ORIGIN™

WiFi can do more.

WiFi-Based Robust Human and Non-human Motion Recognition With Deep Learning

Guozhen Zhu, Beibei Wang, Weihang Gao, Yuqian Hu, Chenshu Wu, K. J. Ray Liu
Origin Research, Rockville, MD, USA
{guozhen.zhu, beibei.wang, weihang.gao, yuqian.hu, chenshu.wu, ray.liu}@originwirelessai.com

Guozhen Zhu

March 15th 2024

Content

- Introduction
- Framework Design
  - WiFi Signal Preprocessing
  - Deep Neural Networks
- Evaluation
- Conclusion

WiFi-based human and non-human recognition

In many indoor environments, the moving subjects can be:

Human motion

Human

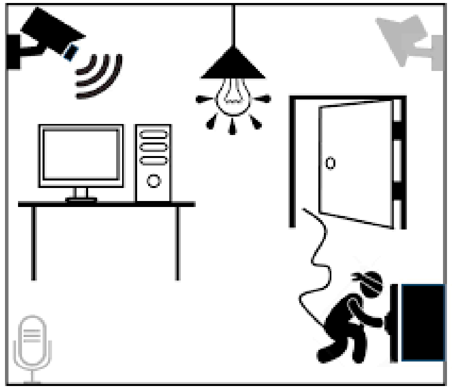
Non-human

Non-human motion

Pets, robots, electrical appliances…

Non-human motion introduce high false alarm rates to intelligent systems and applications:

- Intruder detection
- Occupancy monitoring
- Smart buildings
- Activity recognitions…



WiFi-based human and non-human recognition

- Pets, vacuum machines, and electrical appliances are essential parts of families and extensively exist in various environments.
  - About 85 million families in US have pets in 2022 (American Pet Products Association).
  - The global robotic cleaners market is $5.59 billion in 2021.
  - Electrical appliances such as fans and washing machines are common in households.

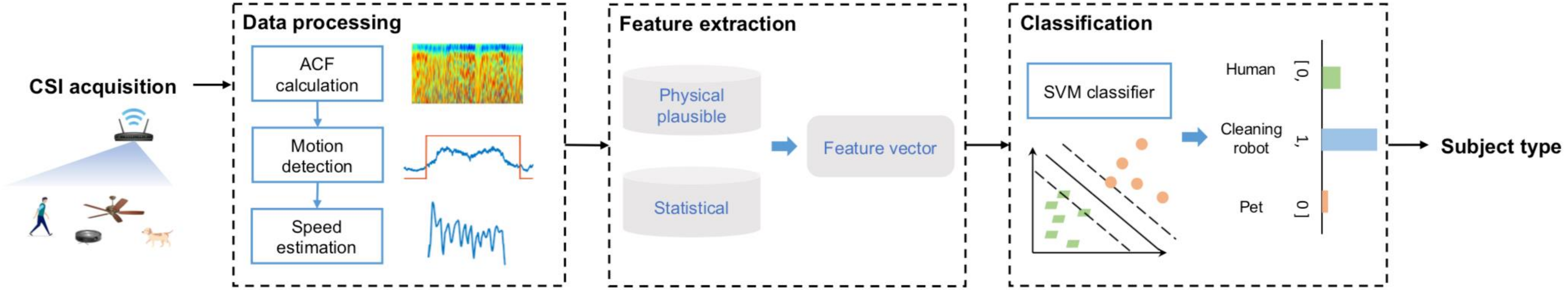
→ Distinguishing human and non-human subjects is crucial for practical indoor intelligent applications and systems.



ORIGIN™

WiFi-based human and non-human recognition

Previous work using hand craft features



Feature vector

- Physical features:
  - existence of gait, stride length, stride cycle time, average speed, speed variance, speed 25 percentile, speed 75 percentile
- Statistical features:
  - ACF peaks mean, ACF valleys mean, ACF peaks interval distance, ACF valleys interval distance, motion statistic mean, motion statistic variance

WiFi-based human and non-human recognition

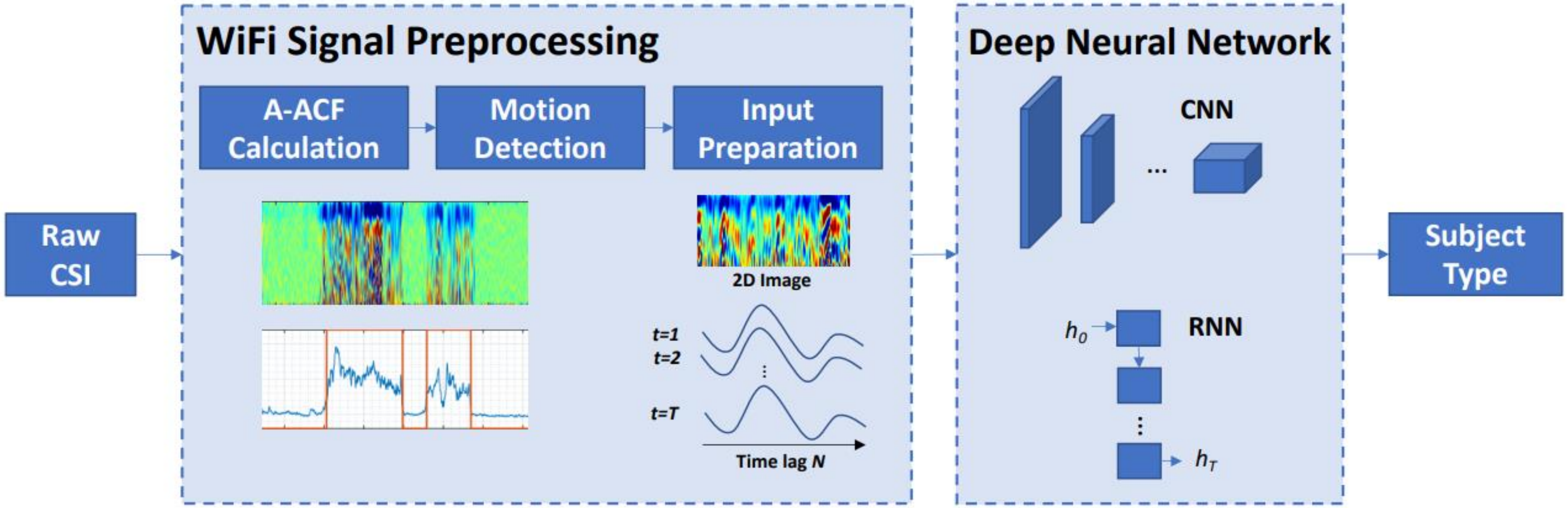
Motivation

- Challenging cases:
  - Features cannot be derived from data:
    - Device limitations:
      - Coverage
      - Noise
    - Motion is complex:
      - Intrusion
      - Multiple people
  - Features show overlap:
    - Pet too large

Need more powerful models that can directly extract features from data!

WiFi-based human and non-human recognition

Framework Overview



System Design

Input preparation

1. Take channel power response of CSI

$$G(t,f) \triangleq |\tilde{H}(t,f)|^2$$
$$= |H(t,f)|^2 + 2\,\mathrm{Re}\{n^*(t,f)H(t,f)$$
$$\exp(-j(\alpha(t)+\beta(t)f))\} + |n(t,f)|^2$$
$$\triangleq |H(t,f)|^2 + \varepsilon(t,f),$$

2. ACF of channel power response

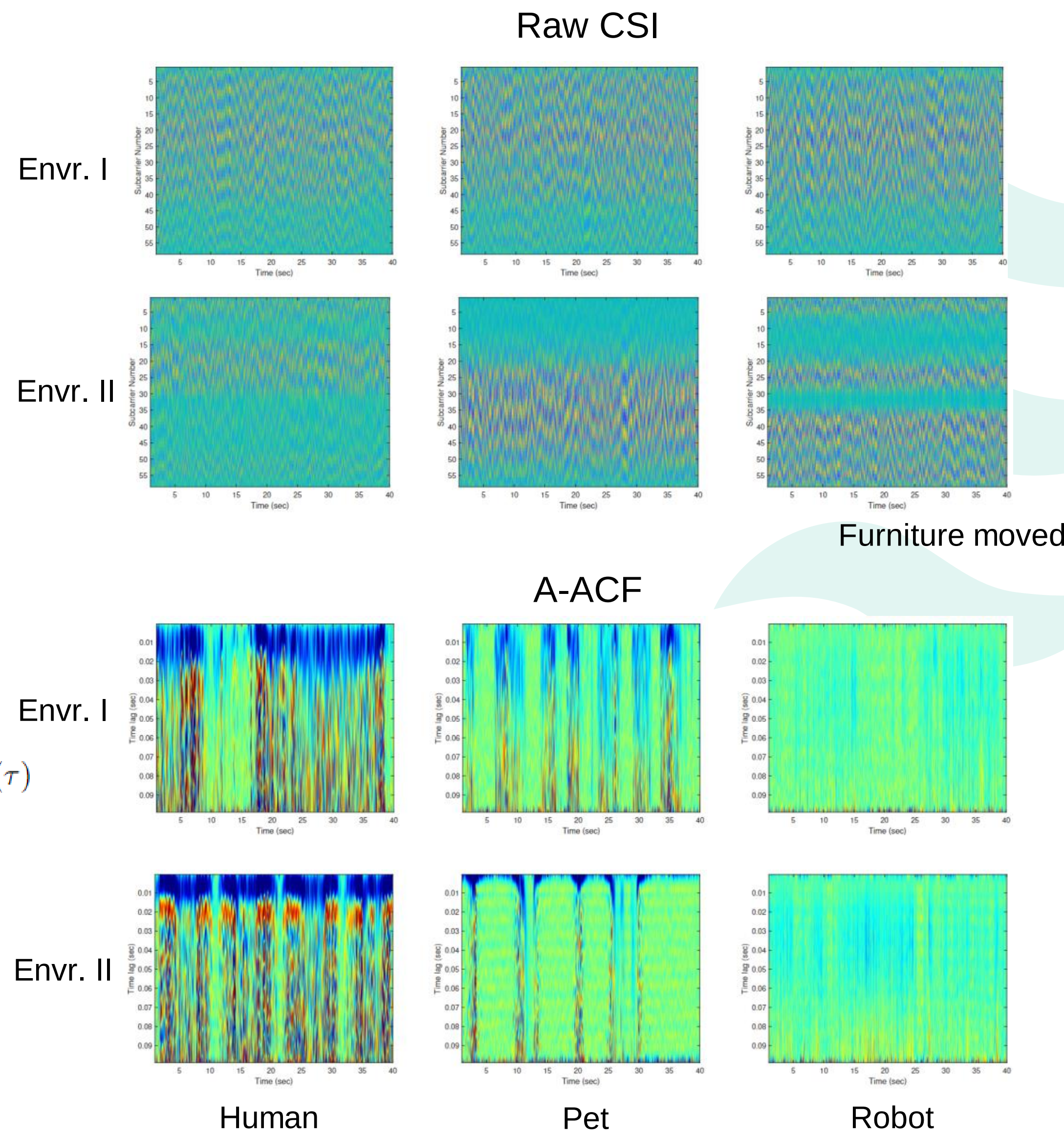
$$\rho_G(\tau,f) = \frac{\mathrm{cov}[G(t,f),G(t+\tau,f)]}{\mathrm{cov}[G(t,f),G(t,f)]}$$
$$= \frac{\sigma_s^2(f)}{\sigma_s^2(f)+\sigma_n^2(f)}\rho_s(\tau,f) + \frac{\sigma_n^2(f)}{\sigma_s^2(f)+\sigma_n^2(f)}\delta(\tau)$$

3. Apply Maximum Ratio Combine (MRC)

$$\hat{\rho}_s(\tau) = \sum_{i=1}^{N_s} \rho_G(\tau = \frac{1}{F_s}, f_i)\rho_G(\tau, f_i)$$

4. Take the differential

$$\Delta\hat{\rho}_s(\tau)$$



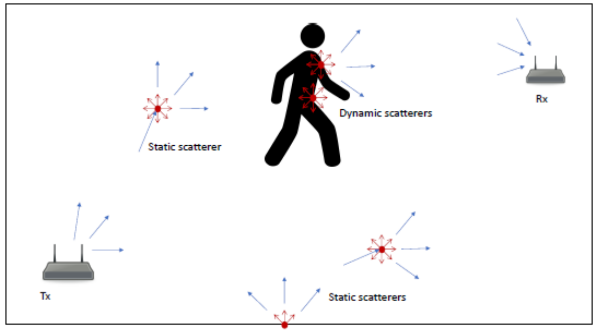
WiFi-based human and non-human recognition

Motion Detection

The autocorrelation function (ACF) of the CSI H(t, f) with the time lag τ:

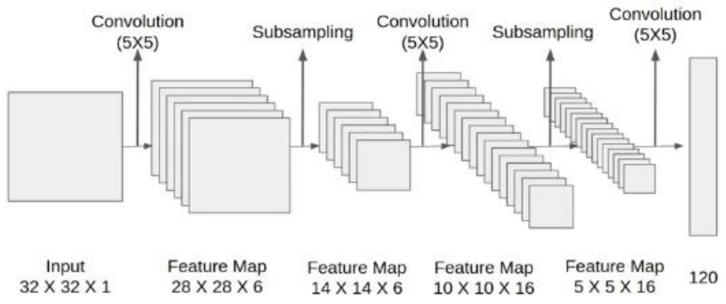
$$\rho_G(\tau, f) = \frac{\mathrm{cov}[G(t,f), G(t+\tau,f)]}{\mathrm{cov}[G(t,f), G(t,f)]} = \frac{\sigma_s^2(f)}{\sigma_s^2(f)+\sigma_n^2(f)}\rho_s(\tau,f) + \frac{\sigma_n^2(f)}{\sigma_s^2(f)+\sigma_n^2(f)}\delta(\tau)$$

→ motion statistic $\phi_G(f) \triangleq \rho_G\left(\tau = \frac{1}{F_s}, f\right)$



[1] F. Zhang, C. Chen, B. Wang and K. J. R. Liu, "WiSpeed: A Statistical Electromagnetic Approach for Device-Free Indoor Speed Estimation," in *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2163-2177, June 2018.

WiFi-based human and non-human recognition



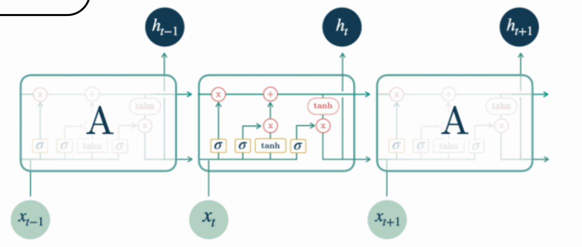
Deep Learning Models

- **Categories:**
  - Convolutional neural network
    - LeNet
    - ResNet (-18, -50, -101)

Pro: extract spatial hierarchies with 2D kernel
Con: focus on local

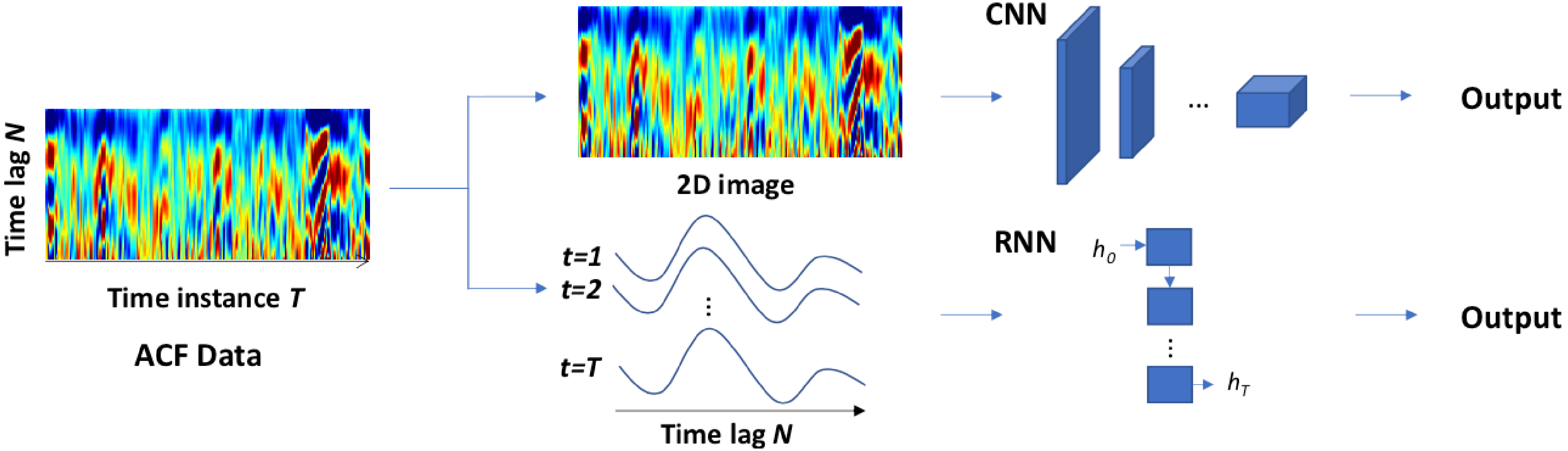
  - Recurrent neural network
    - RNN
    - GRUNet
    - LSTM

Pro: good at extract temporal dependencies
Con: neglect patterns of A-ACF within a time stamp



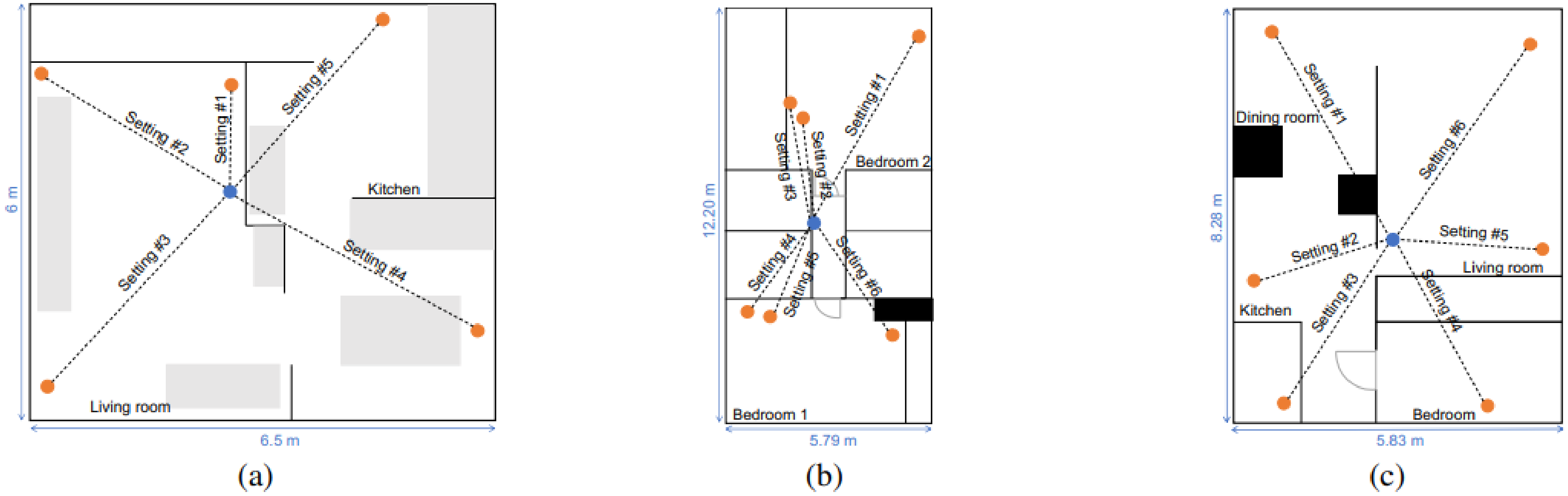
WiFi-based human and non-human recognition

Input Preparation



WiFi-based human and non-human recognition

Evaluation Environments



Floor plan of (a) Scenario I, an apartment, (b) Scenario II, a townhouse, and (c) Scenario III, a single family house.

Evaluation Metrics

$$Recall\ (True\ Possitive\ Rate)\ = \frac{True\ Possitive}{Total\ Possitive}$$

$$Accuracy = \frac{True\ Possitive + True\ Negative}{Total}$$

$$Precision\ (Positive\ Predictive\ Values) = \frac{True\ Possitive}{True\ Possitive + False\ Possitive}$$

$$False\ Alarm\ Rate(False\ Possitive\ Rate)\ = \frac{False\ Possitive}{Total\ Negative}$$

Evaluation

Dataset

Scenario	Human	Pet	Cleaning robot	Fan
I	F1, M1, M2, M3, M4, M5, *M6, M7, M8, F2, F3*	D1, D2, *D3, D4, D5, D6, D7, D8*	iRobot V3	Rotation
II	F1	–	iRobot V3	–
III	M2, M3, *M9*	C1	–	–

- Human:
  - **12 individuals** (9 males, 3 females)
  - Age: 23 to 34 years old
  - Height: **154** cm to 194 cm
  - Participants are free to walk, run, sneak, or stop and have small motions during data collection such as using their phones while walking.
- Pets:
  - **9 different pets**, including 8 dogs and a cat
  - Weight: **17 lb** to **85 lb**
  - Pets are allowed to move freely.
- Cleaning robot
  - iRobot V3 vacuum machine
- Fan:
  - A rotation fan

WiFi-based human and non-human recognition

Evaluation – Classification Performance

CLASSIFICATION ACCURACY OF DEEP LEARNING MODELS FOR HUMAN AND NON-HUMAN MOTION IDENTIFICATION

Model	Exp. I	Exp. II	Exp. III	Average
LeNet	91.29%	88.69%	98.19%	92.72%
ResNet18	94.60%	92.78%	99.34%	95.57%
ResNet50	93.28%	92.46%	98.68%	94.81%
ResNet101	93.66%	90.52%	98.59%	94.25%
GRUNet	75.66%	72.20%	97.65%	81.84%
RNN	75.38%	79.31%	97.86%	84.18%
LSTM	86.46%	82.76%	96.71%	88.64%

- LeNet achieves an average accuracy of 92.72%.
- The ResNet family of models demonstrated superior performance, especially ResNet18, which achieved the highest average accuracy of **95.57%.**
- Other deep learning models like RNN, GRUNet, and LSTM performed comparatively lower, with average accuracies of 84.18%, 81.84%, and 88.64%, respectively.

EVALUATION RESULTS OF FOUR-CLASS MOTION RECOGNITION

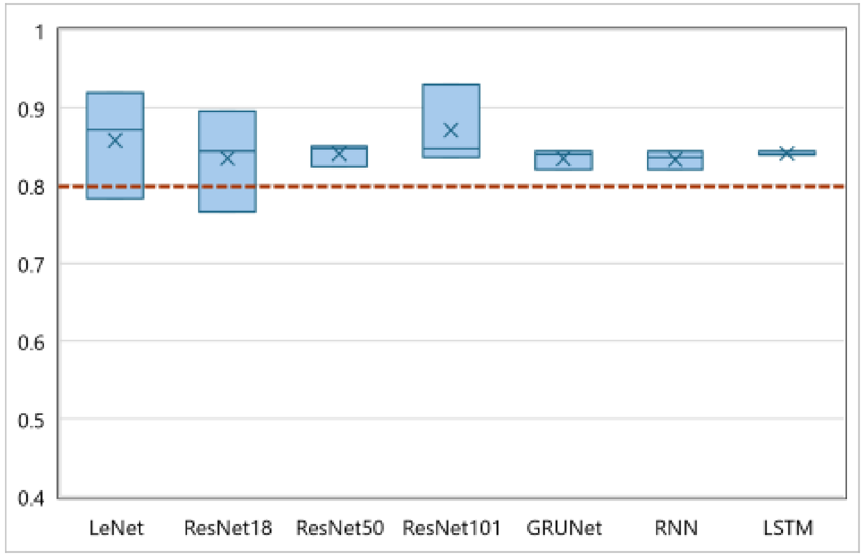
Model	Average Precision	Average Recall	Average F1-score
LeNet	92.00%	91.96%	91.92%
ResNet18	94.23%	93.52%	93.70%
ResNet50	91.42%	91.71%	91.31%
ResNet101	91.82%	90.74%	90.83%
GRUNet	68.67%	81.66%	73.90%
RNN	73.90%	81.86%	77.12%
LSTM	87.73%	87.85%	87.50%

- Overall, the results demonstrate the superiority of our proposed deep learning-based framework in accurately identifying human and non-human subjects through-the-wall with WiFi.

The remarkable performance of the ResNet models affirms the efficacy of the deep learning approach in complex detection tasks.

WiFi-based human and non-human recognition

Evaluation – Robustness



Testing accuracy in unseen environments.

- Followed the **leave one-environment-out methodology** and thus get three testing experiments.
- Each experiment contains a balanced dataset that can reflect the **real-world conditions**, where the data size ratio of human to non-human targets is 5:1.

- ResNet101 achieved an impressive average testing accuracy of **87.09%.**
- **All models surpassed the 80%** mark in average testing accuracy.
- These outcomes validate that our proposed architecture minimizes the impact of factors such as environment, position, or direction on recognition performance.

This enables effective recognition of human and various non-human subjects in new environments without requiring additional training or parameter adjustments

WiFi-based human and non-human recognition

Evaluation – Model Parameter

Model	LeNet	ResNet-18	ResNet-50	ResNet-101	RNN	GRUNet	LSTM
Parameter Size (MB)	14.7	42.7	81.3	158.0	0.082	0.24	0.055

For implementation on edge devices, we need to balance the trade-off between memory requirement and recognition performance.

WiFi-based human and non-human recognition

Conclusion

- We introduce a framework for **recognizing human and various non-human subjects** using **a single pair of WiFi devices** with deep learning.
- Our system addresses the common challenge of non-human interference by leveraging **a unique and context-agnostic feature A-ACF** as input for deep neural networks, which ensures **robust performance regardless of environment, location, or direction**.
- Leverage a deep neural network to extract features from **a designed statistic, A-ACF**, Rigorous tests across diverse scenarios not only affirm the **high accuracy of the proposed system on four types moving subjects recognition in complex settings** but also **guide the choice of the most suitable deep learning architectures for WiFi sensing tasks**.

Thank You