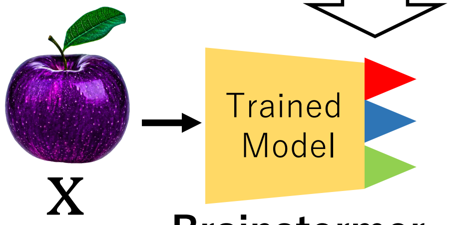
Collaborative Multi-Task Learning across Internet Edges with Device-to-Device Communications

Available Supervise Labels (y, z, w) for x

Input (x)	Person	Object Class (y)	Color (z)	Size (w)
	Person A	Apple	N/A	N/A
	Person B	N/A	Purple	N/A
	Person C	N/A	N/A	Big

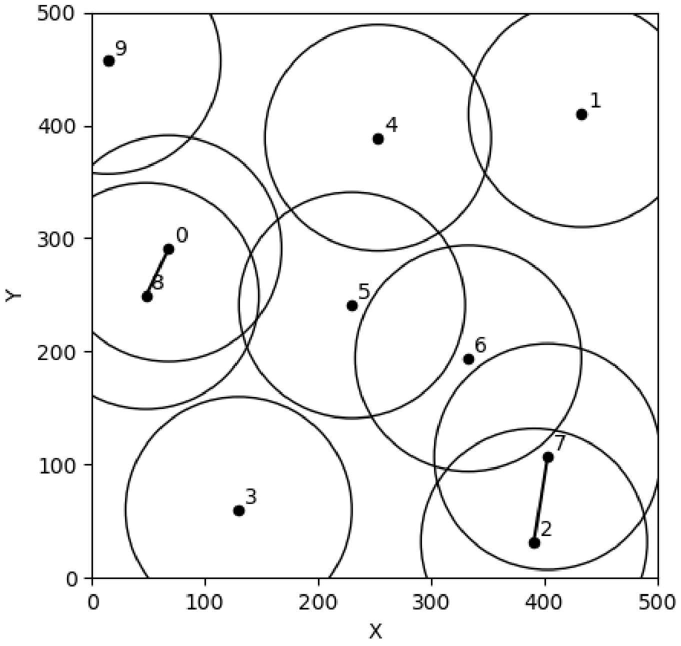
Collaborative Model Training



Brainstormer

(Apple, Purple, Big)
y z w

Device-to-Device Contacts for Communication



Ryusei Higuchi, Hiroshi Esaki, **Hideya Ochiai**

The University of Tokyo, Japan

Table of Contents

- Introduction
- Previous Studies
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Demonstration
- Conclusion

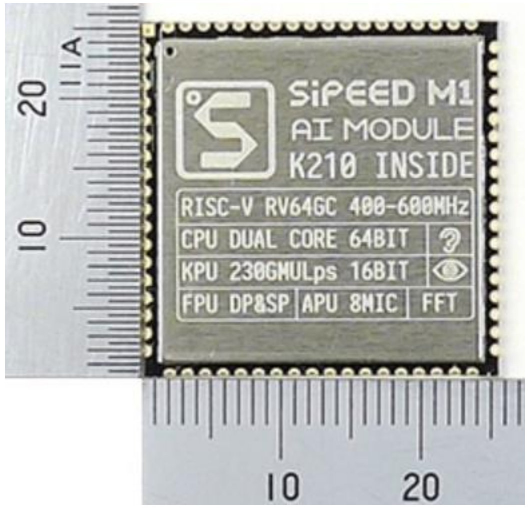
Table of Contents

- Introduction
- Previous Studies
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Demonstration
- Conclusion

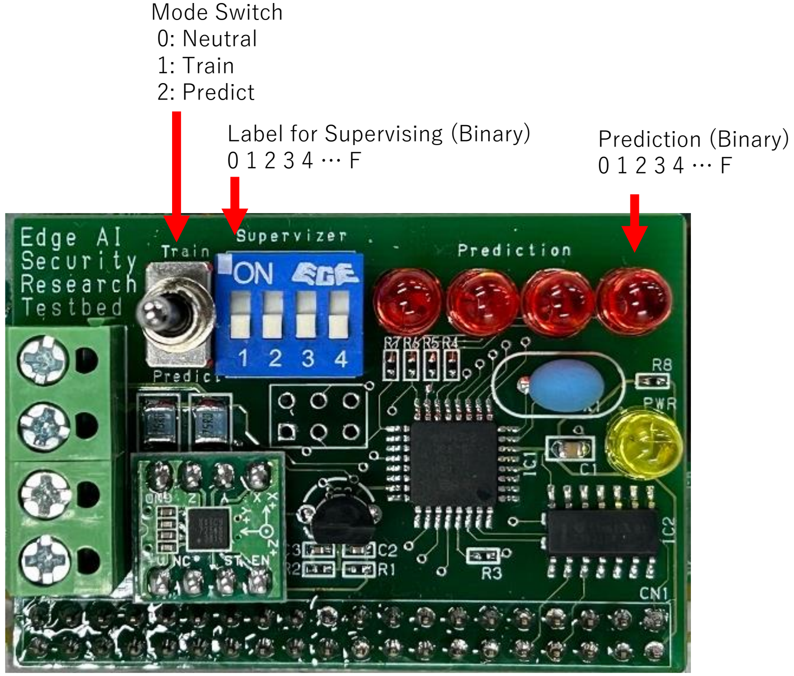
Toward On-Device Training

- Machine Learning has evolved with cloud computing.
- Nowadays :
  - AI Chips are available for 10 USD.
  - Machine Learning shifts onto IoT Edges.

- Issues:
  - Learning on an Edge may lead to model overfit.
- Approach of this issue:
  - Collaborative Training with Device-to-Device Communication.
- Focus of this research:
  - Training a Multi-Task Model
  - Using a Standard Benchmark Dataset



AI Chip (10USD)

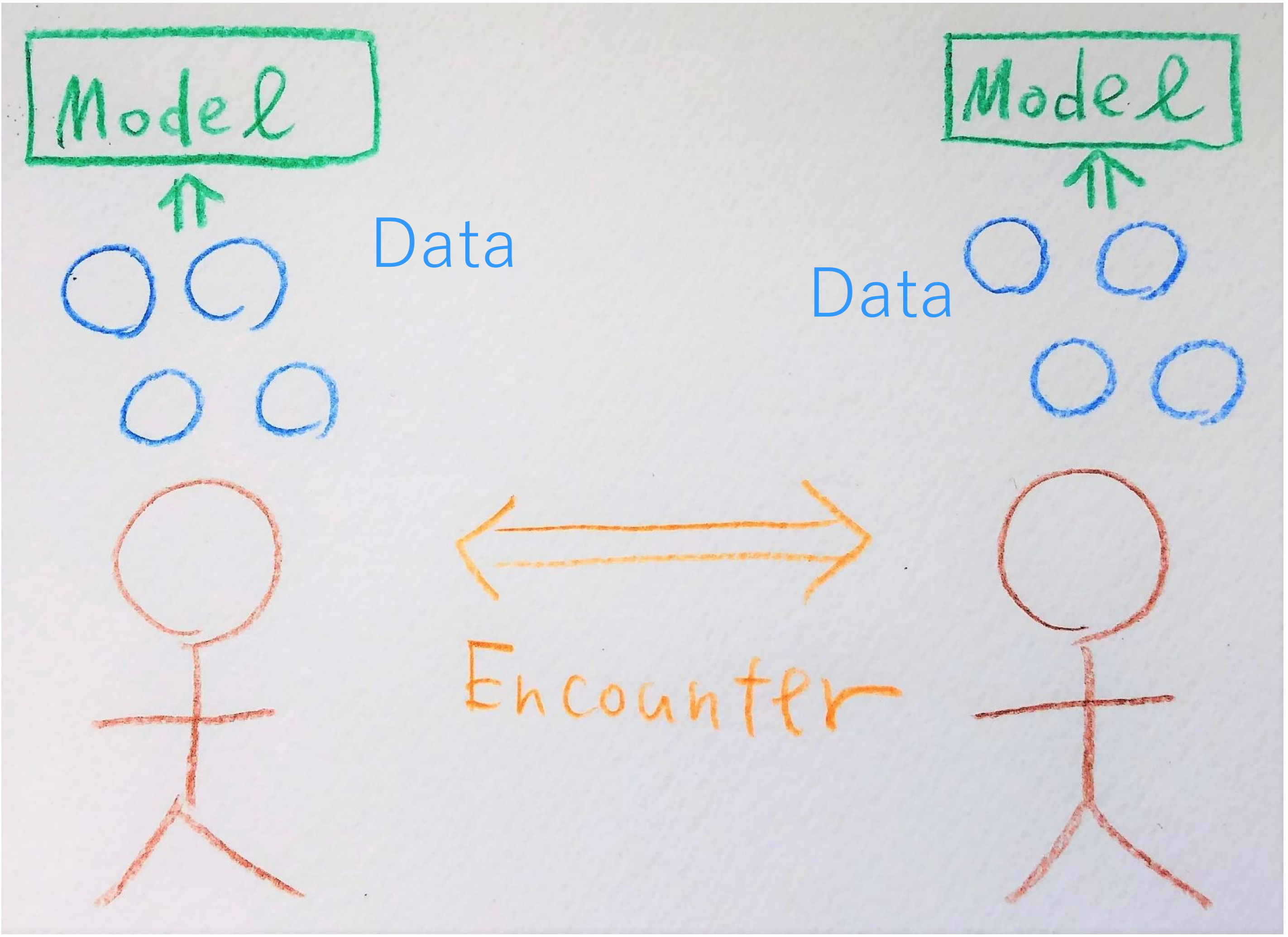


Edge AI Testbed

Table of Contents

- Introduction
- **Previous Studies**
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Demonstration
- Conclusion

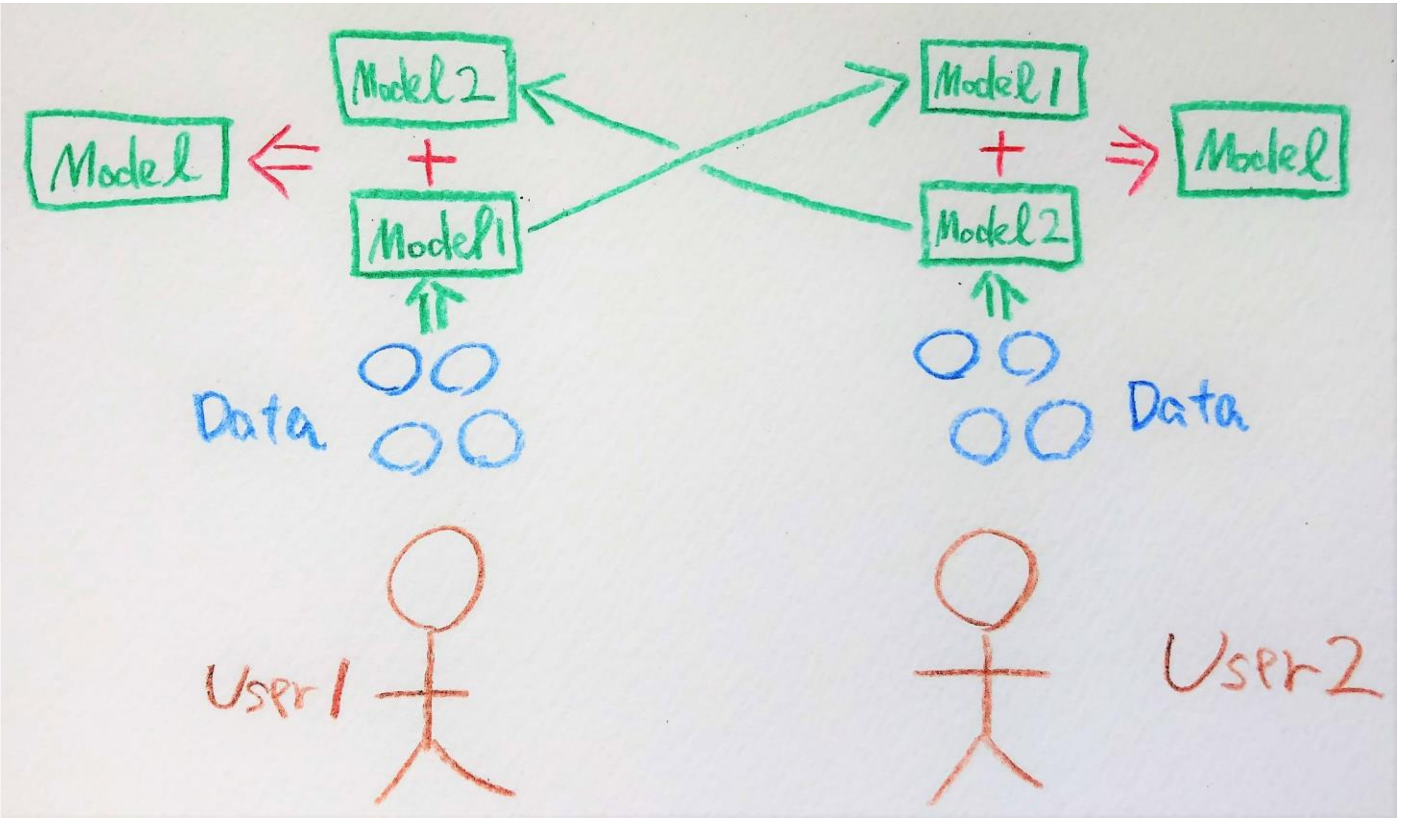
Previous Studies (1/5): Wireless Ad Hoc Federated Learning (WAFL) [1] with Device-to-Device Communication



1. Each node individually trains its ML model using its local data.
2. Each node encounters the other.
3. They can communicate with local wireless communication media such as <u>Wi-Fi Ad Hoc mode</u> or <u>Bluetooth</u>

[1] Ochiai, Hideya, et al. "Wireless ad hoc federated learning: A fully distributed cooperative machine learning." *arXiv preprint arXiv:2205.11779* (2022).

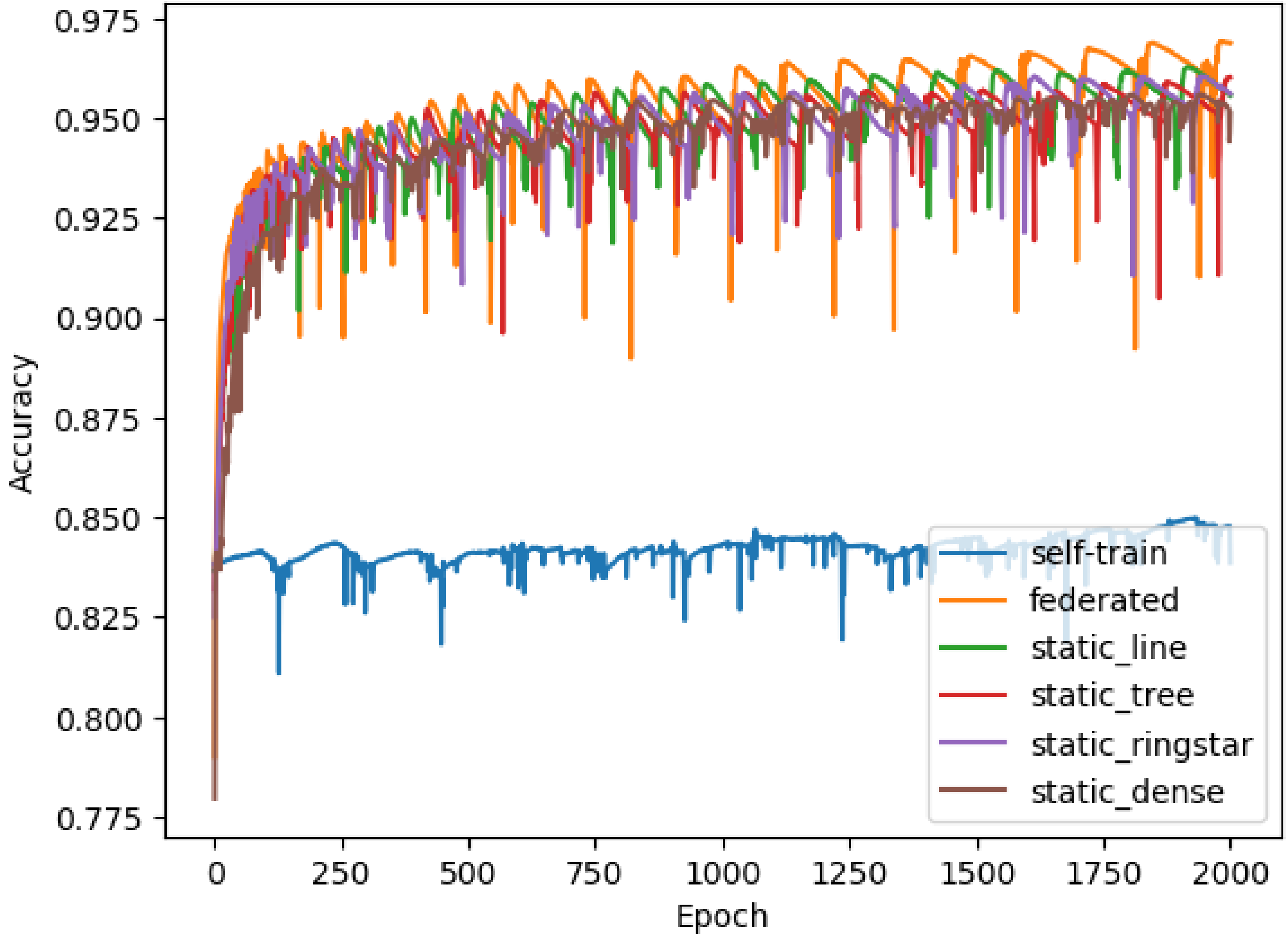
Previous Studies (2/5): Wireless Ad Hoc Federated Learning (WAFL) [1] with Device-to-Device Communication



1. Each node individually trains its ML model using its local data.
2. Each node encounters the other.
3. They can communicate with local wireless communication media such as Wi-Fi Ad Hoc mode or Bluetooth
4. They exchange and aggregate the models to develop a new model.
5. This enables collaborative training.

[1] Ochiai, Hideya, et al. "Wireless ad hoc federated learning: A fully distributed cooperative machine learning." *arXiv preprint arXiv:2205.11779* (2022).

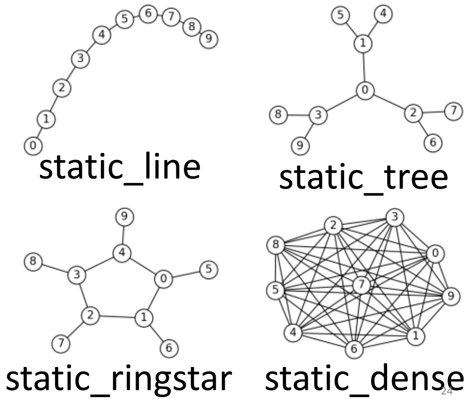
Previous Studies (3/5): Wireless Ad Hoc Federated Learning (WAFL) [1] with Device-to-Device Communication

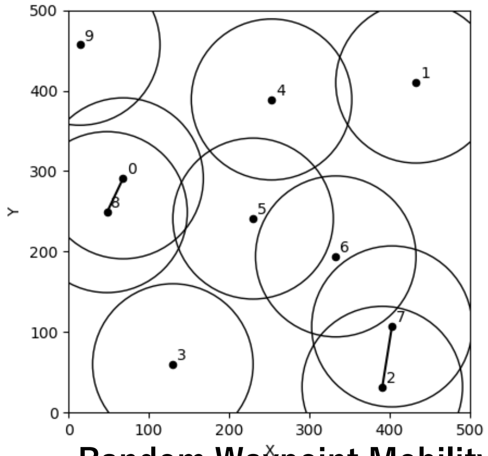


WAFL has achieved almost the same performance as conventional federated learning.

10.0-11.5% Improvement

Accuracy of Self-Train



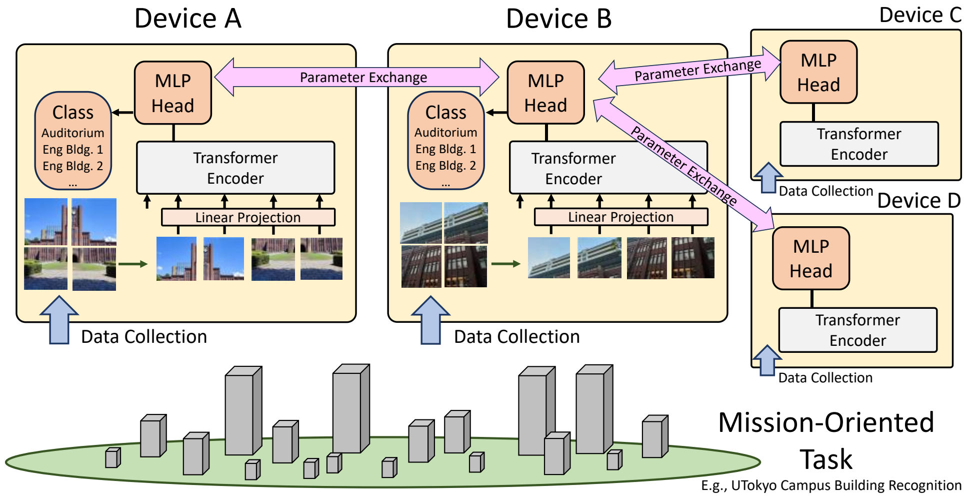


Random Waypoint Mobility

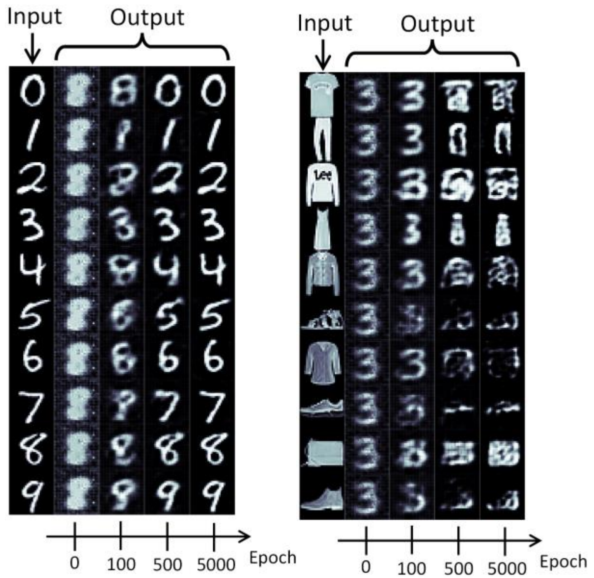
Ochiai, Hideya, et al. "Wireless ad hoc federated learning: A fully distributed cooperative machine learning." *arXiv preprint arXiv:2205.11779* (2022).

Previous Studies (4/5): Extensions and Variations of WAFL (1/2)

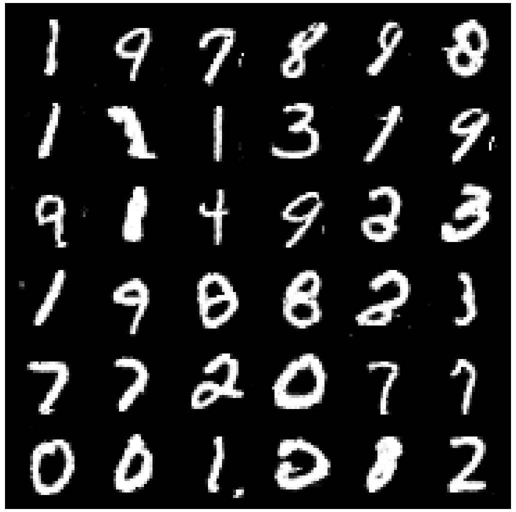
WAFL-Vision Transformer [1]



WAFL-Autoencoder[2]



WAFL-GAN[3]



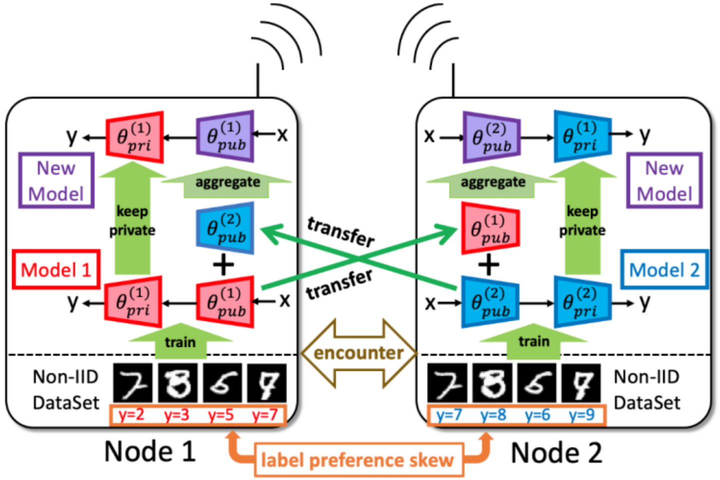
[1] Hideya Ochiai, Atsuya Muramatsu, Yudai Ueda, Ryuhei Yamaguchi, Kazuhiro Katoh, and Hiroshi Esaki, "Tuning Vision Transformer with Device-to-Device Communication for Targeted Image Recognition", IEEE World Forum on Internet of Things, 2023 (**Best Paper Award**).

[2] Hideya Ochiai, Riku Nishihata, Eisuke Tomiyama, Yuwei Sun, and Hiroshi Esaki, "Detection of Global Anomalies on Distributed IoT Edges with Device-to-Device Communication", ACM MobiHoc AIoT workshop, 2023.

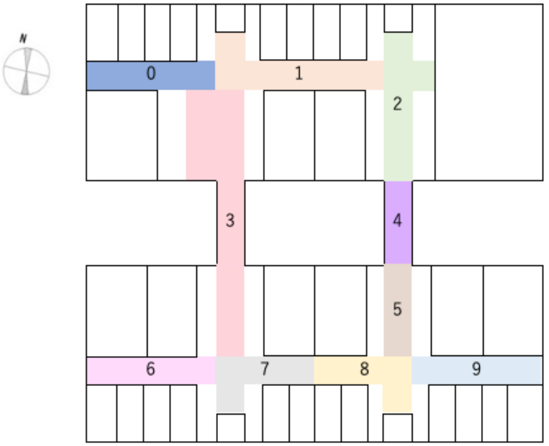
[3] Eisuke Tomiyama, Hiroshi Esaki, Hideya Ochiai, "WAFL-GAN: Wireless Ad Hoc Federated Learning for Distributed Generative Adversarial Networks", IEEE International Conference on Knowledge and Smart Technology, 2023.

Previous Studies (5/5): Extensions and Variations of WAFL (2/2)

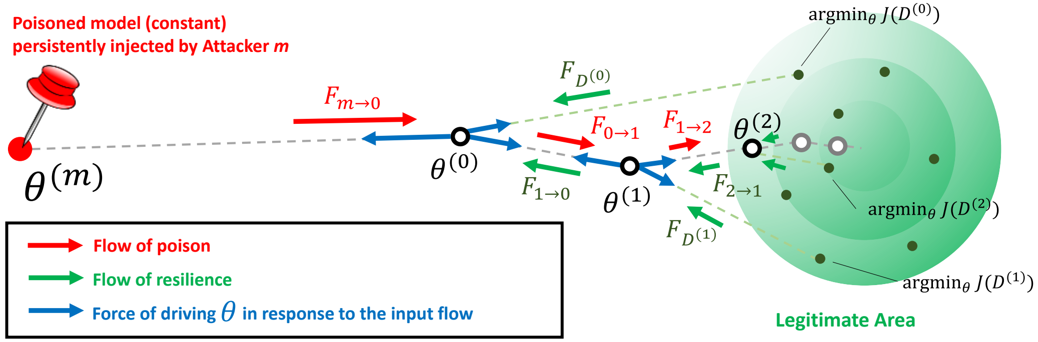
WAFL-Personalization[4]



WAFL-Localization[5]



WAFL-Security[6]



[4] Ryusei Higuchi, Hiroshi Esaki, and Hideya Ochiai, "Personalized Wireless Ad Hoc Federated Learning for Label Preference Skew", IEEE World Forum on Internet of Things, 2023.

[5] Yusuke Sugizaki, Hideya Ochiai, Muhammad Asad, Manabu Tsukada, and Hiroshi Esaki, "Wireless Ad-Hoc Federated Learning for Cooperative Map Creation and Localization Models", IEEE World Forum on Internet of Things, 2023.

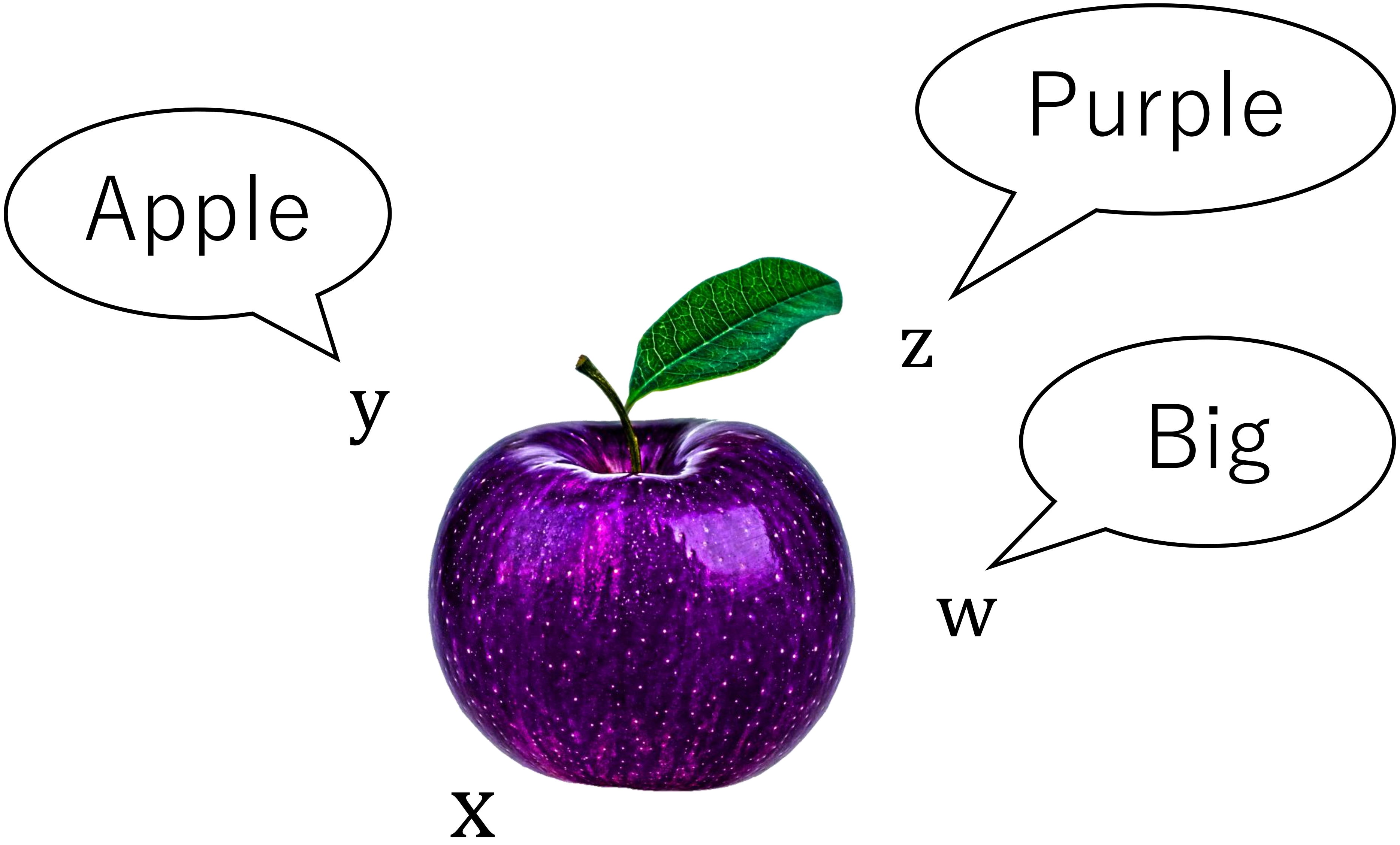
[6] Naoya Tezuka, Hideya Ochiai, Yuwei Sun, Hiroshi Esaki, "Resilience of Wireless Ad Hoc Federated Learning against Model Poisoning Attacks", IEEE International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA), 2022.

We propose MT-WAFL and demonstrate the effectiveness in training a Brainstormer.

Table of Contents

- Introduction
- Previous Studies
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Demonstration
- Conclusion

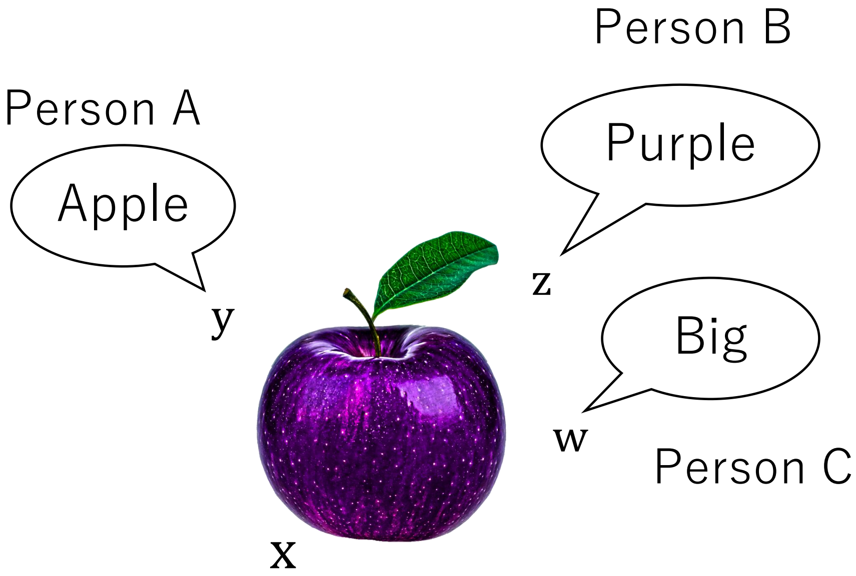
Multiple Cognitive Axes



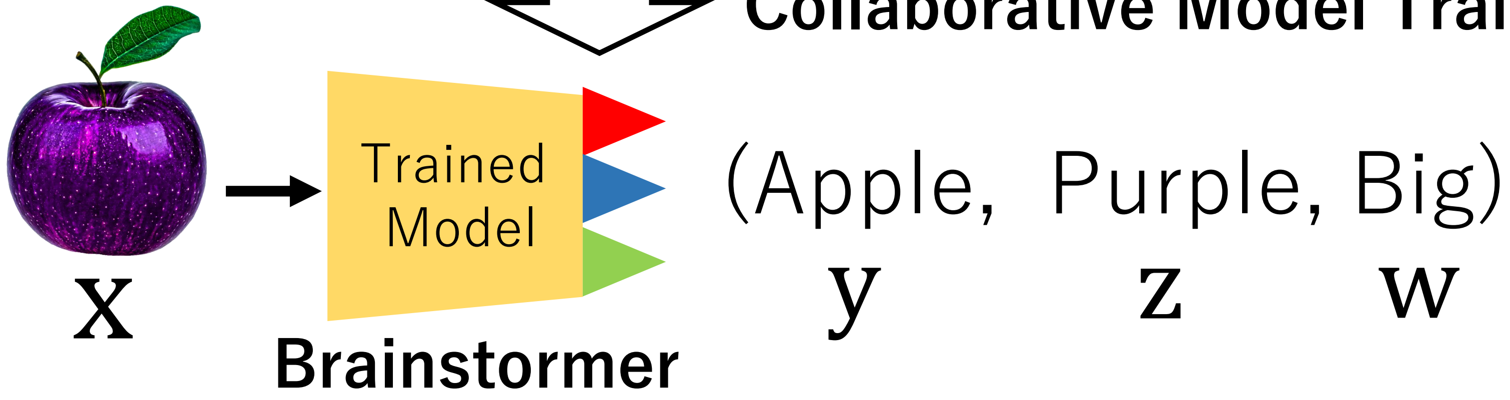
Training Brainstormer from Weekly-Labeled Distributed Data on Edge Devices

Available Supervise Labels (y, z, w) for x

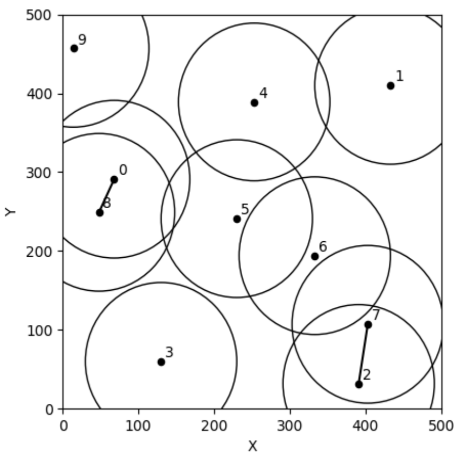
Input (x)	Person	Object Class (y)	Color (z)	Size (w)
	Person A	Apple	N/A	N/A
	Person B	N/A	Purple	N/A
	Person C	N/A	N/A	Big



Collaborative Model Training



Multi-Task Learning Model Integration with Wireless Ad Hoc Federated Learning



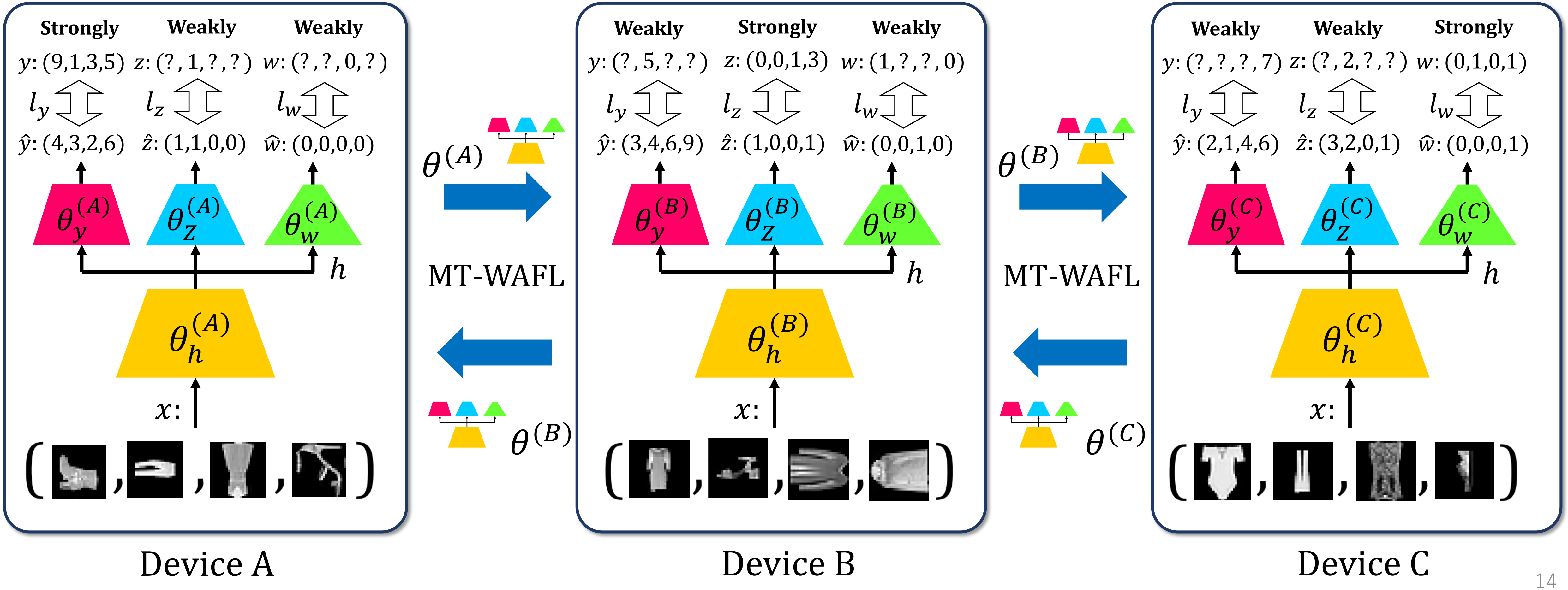
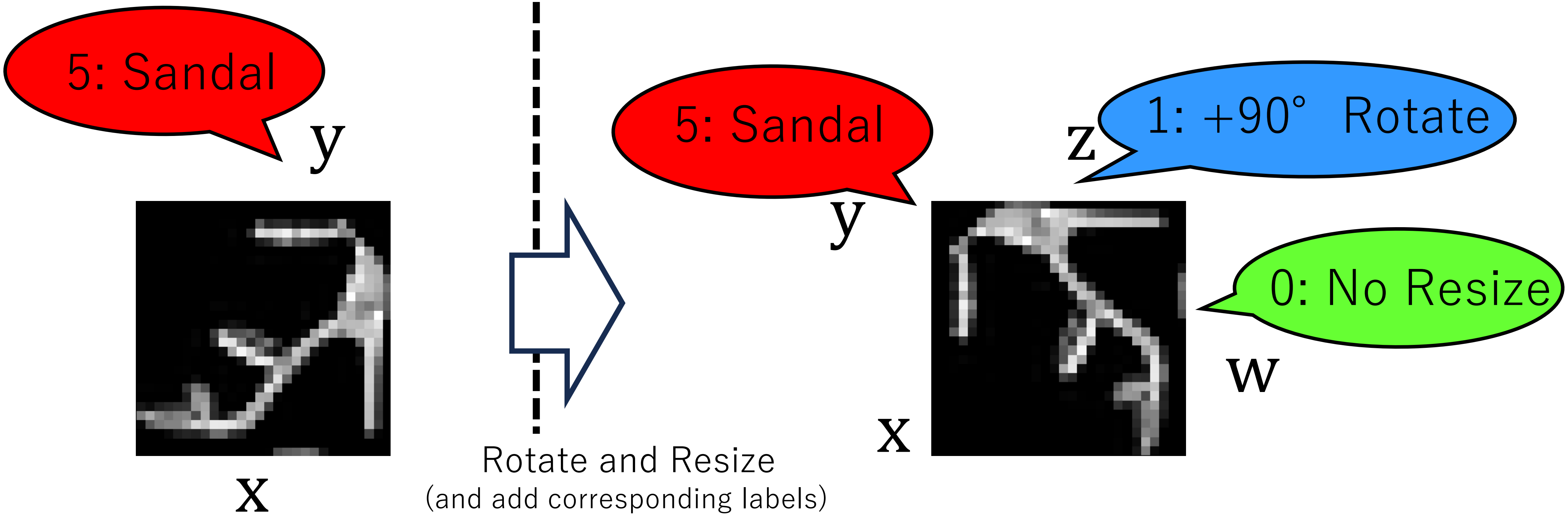


Table of Contents

- Introduction
- Previous Studies
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Demonstration
- Conclusion

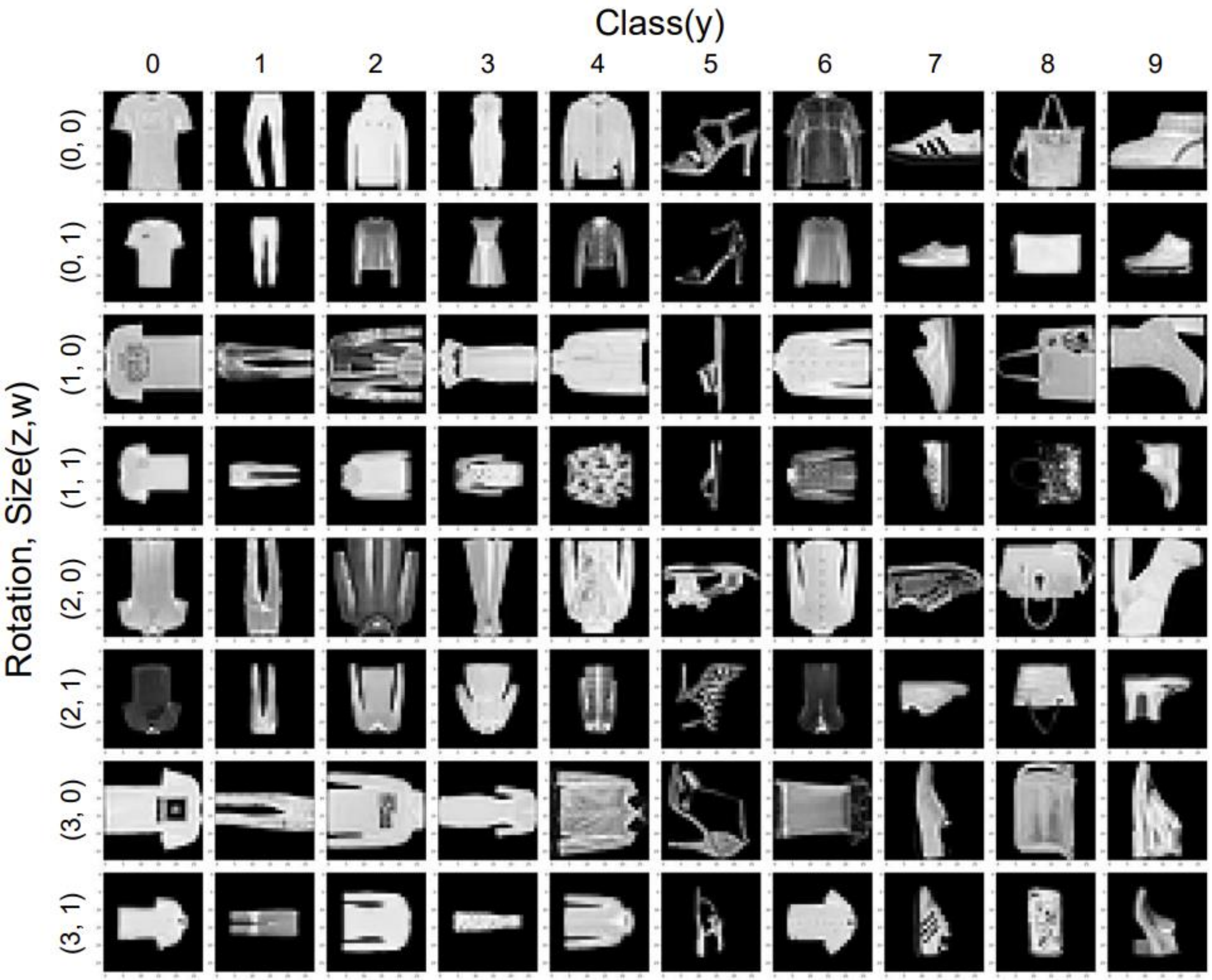
Extension of Fashion-MNIST for Multi-Task Learning

Original Fashion-MNIST | Rotated and Resized Fashion-MNIST



Single Axis Dataset | Multiple Axes Dataset

Examples of the Expanded Fashion-MNIST

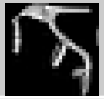


After the rotation and resizing,

1. We split them into train and test data.
2. We distributed the train data to 10 mobile devices.
3. We removed some labels in the training data on the device to reveal their major cognitive axis.

Major Cognitive Axis Skew in our study

Available Supervise Labels (y, z, w) for x

Input (x)	Person	Object Class (y)	Rotation (z)	Size (w)
	Device 0-4	5	N/A	N/A
	Device 5-7	N/A	1	N/A
	Device 8-9	N/A	N/A	1

Distribution of Availability

Availability of Supervised Labels

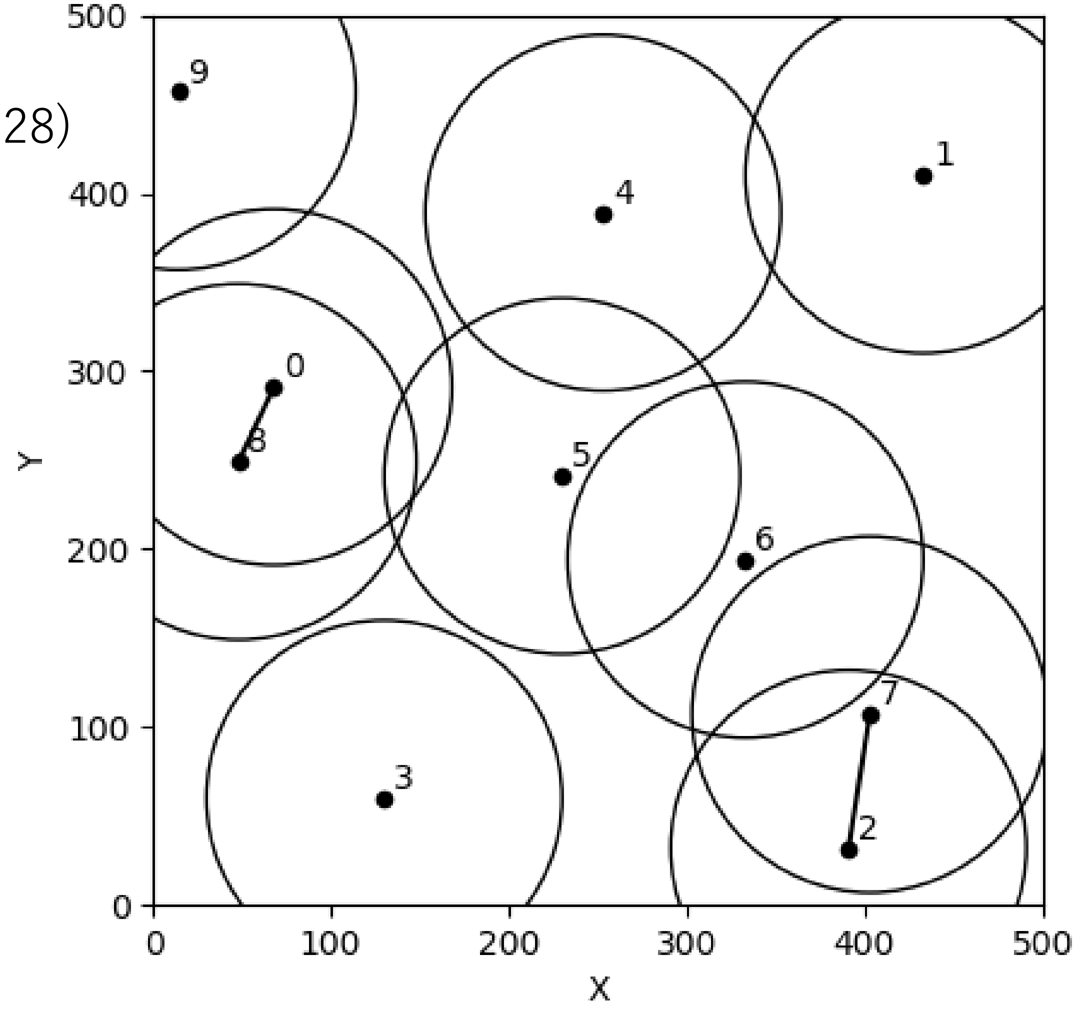
	Y	Z	W
Device 0-4	99%	1%	1%
Device 5-7	1%	99%	1%
Device 8-9	1%	1%	99%

Table of Contents

- Introduction
- Previous Studies
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Demonstration
- Conclusion

Demonstration: Experiment Settings

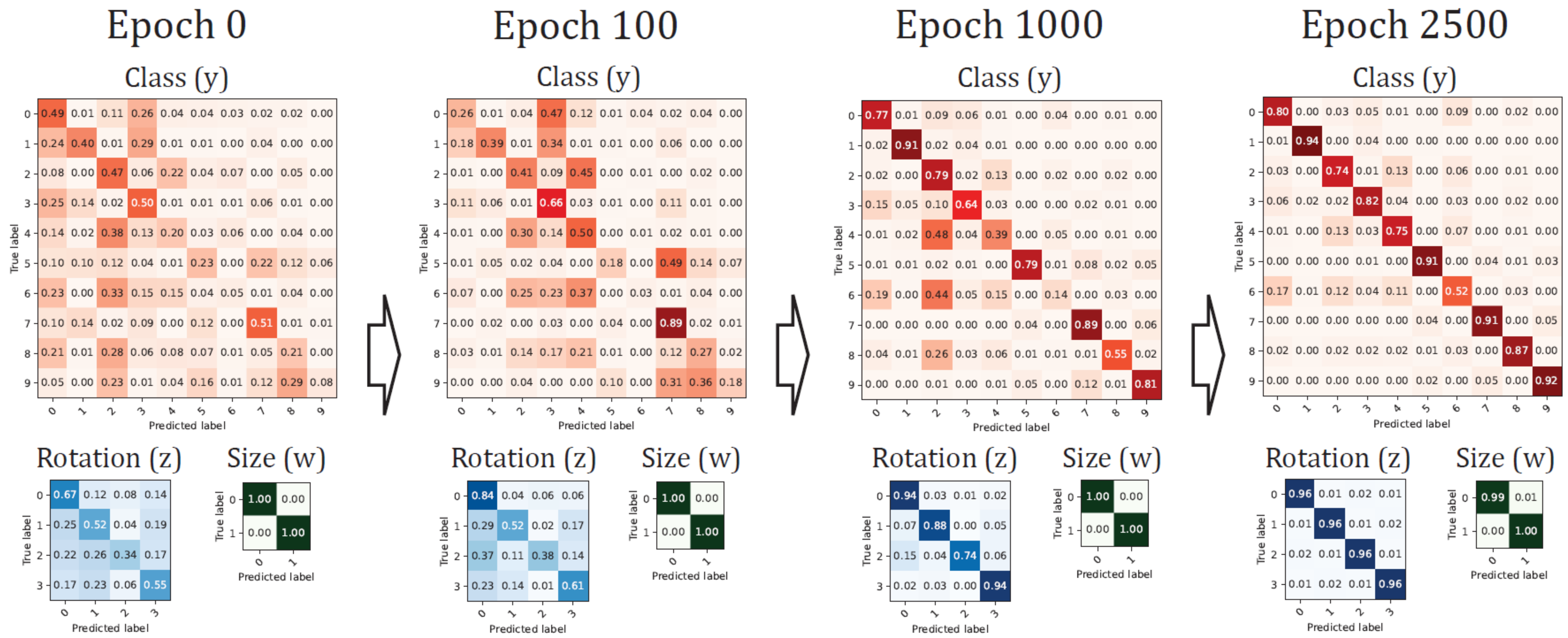
- ML model (MLP-based)
  - θh : FC ($i = 784$, $o = 256$) - ReLU - FC ($i = 256$, $o = 128$)
  - θy : FC ($i = 128$, $o = 64$) - ReLU - FC ($i = 64$, $o = 10$)
  - θz : FC ($i = 128$, $o = 32$) - ReLU - FC ($i = 32$, $o = 4$)
  - θw : FC ($i = 128$, $o = 32$) - ReLU - FC ($i = 32$, $o = 2$)
- Dataset
  - Expanded Fashion-MNIST
- Mobility Pattern
  - Random Waypoint Mobility (RWP)
- Simulation
  - We carried out the experiment by simulation on a single computer.



Random Waypoint Mobility

Integration of Multiple Perspectives by WAFL

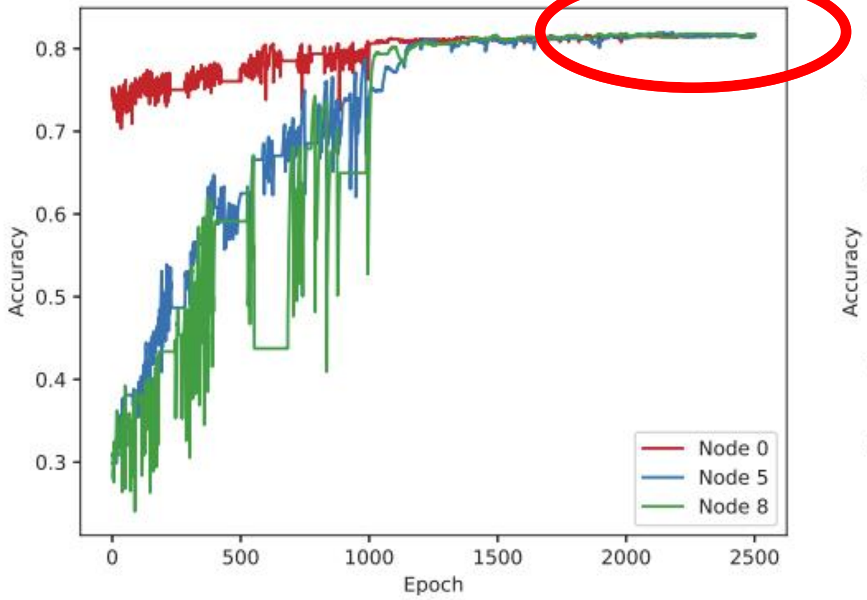
Device 8's Predictions for Object Class, Rotation, and Size



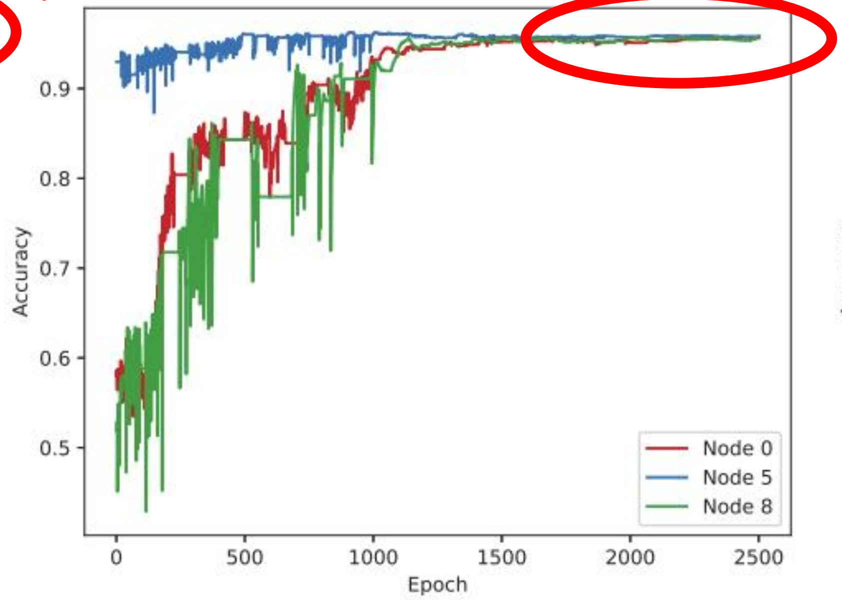
Device 8 originally has many labels in Size(w) perspective.
As the training proceeds, misclassifications in Object Class and Rotation predictions have improved.

Accuracy of Class, Rotation, and Size @Device (0, 5, 8)

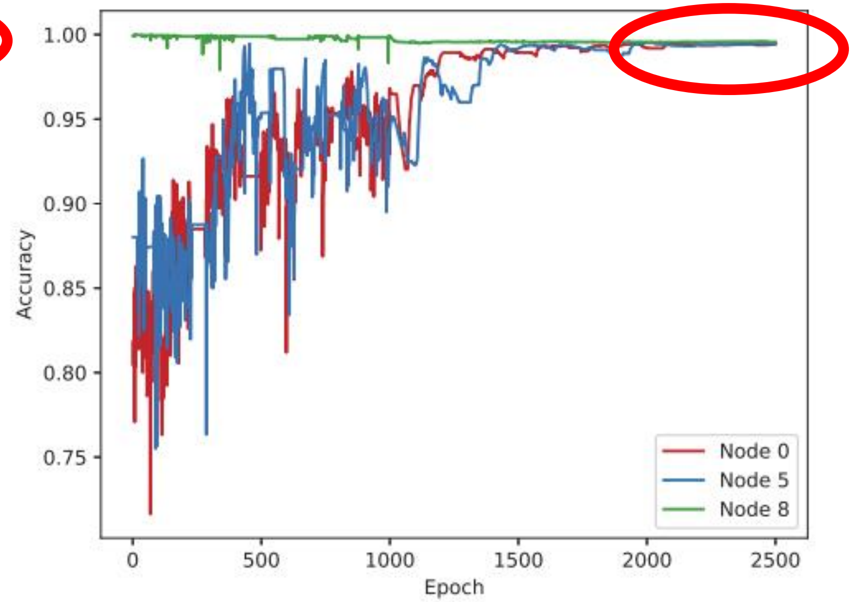
MT-WAFL (our proposal) achieved identical accuracies



(a) Test Accuracy for Class (Y)



(b) Test Accuracy for Rotation (Z)



(c) Test Accuracy for Size (W)

Device 0 has Class (Y) labels.

Device 5 has rotation (Z) labels.

Device 8 has size (W) labels.

Accuracy has been improved in all the perspectives.

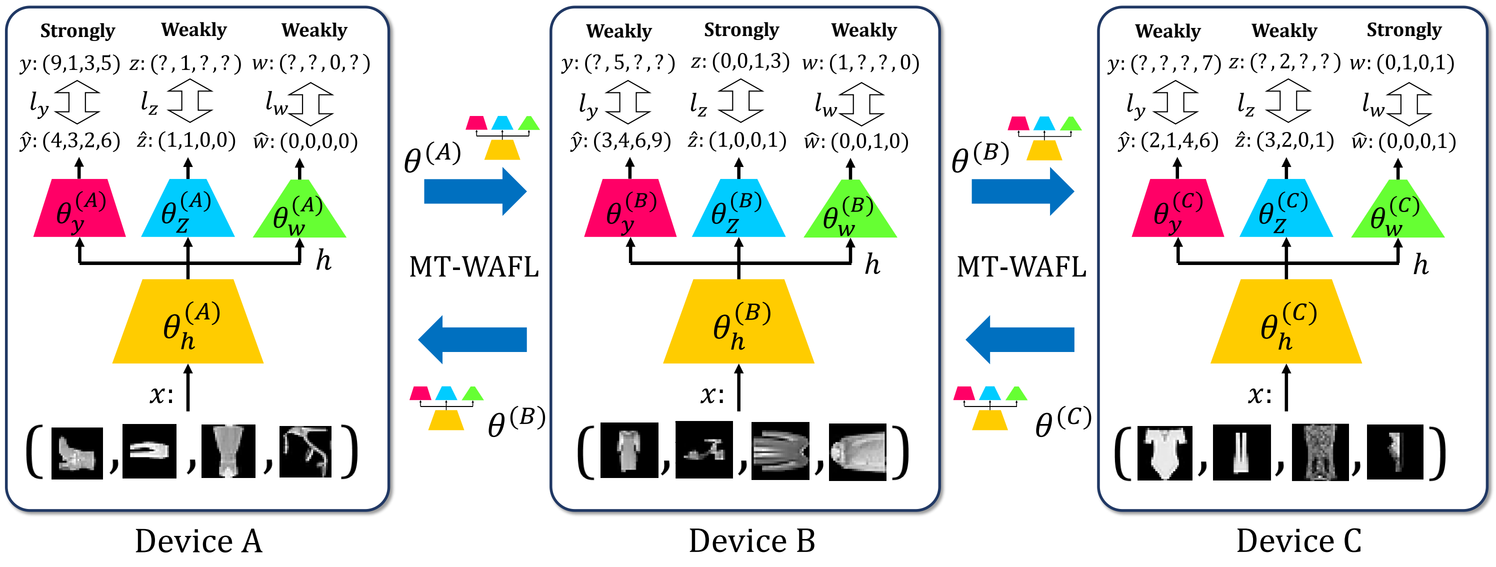
Table of Contents

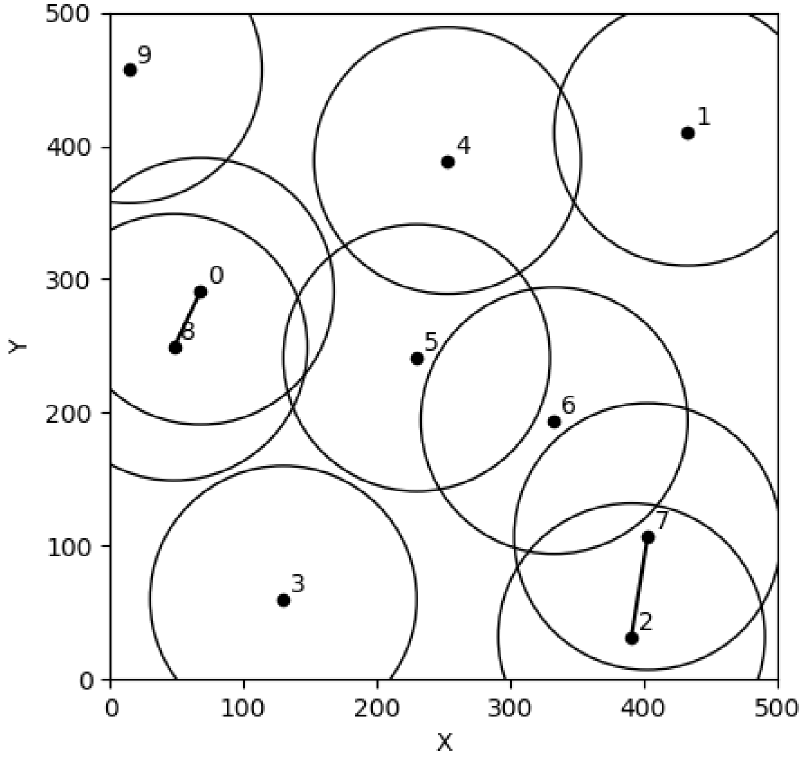
- Introduction
- Previous Studies
- Multi-Task Wireless Ad Hoc Federated Learning
- Rotated and Resized Fashion-MNIST Dataset for MTL
- Evaluation
- Conclusion

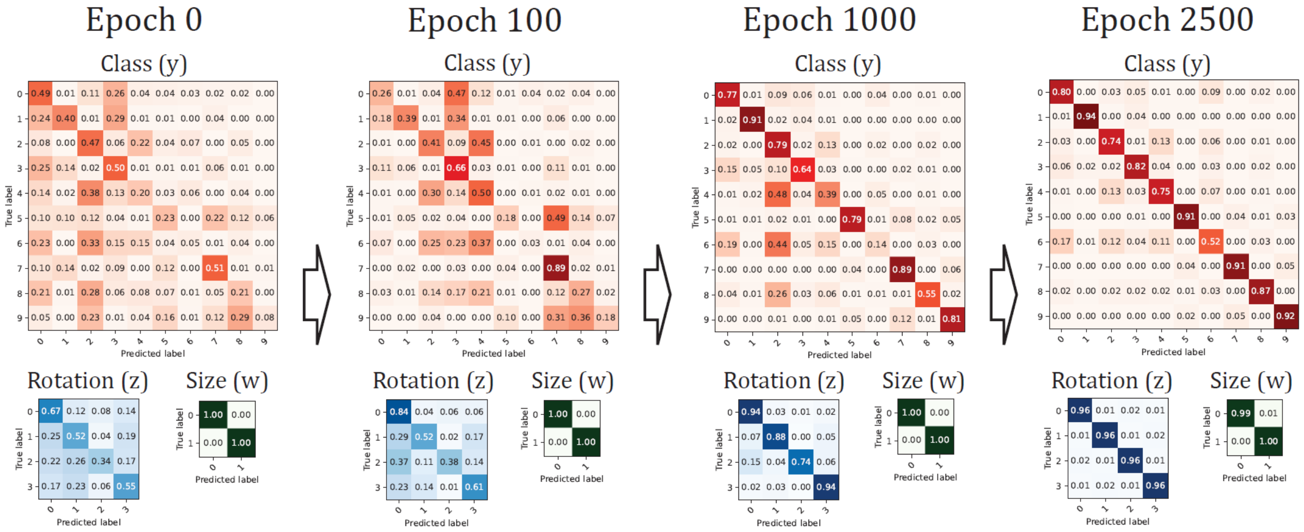
Conclusion

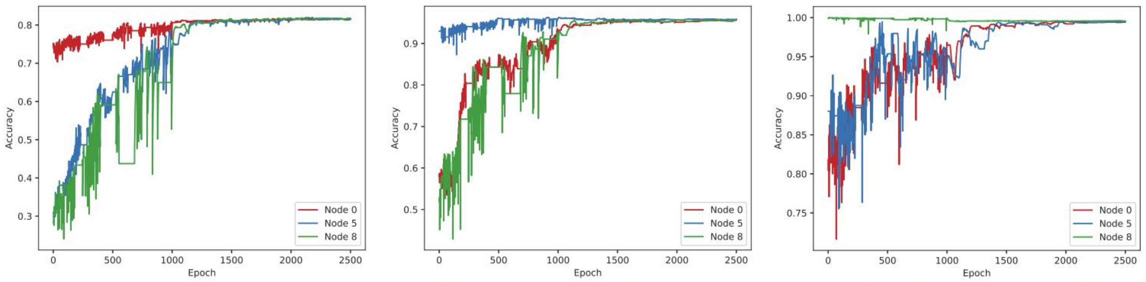
- Multi-Task WAFL (MT-WAFL) has successfully integrated multi-task models with other devices through D2D communication.
- We demonstrated a brainstormer trained by MT-WAFL in the case of our expanded Fashion-MNIST dataset.
- Further applications could be considered using the realistic photos. The trained model will help the people of the community.

Thank you.

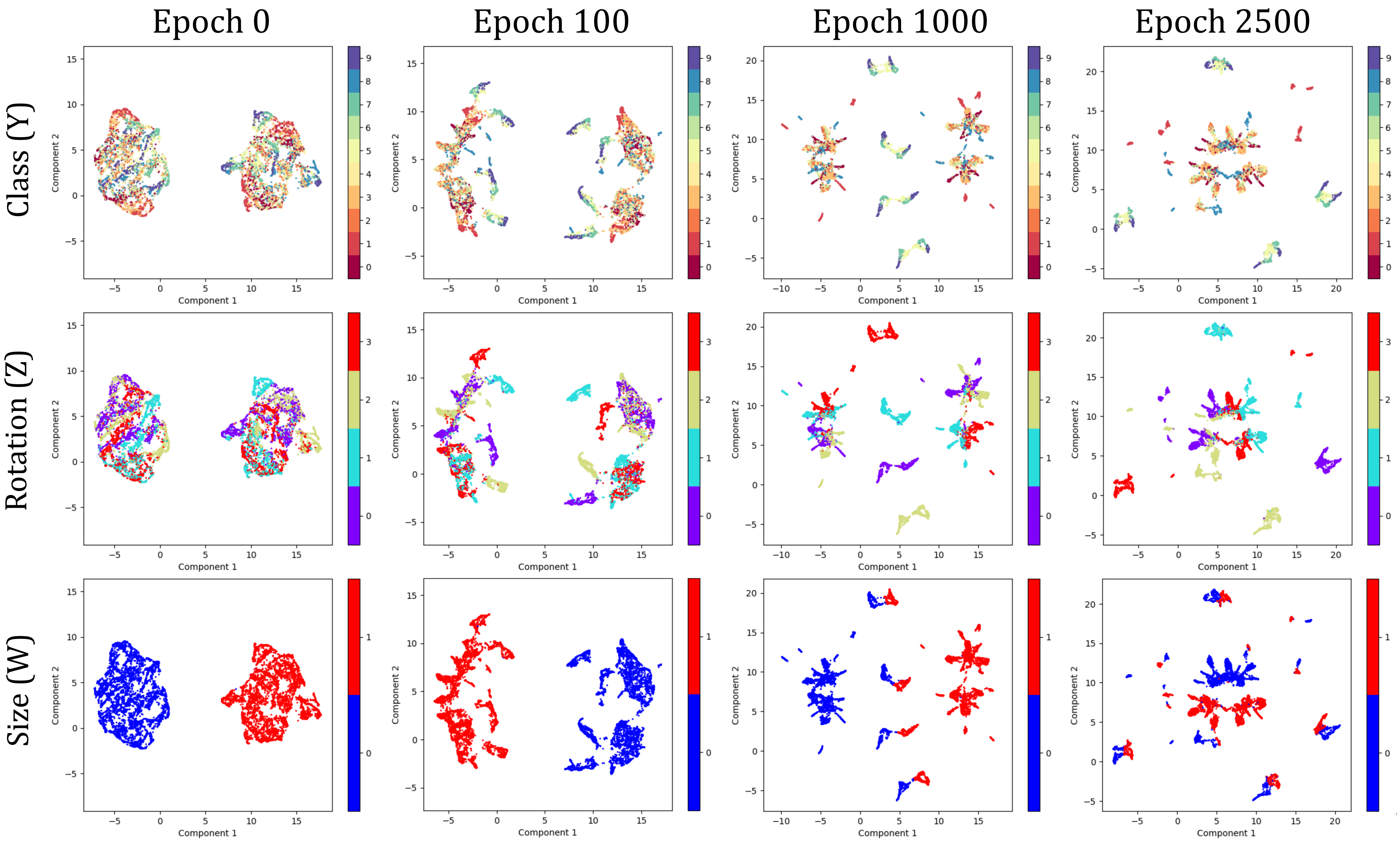








(a) Test Accuracy for Class (Y) (b) Test Accuracy for Rotation (Z) (c) Test Accuracy for Size (W)



Epoch 0
Epoch 100
Epoch 1000
Epoch 2500
Class (Y)
Rotation (Z)
Size (W)
Component 1
Component 2