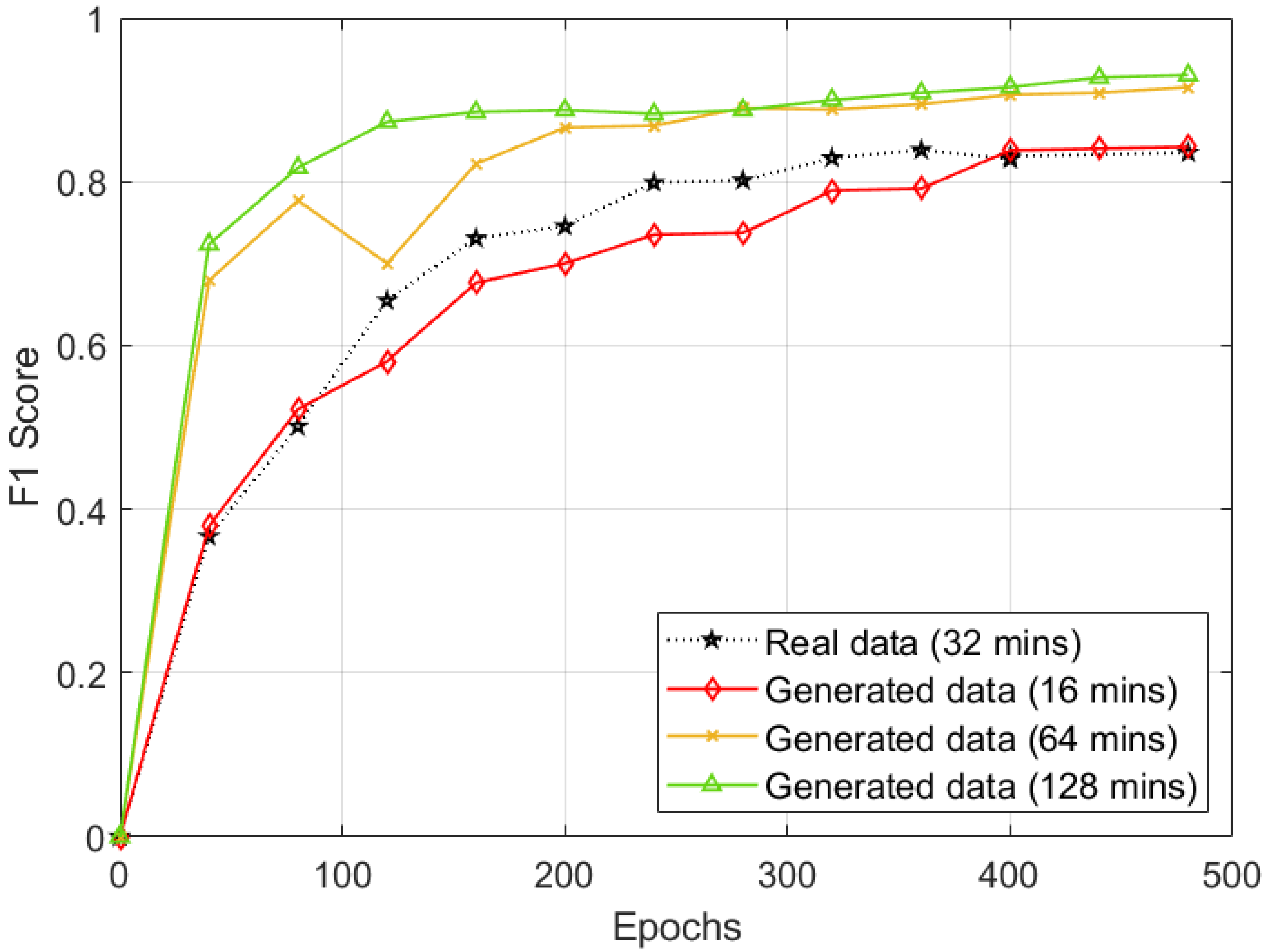


Figure 3. The confusion matrices obtained with CNN models trained on 32 minutes of real data (left), 16 minutes of RFID-ACCLDM generated data (middle), and 64 minutes of RFID-ACCLDM generated data (right).

Downstream Task: Human Activity Recognition (cont'd)



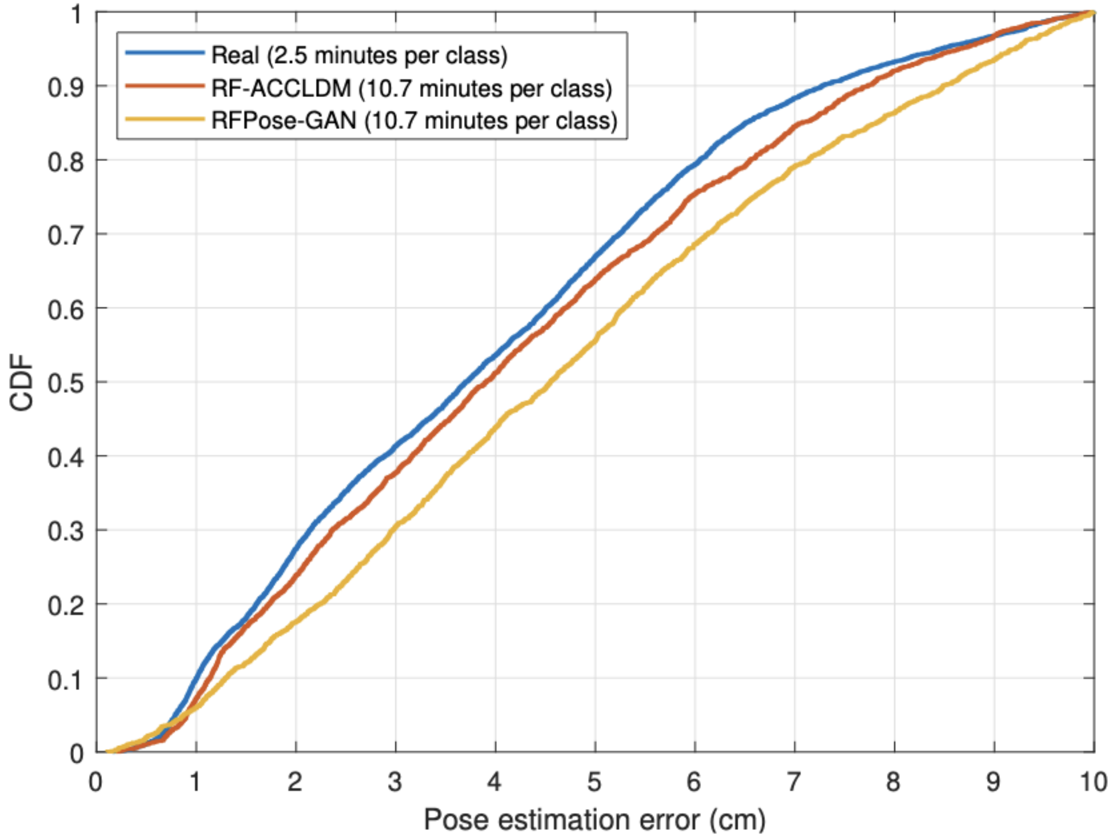
Downstream Task: Human Pose Estimation

How to generate labeled data for supervised training:

- RF-ACCLDM: we first use a pre-trained RFID-Pose model to estimate synthetic poses from ACCLDM generated data, and then employ pairs of generated RFID data and estimated pose for the supervised training
- RFPose-GAN: we pair GAN synthesized RFID data with its input pose, i.e., the simulated pose data

The mean per joint position error (MPJPE):

$$MPJPE = \frac{1}{N}\sum_{n=1}^{N}\left\|\hat{P}_n^t - P_n^t\right\|$$



Overall pose estimation performance regarding complex activities in the form of CDF of estimation errors

Z. Wang and S. Mao, “AIGC for RF sensing: The case of RFID-based human activity recognition,” in *Proc. ICNC 2024*, Big Island, HI, Feb. 2024, pp. 1–6.

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Outline

- Human pose tracking: preliminaries an
- RFID-Pose: 3D human pose monitoring extensions [2,3]
- Generative AI for data augmentation [4
- **Generative AI for 3D pose augmentati**
- Conclusions

[1] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reli*

[2] C. Yang, L. Wang, X. Wang, and S. Mao, "Environment adaptive RFID based 3D human pose tracking with a meta-learning 10.1109/JRFID.2022.3140256.

[3] C. Yang, X. Wang, and S. Mao, "TARF: Technology-agnostic RF sensing for human activity recognition," *IEEE Journal of Bi*

[4] Z. Wang, C. Yang, and S. Mao, "Data augmentation for RFID-based 3D human pose tracking," in *Proc. IEEE VTC-Fall 2022,*

[5] C. Yang, Z. Wang, and S. Mao, "RFPose-GAN: Data augmentation for RFID based 3D human pose tracking," in *Proc. The 1 2022*), Cagliari, Italy, Sept. 2022, pp.138-141.

[6] Z. Wang and S. Mao, "AIGC for RF sensing: The case of RFID-based human activity recognition," in *Proc. ICNC 2024*, Big Is

[7] Z. Wang and S. Mao, "AIGC for wireless data: The case of RFID-based human activity recognition," in *Proc. IEEE ICC 2024*

[8] Z. Wang, C. Yang, and S. Mao, "AIGC for RF-based human activity sensing," *IEEE Internet of Things Journal*, vol.12, no.4, p

[9] Z. Wang and S. Mao, "AIGC for Wireless Sensing: Diffusion-empowered Human Activity Recognition," *IEEE Transactions*

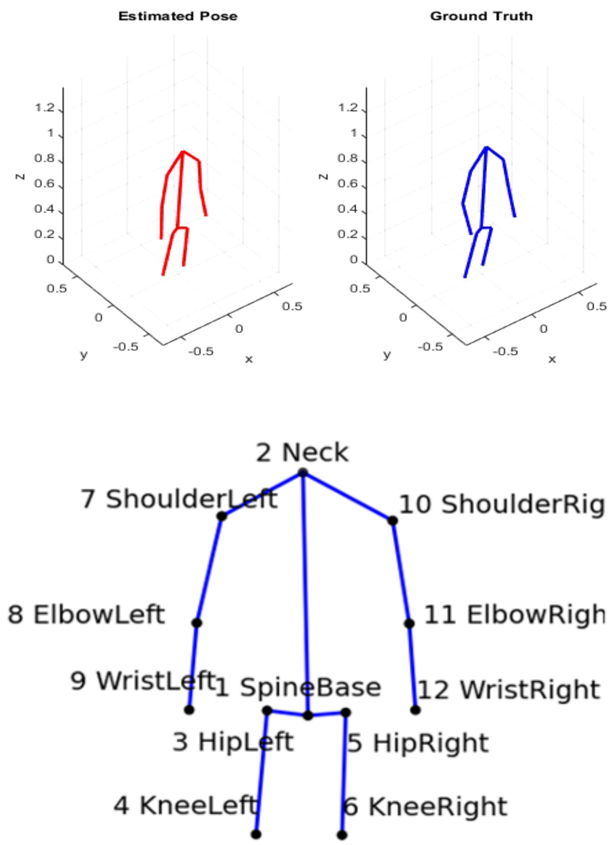
[10] Z. Wang and S. Mao, "Generative AI for 3D human pose completion under RFID sensing constraints," in *Proc. ICNC 2025*

[11] Z. Wang and S. Mao, "Generative AI-empowered RFID sensing for 3D human pose augmentation and completion," *IEEE*

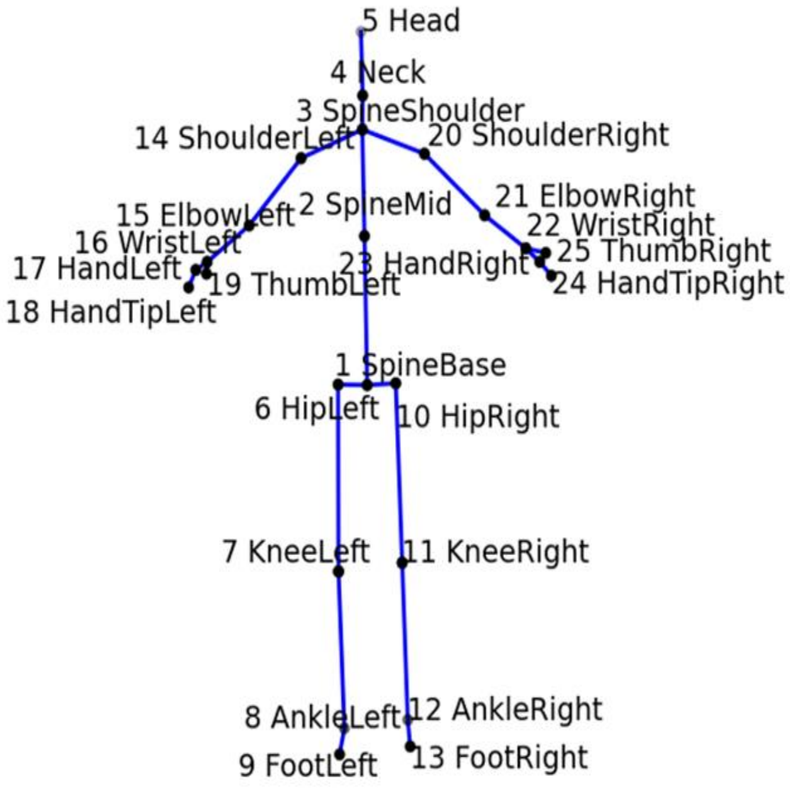
AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Critical Challenges for Deployment in Reality

Challenges: (i) lacking sufficient training data/high cost on collecting training data; (ii) low sampling rate of RFID; (iii) partial pose detected/occlusion



RFID-captured partial pose observation (12 joints)

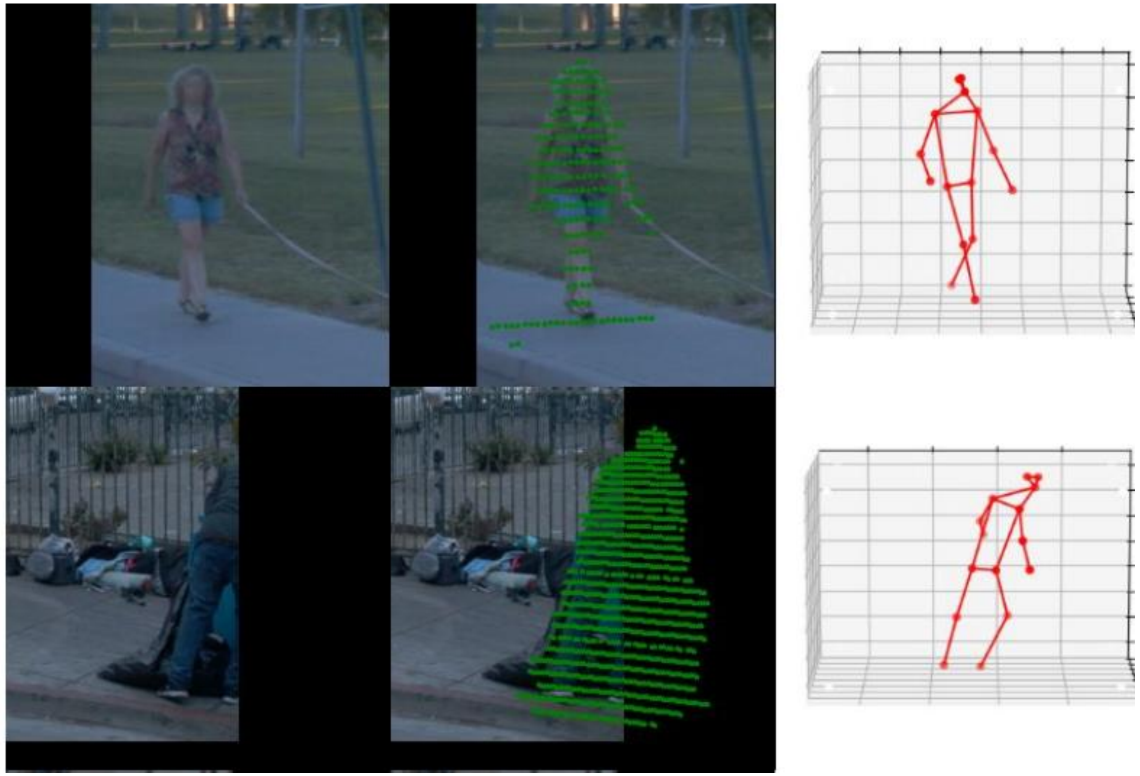


Kinect-captured full pose (25 joints)

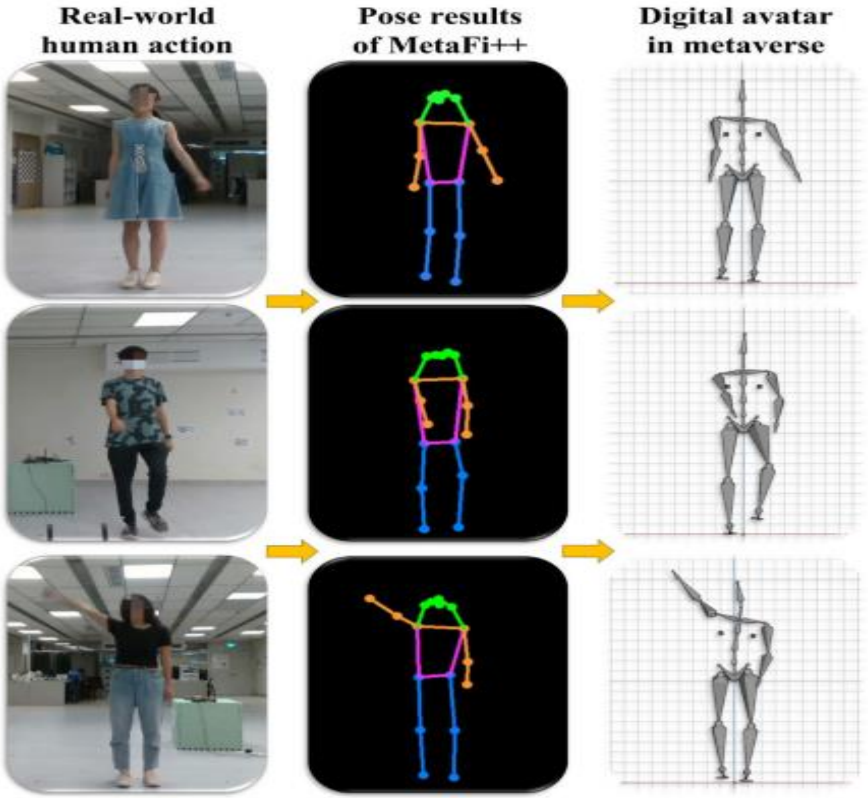
AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Necessity of Full-Body Pose Estimations

- Self-driving companies such as Waymo are stressing the importance of full-body pose estimation under sensing constraints for pedestrian behavior analysis



- A full-body pose with detected joints in the head region is essential to VR/AR related 3D human pose applications



Zanfir, M. Zanfir, A. Gorban, J. Ji, Y. Zhou, D. Anguelov, and C. Sminchisescu, “HUM3DIL: Semi-supervised multi-modal 3Dhumanpose estimation for autonomous driving,” in *Proc. 6th Annual Conf. Robot Learning*, Auckland, NZ, Dec. 2022.
Y. Zhou, H. Huang, S. Yuan, H. Zou, L. Xie, and J. Yang, "MetaFi++: WiFi-Enabled Transformer-Based Human Pose Estimation for Metaverse Avatar Simulation," in *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14128-14136, Aug. 2023.

Problem Statement

- **Pose augmentation**:
  - Generating high-fidelity and temporally smooth synthetic RFID data

$$z_{RFID} \sim p(z_{RFID} \mid \alpha)$$

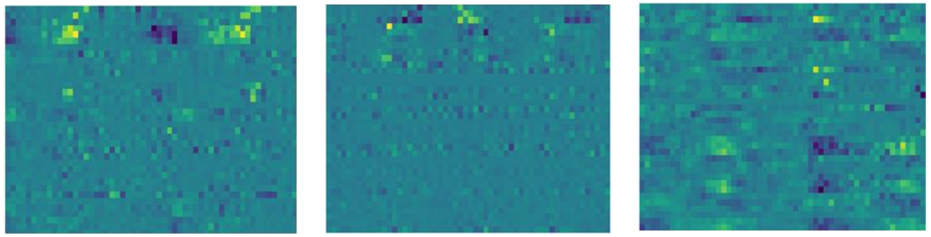
  - Estimating corresponding 3D human pose from this synthetic data using a kinematics predictor.

$$\hat{P}_p = f_{\text{kin}}(\psi_{RFID}(z_{RFID}))$$

- **Pose completion:**
  - Structural: generates a complete 3D pose from partial observations, leveraging the latent representation of partial poses and activity labels

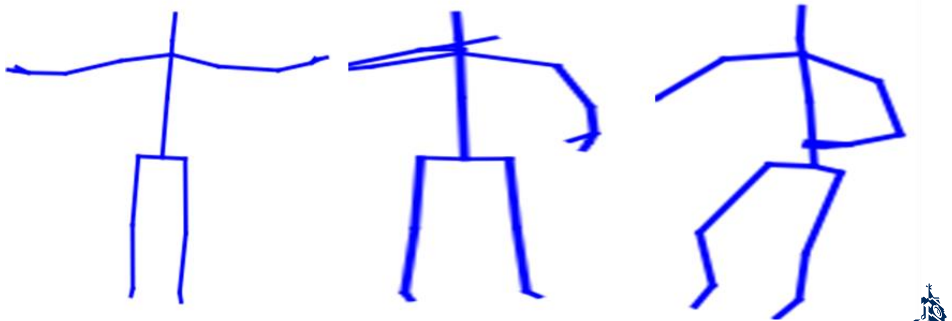
$$z_f \sim p(z_f \mid z_p, \alpha),\ \hat{P}_f = \psi_{Pose}(z_f),$$

  - Temporal: increase frame rate to obtain smooth transitions and coherent motion sequences



Generated RFID Representation

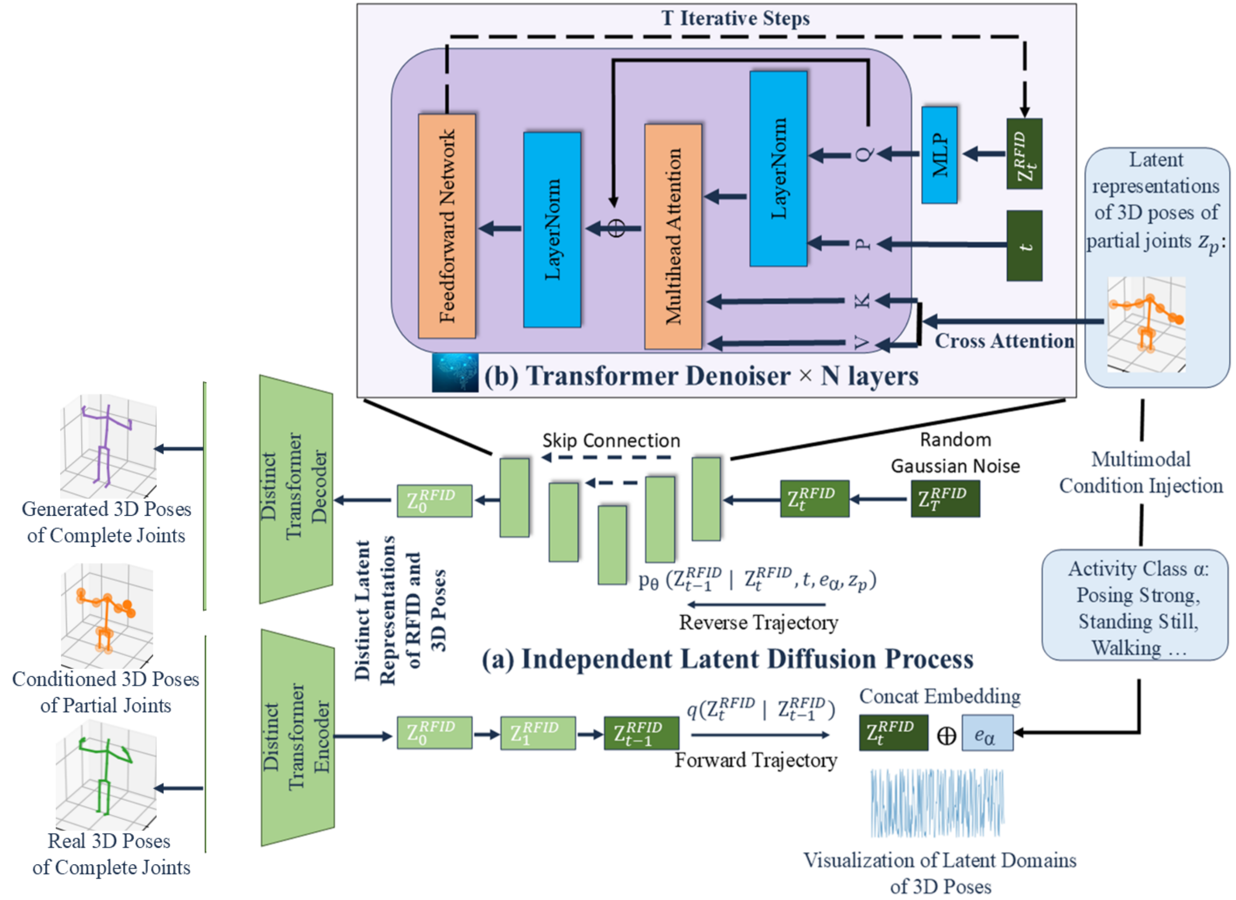
Estimated Partial Pose



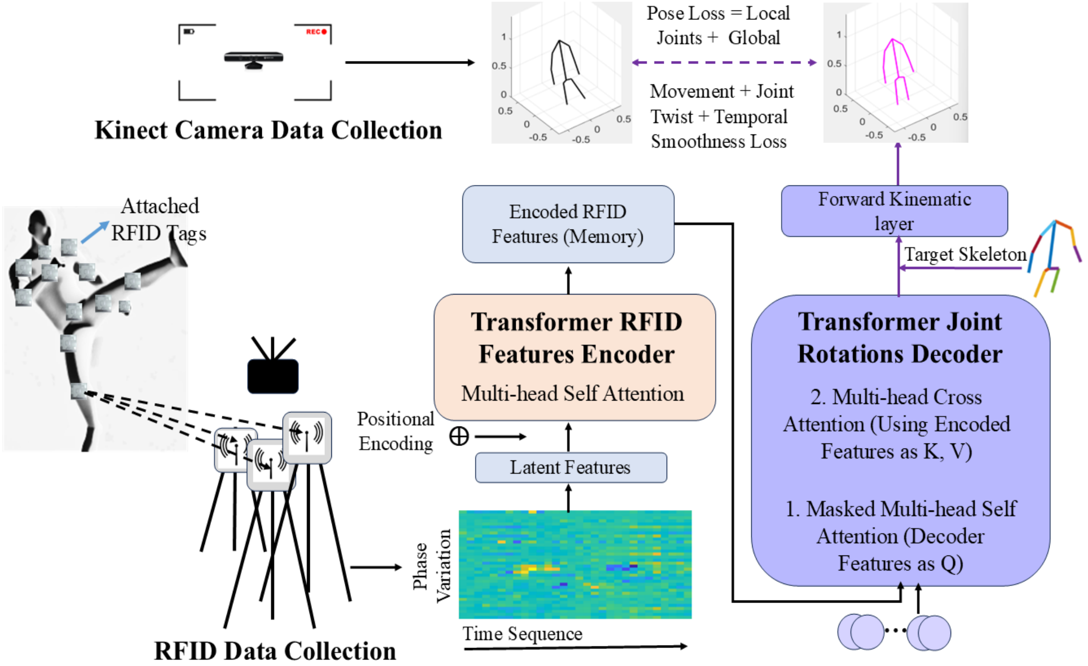
Generated Full Pose

System Design

Transformer-based Latent Diffusion Model



Transformer-based Kinematics Neural Network



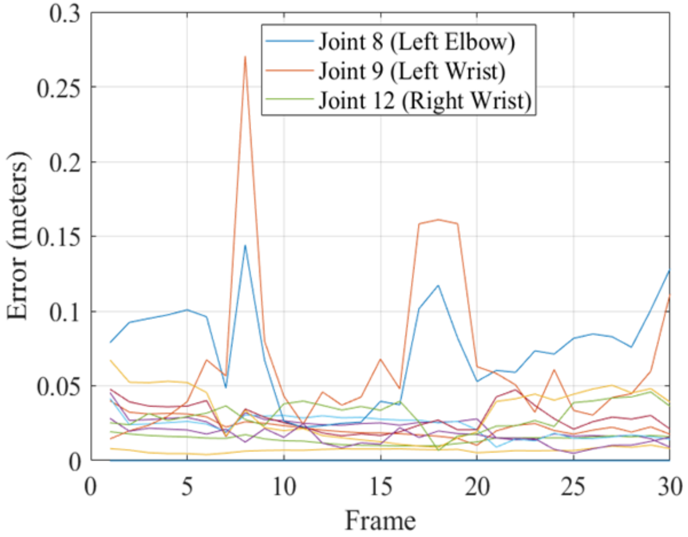
AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Pose Estimation: Train-on-Real, Test-on-Real

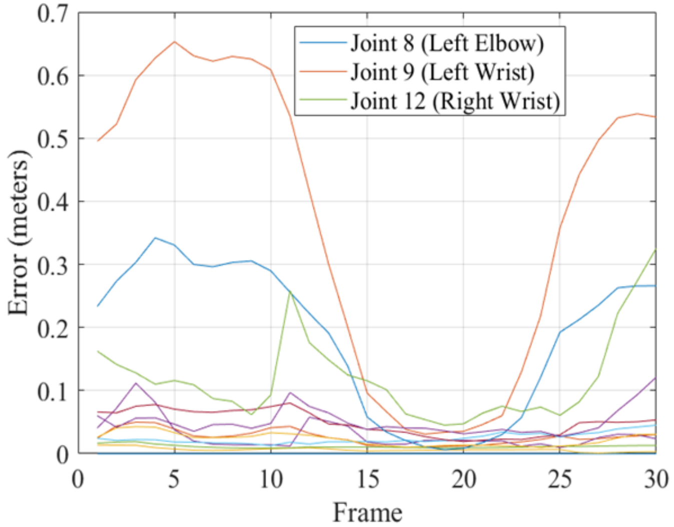
Subject Index	Estimation Error RFID-Pose (cm)	Estimation Error Cycle-Pose (cm)	Estimation Error Proposed (cm)
Subject 1	3.75	4.12	3.34
Subject 2	4.55	4.43	3.47
Subject 3	3.58	3.79	3.05
Subject 4	5.32	4.51	4.91
Subject 5	8.17	4.97	5.65

[1] Note: Subjects 1-3 are trained; Subjects 4-5 are untrained.

LDT



RNN Baseline



Two baselines:

[1] Chao Yang, Xuyu Wang, and Shiwen Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, vol.70, no.3, pp.1218-1231, Sept. 2021.

[2] Chao Yang, Xuyu Wang, and Shiwen Mao, "RFID based 3D human pose tracking: A subject generalization approach," *Elsevier/KeAi Digital Communications and Networks*, vol.8, no.3, pp.278-288, Aug. 2022.

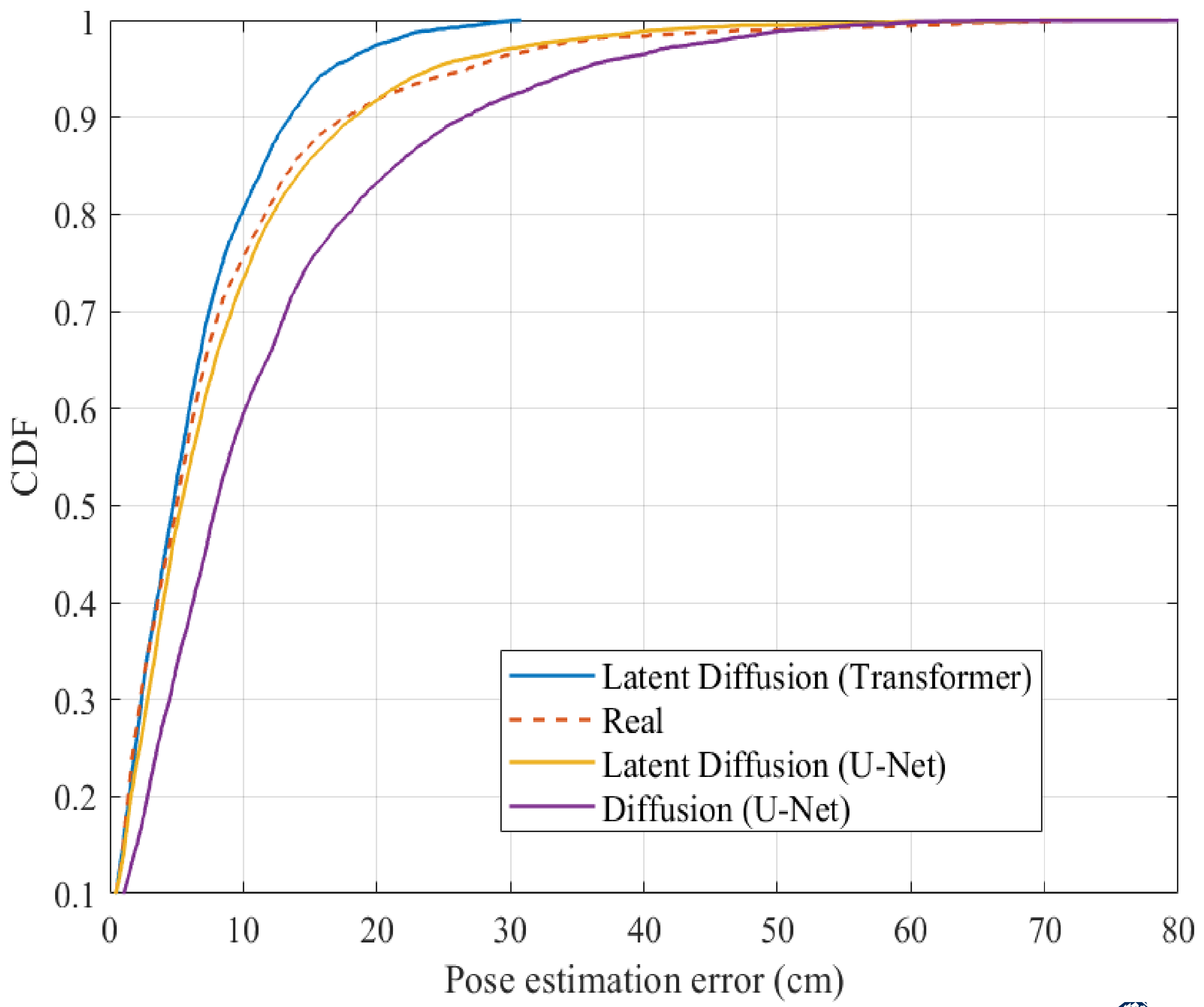
Pose Detection: Train-on-Synthetic, Test-on-Real

Table 3. **Evaluation Metrics for Estimated 3D Human Poses Using LDT Generated RFID Data**

Metrics	LDT	Metrics	LDT
Average joint error (cm)	8.99	FID	1.42
Bone consistency (cm)	2.25	GT FID	0.73
Joint angle error (°)	6.91	Diversity	10.98
Smoothness (cm/frame)	1.51	GT Diversity	10.35
GT Smoothness (cm/frame)	1.40		

Table 4. **Evaluation Metrics for Estimated 3D Human Poses Using LDT Generated WiFi CSI Data**

Metrics	LDT	Metrics	LDT
Average joint error (cm)	9.33	FID	4.46
Bone consistency (cm)	2.33	GT FID	0.85
Joint angle error (°)	7.52	Diversity	11.53
Smoothness (cm/frame)	1.03	GT Diversity	11.75
GT Smoothness (cm/frame)	1.38		



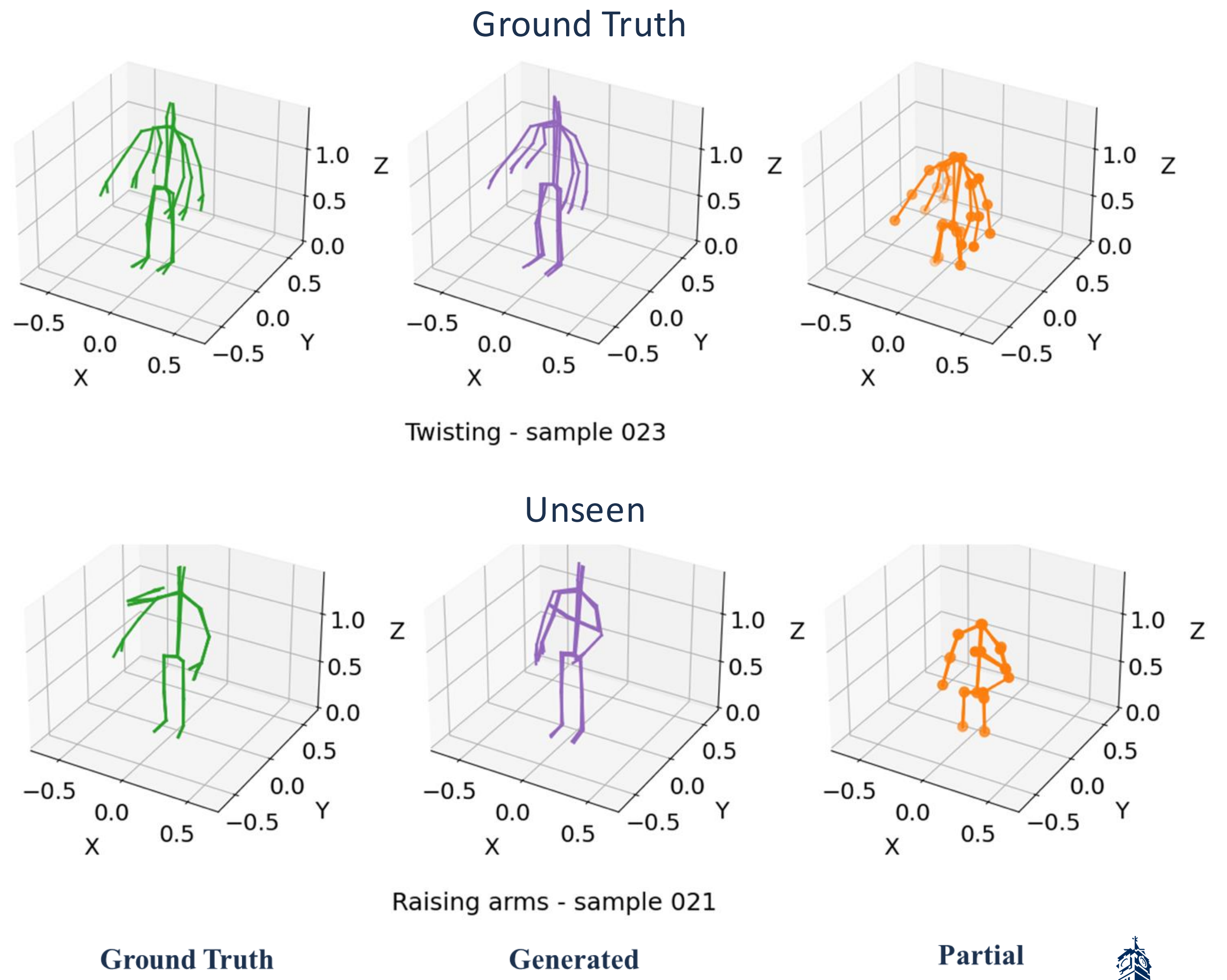
AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Structural Completion

- A two-stage motion-aligned generation process
  - Initial generation with attention capture
  - Motion-aligned refinement
- Generates anatomically consistent and temporally aligned full-body 3D poses

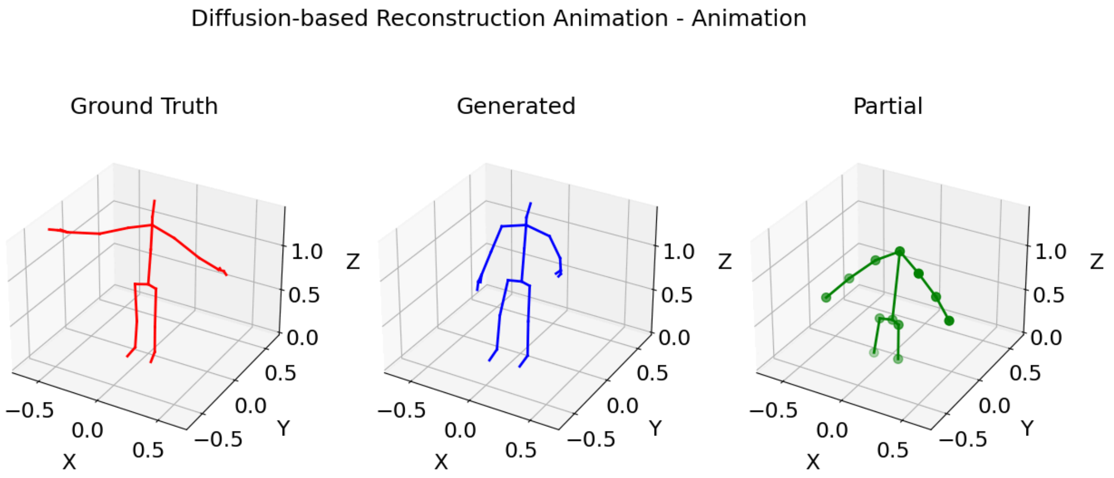
Table 5. Evaluation Metrics for 3D Pose Completion with Ground Truth and Unseen Partial Pose Conditioning

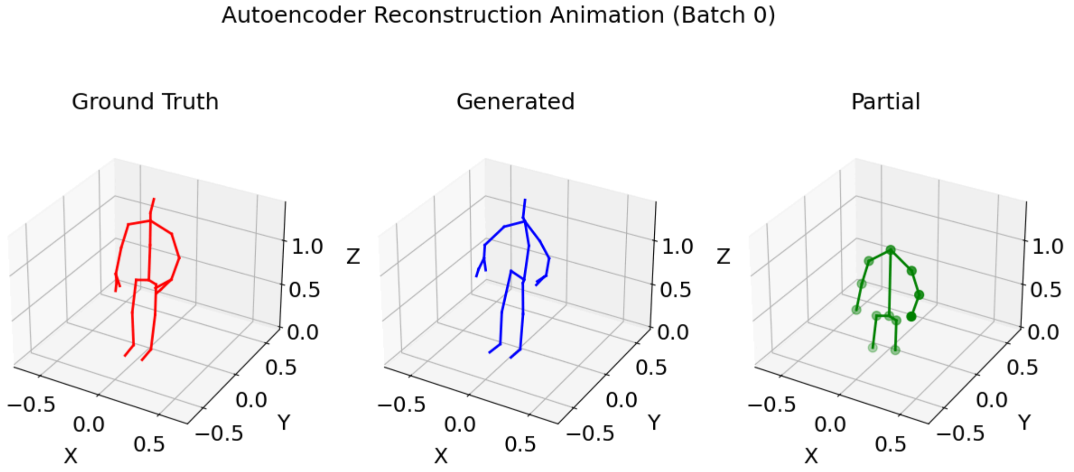
Metrics	Ground Truth	Unseen
Avg joint error (cm)	11.74	19.23
Bone consistency (cm)	1.77	2.12
Joint angle error (°)	6.65	11.13
Smoothness (cm/frame)	2.46	1.90
FID (-)	0.87	4.67
Diversity (-)	26.59	13.71
Trajectory joint error (cm) compared with partial pose	7.24	8.11
Trajectory velocity error (cm/frame) compared with partial pose	7.56	7.80

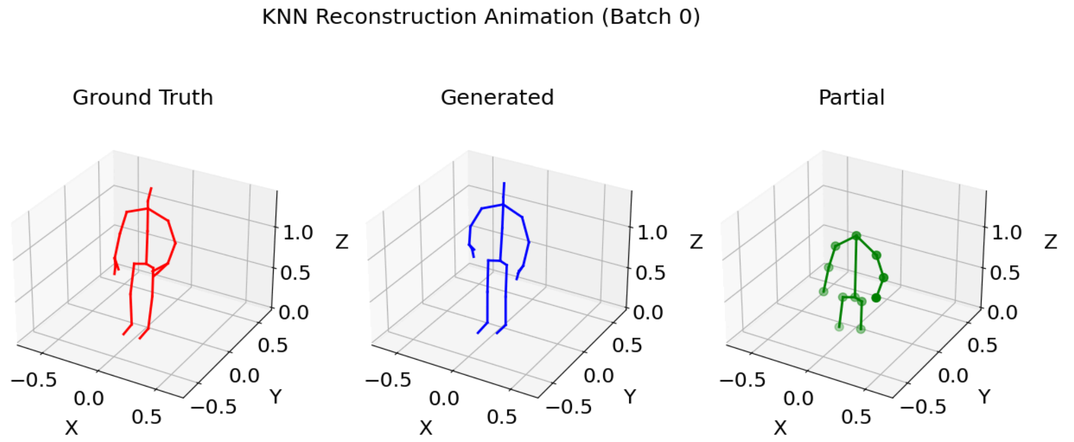


AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Comparison with Baselines

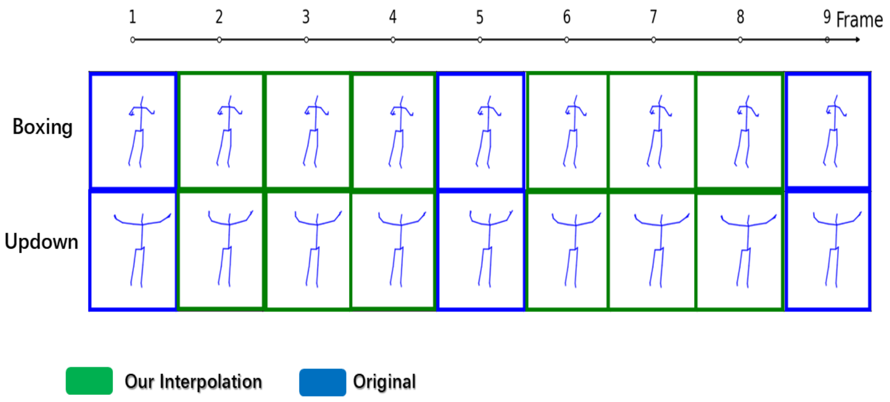


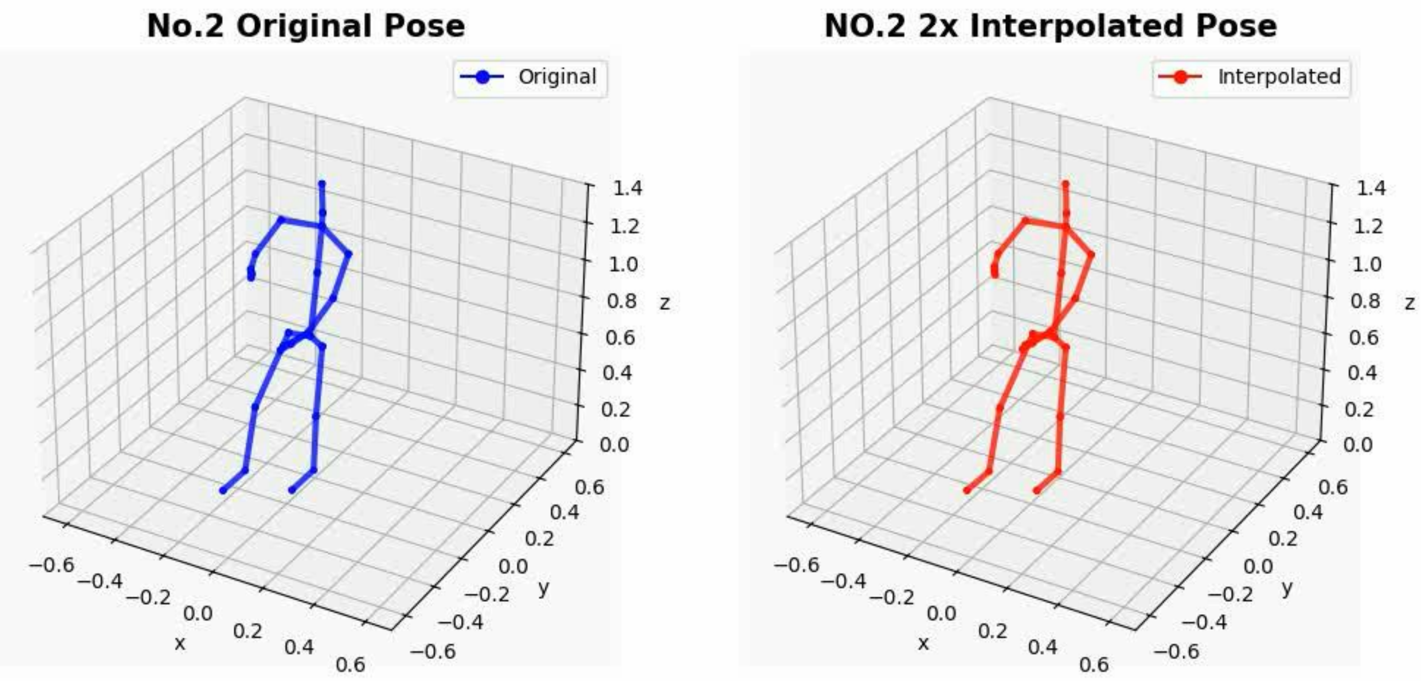




3D Human Pose Frame Interpolation

- A 2D U-Net-based frame interpolation method to up-sample the estimated poses by up to 30 Hz
  - It takes several frames before and after the target interval as input to predict the intermediate pose frames
- Achieves smaller temporal smoothness errors than traditional methods such as linear and cubic interpolation





AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

Conclusions

- RF sensing for 3D human pose tracking
- Real-time 3D human pose tracking and classification with commodity RFID devices, and its enhancements
- Data augmentation for RF sensing: GAN, Diffusion, and Stable Diffusion based approaches
- Pose augmentation and completion: Latent Diffusion Transformer based approach, structural and temporal completion

This work is supported in part by the National Science Foundation (NSF) under Grants CNS-2107190, CNS-2148382, IIS-2306789, and CCSS-2245608, and by the Wireless Engineering Research and Education Center at Auburn University

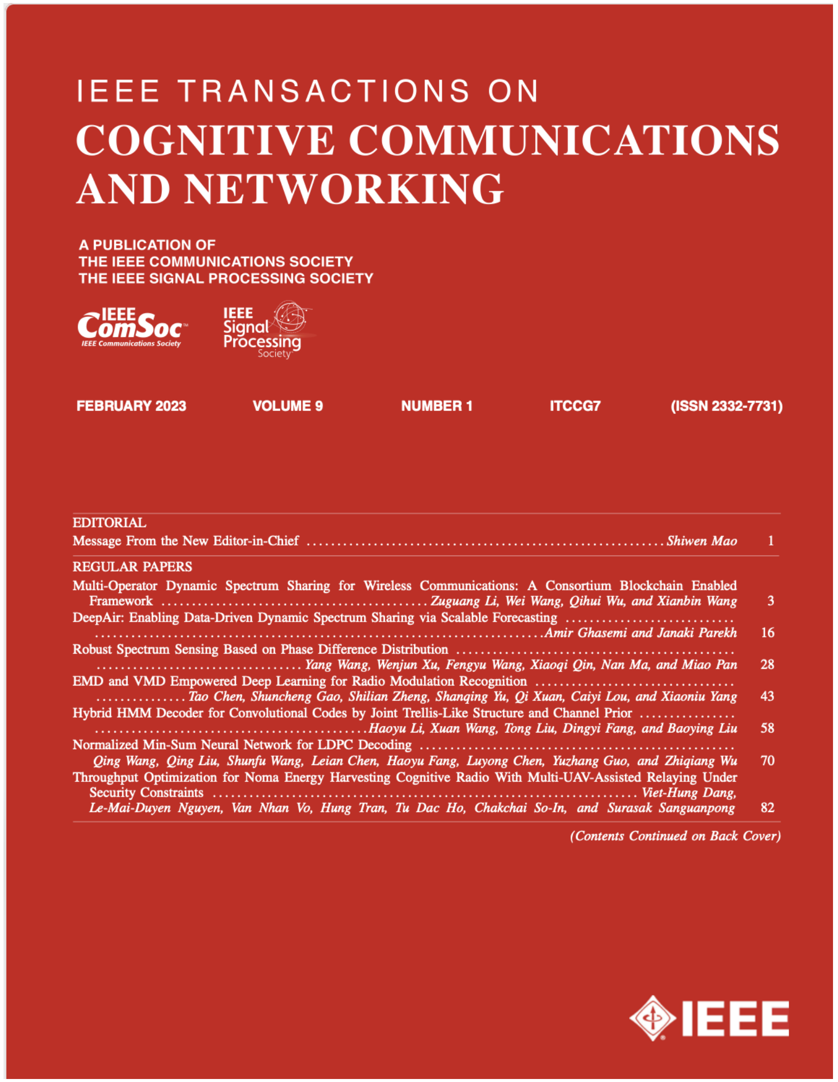
Thanks my hosts of this ComSoc DL tour: Dr. Takaya Yamazato and Dr. Noriaki Kamiyama (IEEE ComSoc Tokyo (Joint) Chapter), and Dr. Nei Kato (IEEE ComSoc Sendai Chapter), and many others who made this DL tour possible 🥹

For more information: http://www.eng.auburn.edu/~szm0001/

AUBURN UNIVERSITY
Wireless Engineering Research and Education Center

IEEE Transactions on Cognitive Communications and Networking

- **Scope**: Cognitive radios, spectrum, AI-empowered communications and networking, and resource allocation
- **Areas**:
  - Cognitive Radio
  - AI-empowered Communications
  - AI-empowered Networking
  - AI-empowered Resource Allocation
- **Impact factor**: 4.574 (2019) ➔ 7.0 (2024)
  - Q1 in Telecommunications (17/120=14.2%), Artificial Intelligence, Computer Networks and Communications, and Hardware & Architecture
  - Top 48 among all IEEE publications: 48/215=22.3%
  - Journal Ranking by the National Science and Technology Library, Chinese Academy of Sciences (CAS): Top Tie 1
- **Special issues**: backscatter communications, generative AI, ISAC and RIS, Near-Field technologies, Low-altitude economy, Fluid antenna (Oct. 30, 2025), Agentic AI (Jan. 1, 2026)
- **Home**: https://www.comsoc.org/publications/journals/ieee-tccn
- **Submission**: IEEE Author Portal


IEEE TRANSACTIONS ON
COGNITIVE COMMUNICATIONS AND NETWORKING

A PUBLICATION OF
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY

FEBRUARY 2023 VOLUME 9 NUMBER 1 ITCCG7 (ISSN 2332-7731)

EDITORIAL

REGULAR PAPERS

(Contents Continued on Back Cover)

IEEE