One is Enough: Enabling One-shot Device-free Gesture Recognition with COTS WiFi

Leqi Zhao, Rui Xiao, *Jianwei Liu, *Jinsong Han

Zhejiang University, China

INFOCOM 2024

Human Gesture Recognition

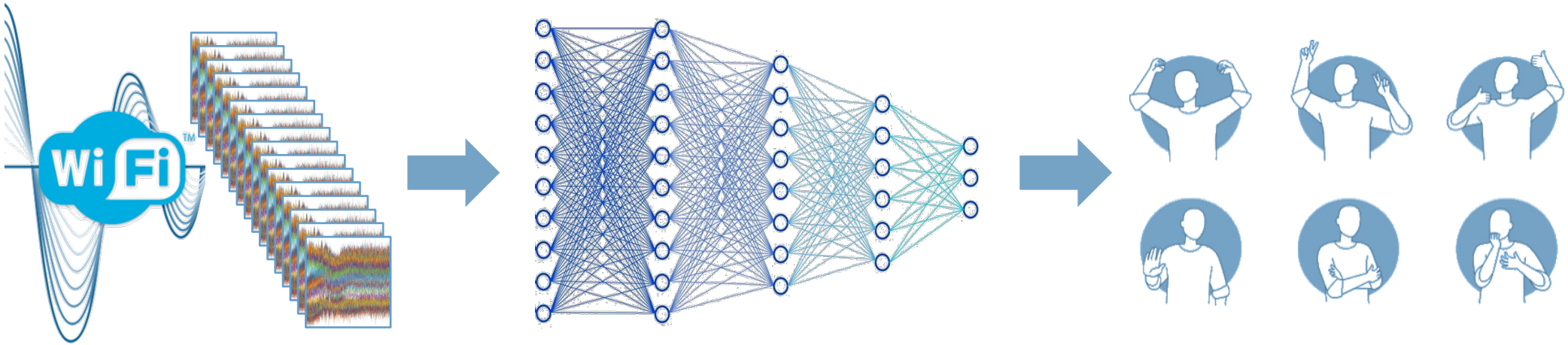
Visual Privacy Preserving

Robust to Occlussion

Non-intrusive

Widely Deployed

WiFi-based Gesture Recognition (WGR)



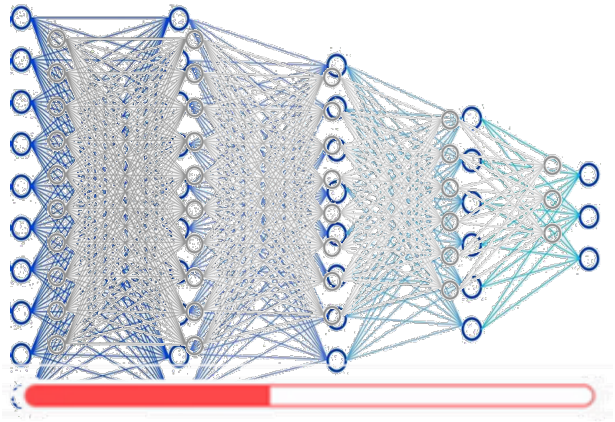
WiFi Signals

Traditional Learning

Gesture Classes

Drawbacks of Existing WGR

Collect signals of new gestures



Train the model from scratch

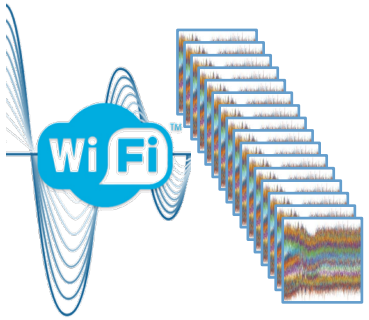
Unseen Gestures

Gesture Classes

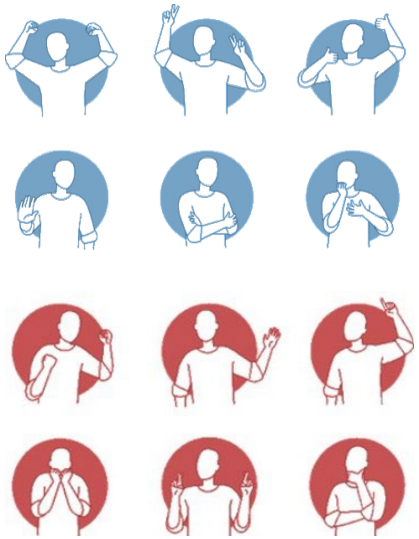
Data Collection Overhead

Model Training Overhead

Our Goal



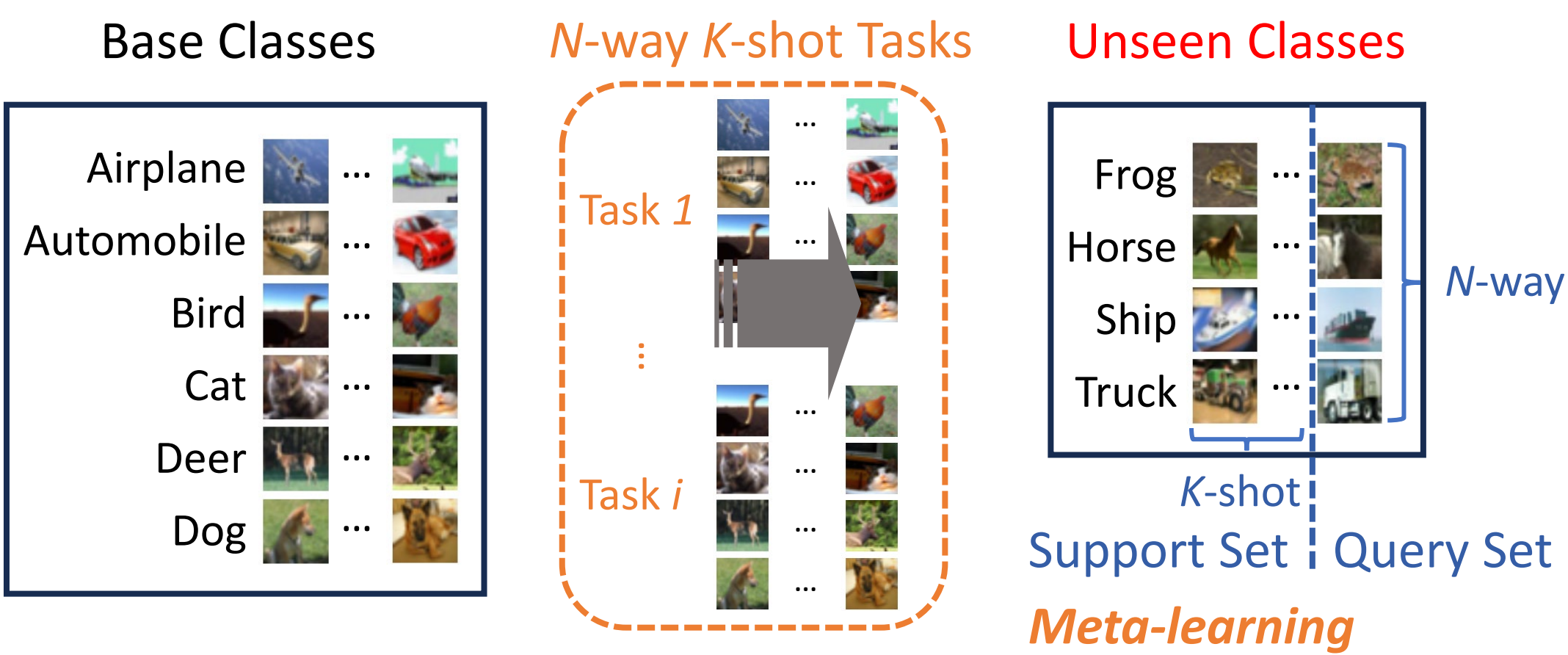
A possible solution:
Few-shot Learning?



A **scalable** WiFi-based gesture recognition system that adapts to **unseen** gestures with **low**

Data Collection Overhead & **Model Training Overhead**

Few-shot Learning (FSL)



Challenges

Directly using classic FSL methods such as *Meta-learning*

does **NOT** work well in WGR

Three Challenges

- Require a diverse training dataset

  ➡ **Laborious data collection (C1)**

- Construct copious few-shot tasks

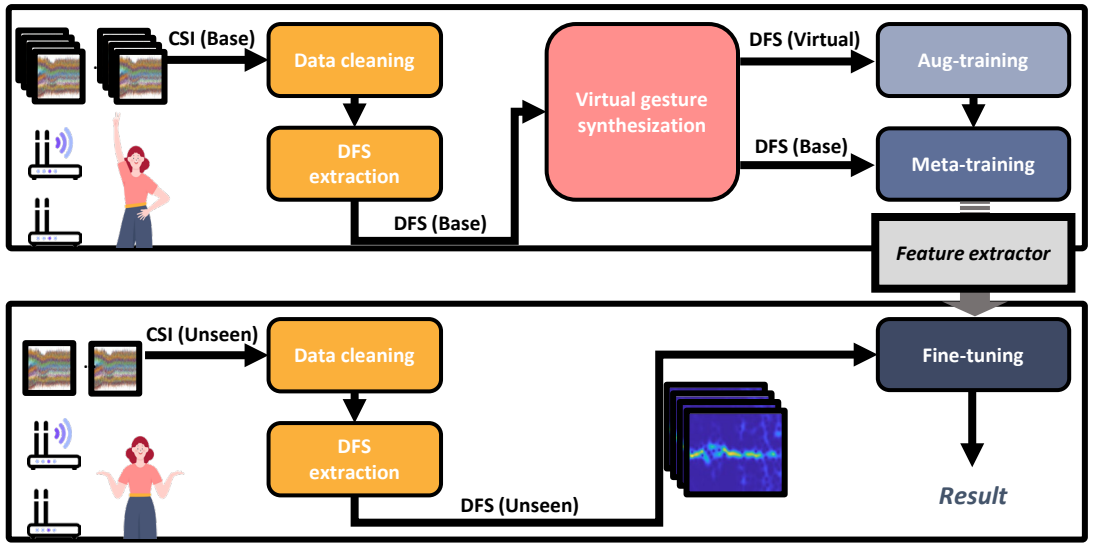
  ➡ **High training overhead (C2)**

- Fix the number of ways

  ➡ **Inflexibility for evolving class numbers (C3)**

Our Solution: *OneSense*

A Novel **One-shot WiFi-based** Gesture Recognition System that solves these **three challenges**



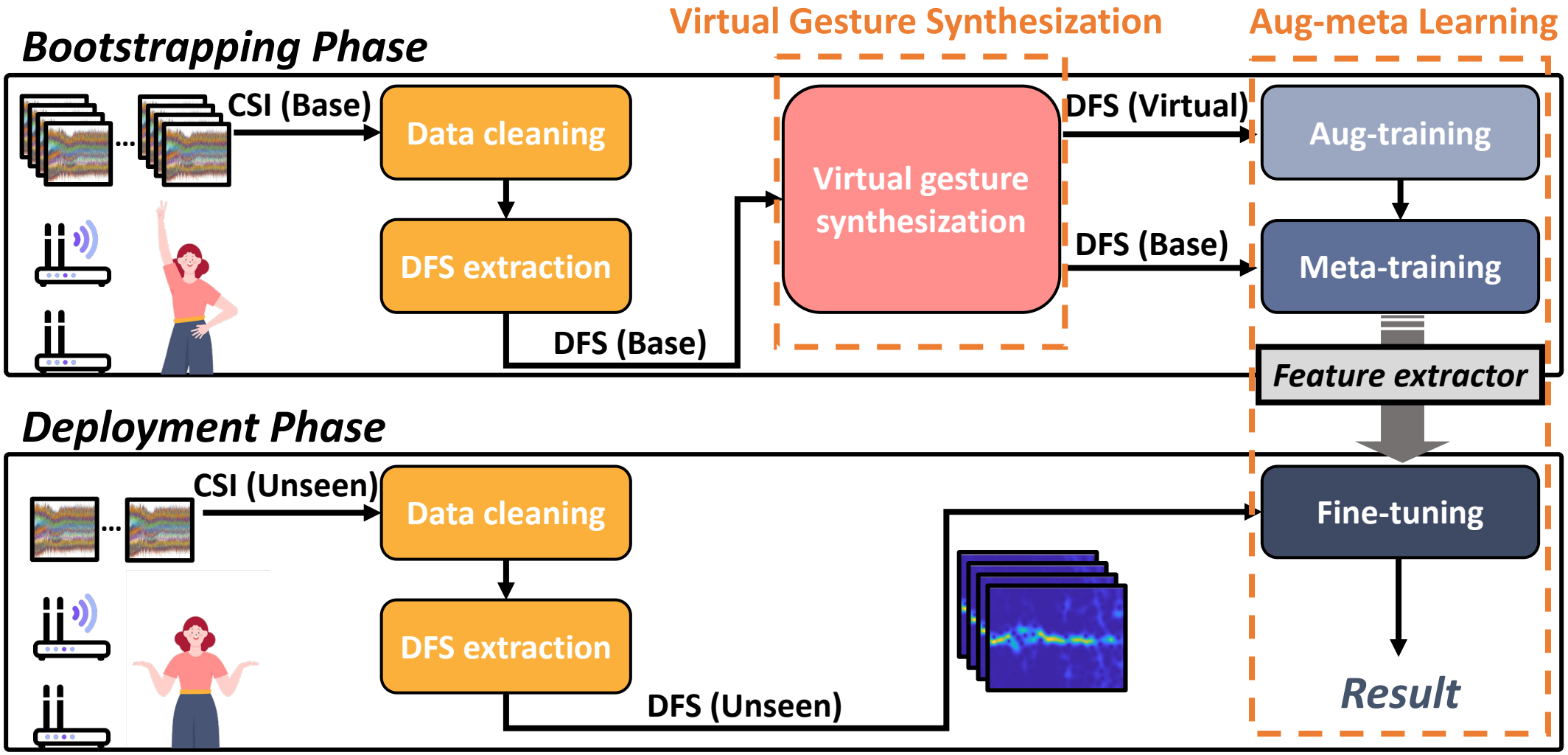
- Labor-intensive data collection (C1)
- High training overhead (C2)
- Inflexibility for evolving gesture classes (C3)

Small-scale real-world data collection

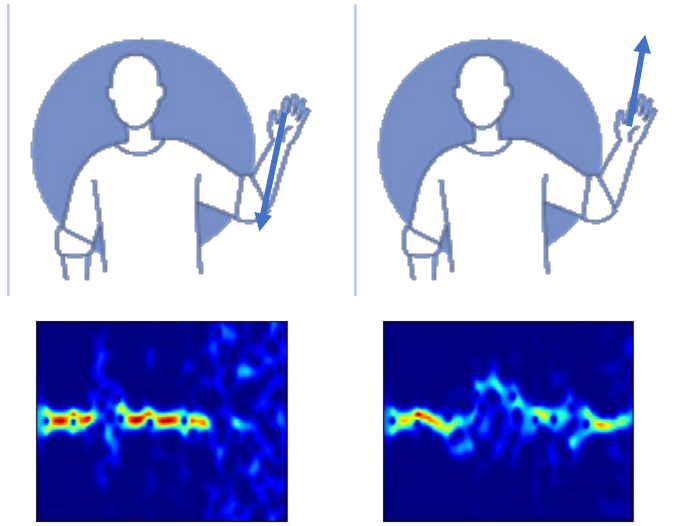
Low training overhead

Adaptability to various class numbers

System Design of *OneSense*

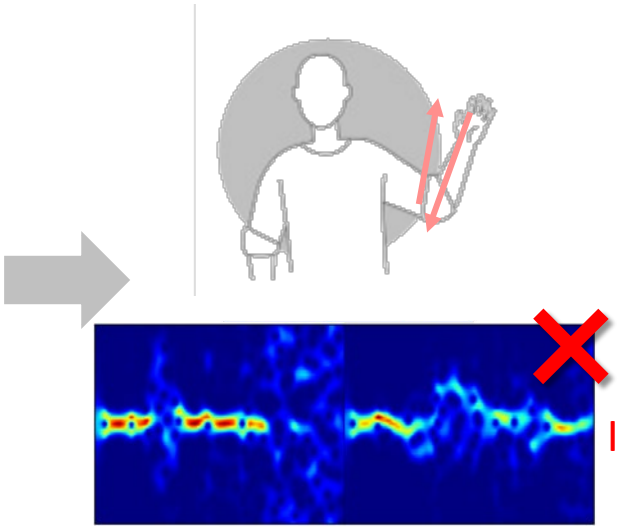


Virtual Gesture Synthesization



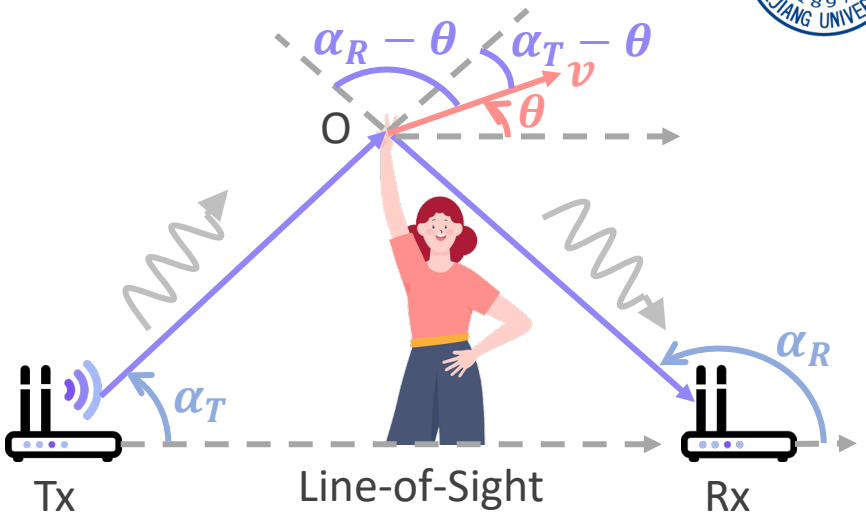
$S_A(f,t)$
$t \in [0,T]$

$S_B(f,t)$,
$t \in [0,T]$



Inconsistent sample shape

Step 1: Acceleration



Step 2: Combination

$$S_{A+B}(f,t) = \begin{cases} \left(\quad\right), 0 \le t < \frac{T}{2} \\ \left(\quad - T\right), \frac{T}{2} \le t \le T \end{cases}$$

Violate physical laws

Doppler frequency shift

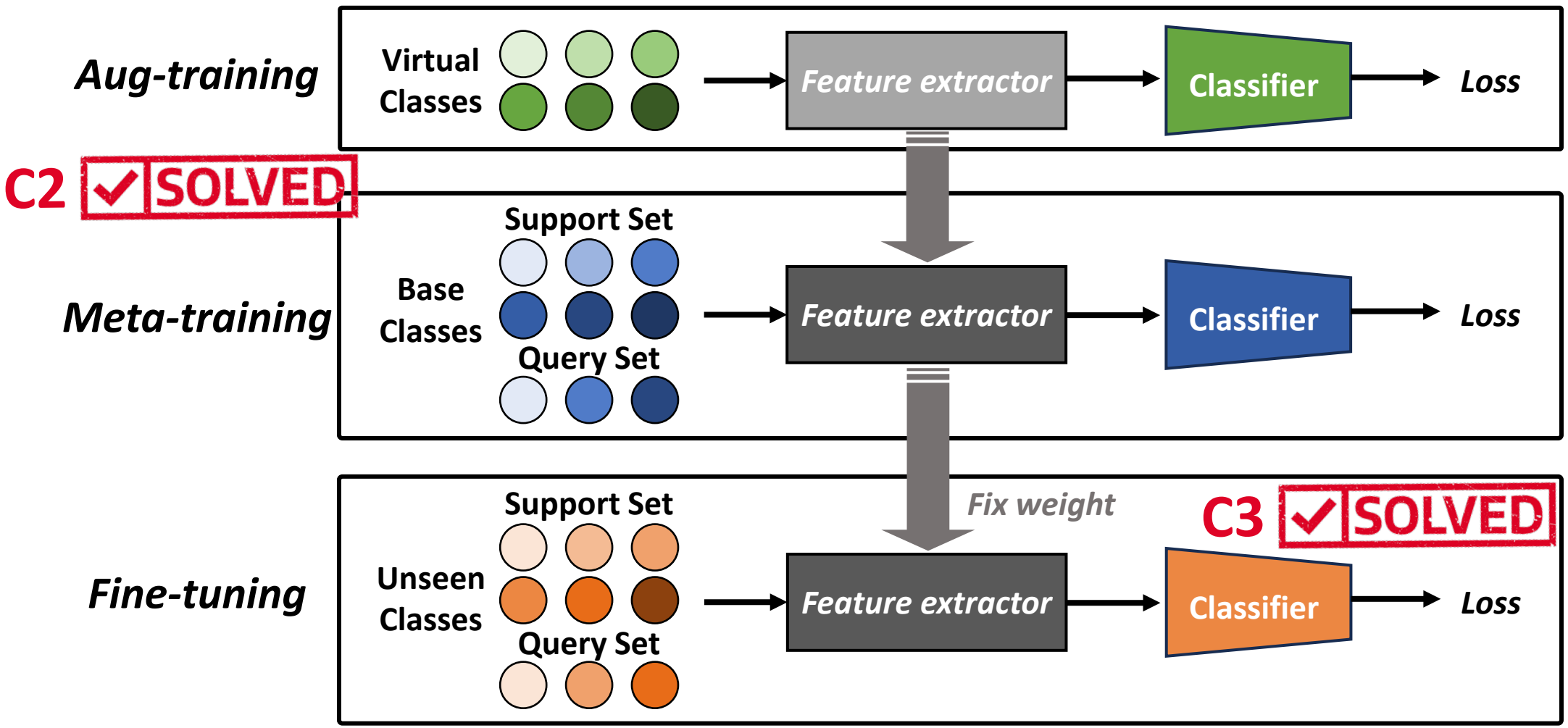
$$f_{D_{acc}}(t) = 2f_D(2t)$$

Doppler spectrogram

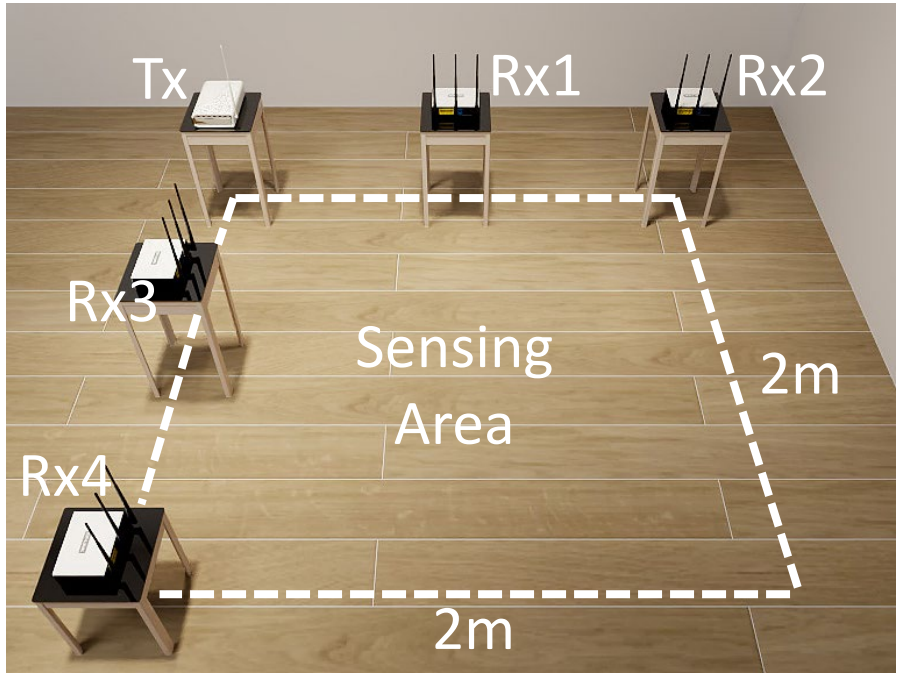
$$S_{acc}(f,t) = S\left(\frac{f}{2}, 2t\right)$$

C1 ✓ SOLVED

Aug-meta Learning (AML)



Prototype and Settings



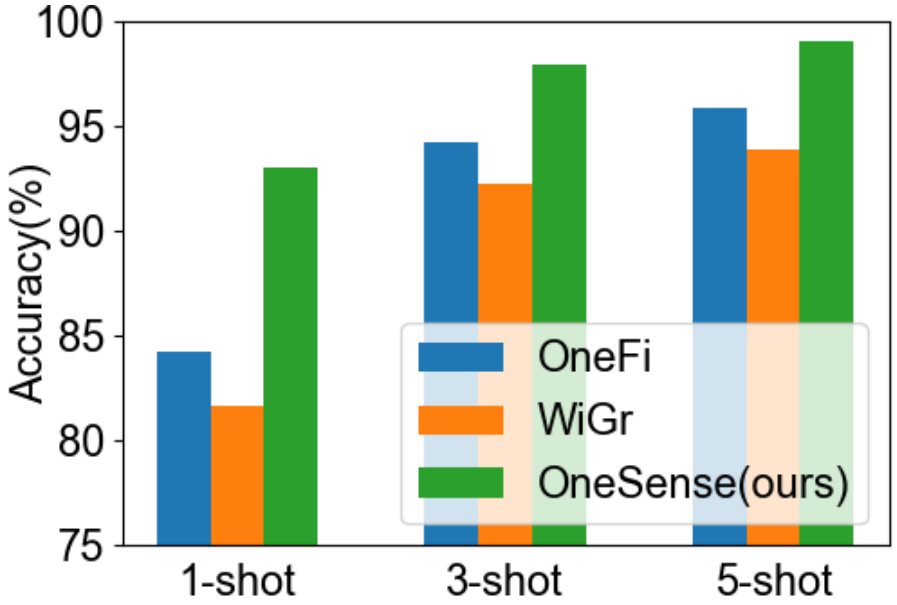
Laptops + Intel 5300 NIC

- Tx (one antenna) ×1
  1000 packets/s
- Rx (three antennas) ×4

40 gesture classes

- 20 **base** gestures
  (draw ‘1’, draw ‘2’, draw ‘3’, …)
- 20 **unseen** gestures
  (push & pull, sweep, slide, …)

Overall Accuracy



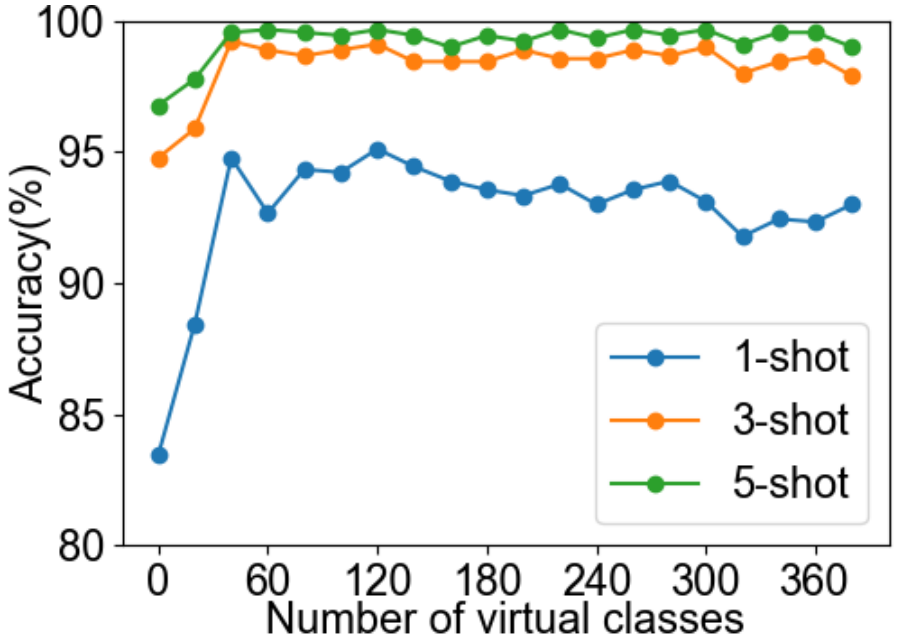
Baselines:

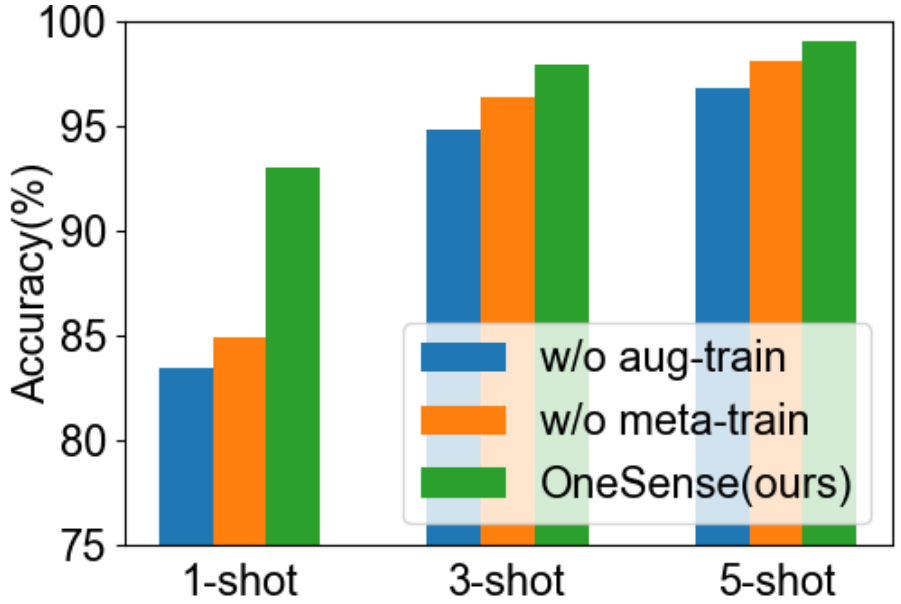
- *OneFi* (Sensys'21)
  data augmentation + transfer learning
- *WiGr* (IoTJ'22)
  modified prototypical network

Test on 6 unseen classes
(push & pull, sweep, slide, clap, draw zig-zag, draw triangle)

OneSense achieves **93% one-shot** recognition accuracy,
outperforming *OneFi* and *WiGr* in **one/few-shot** cases

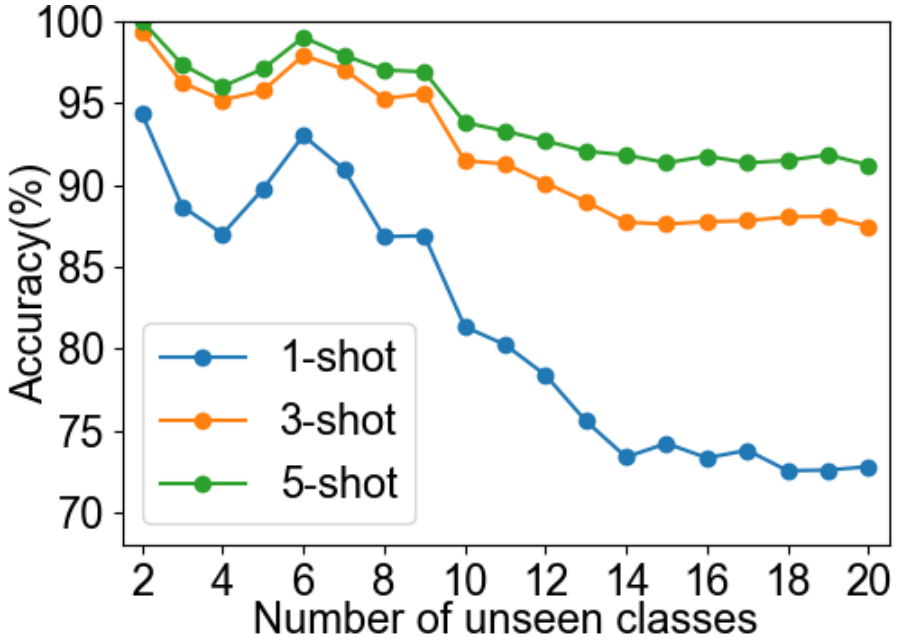
Effect of *OneSense* Modules





Virtual gesture synthesization and **aug-meta learning** framework are beneficial for improving few-shot recognition performance

Performance for Evolving Gestures



The **one-shot** recognition accuracy remains **85%+** for **2~9 unseen** classes

The **five-shot** recognition accuracy remains **91%+** for **2~20 unseen** classes

After **once-and-for-all training**,
OneSense adapts well to **evolving gestures** with **low-cost fine-tuning**

Summary of Other Evaluation Results

- AML framework **reduces the pre-training latency** by **86.1%+** compared to the classic meta-learning algorithm MAML.
- OneSense performs well with only **small number** of **base classes** and **receivers**, showing ability in **resource-intensive** cases.
- OneSense shows robustness in **cross-domain** scenarios.

Conclusion

- We propose ***OneSense,*** a novel **one-shot WGR** system. Extensive experiments demonstrate its high accuracy and scalability in recognizing **unseen** gestures.

- We design a **virtual gesture synthesization** technique, significantly reducing real-world data collection overhead.

- We propose **aug-meta learning**, a novel one-shot learning framework to enable efficient and scalable few-shot recognition.

浙江大學
ZHEJIANG UNIVERSITY

Thank you!