

# Many-Task Federated Fine-Tuning via Unified Task Vectors

Vasileios Tsouvalas · Tanir Ozcelebi · Nirvana Meratnia

Eindhoven University of Technology

FedGenAI @ IJCAI 2025

## 01 Motivation

### The problem

Many-task federated learning (FL) must share knowledge without forcing dissimilar tasks into one update. Separate per-task models increase storage and communication, while clients may hold several tasks simultaneously.

### Contributions

- **Compact representation:** combine task vectors with lightweight masks and scalar rescalers.
- **Related-task transfer:** aggregate within tasks and across tasks with aligned update signs.
- **Stateless server:** avoid retaining client-specific model weights between rounds.

### Beyond one task per client

Clients can hold multiple overlapping tasks, each with its own data. This task heterogeneity creates interference within clients as well as across clients.

### A shared reference point

Task vectors measure changes from the same pretrained weights. Conflicting update signs indicate potential interference; agreement provides a data-free signal of task compatibility.

### Scope

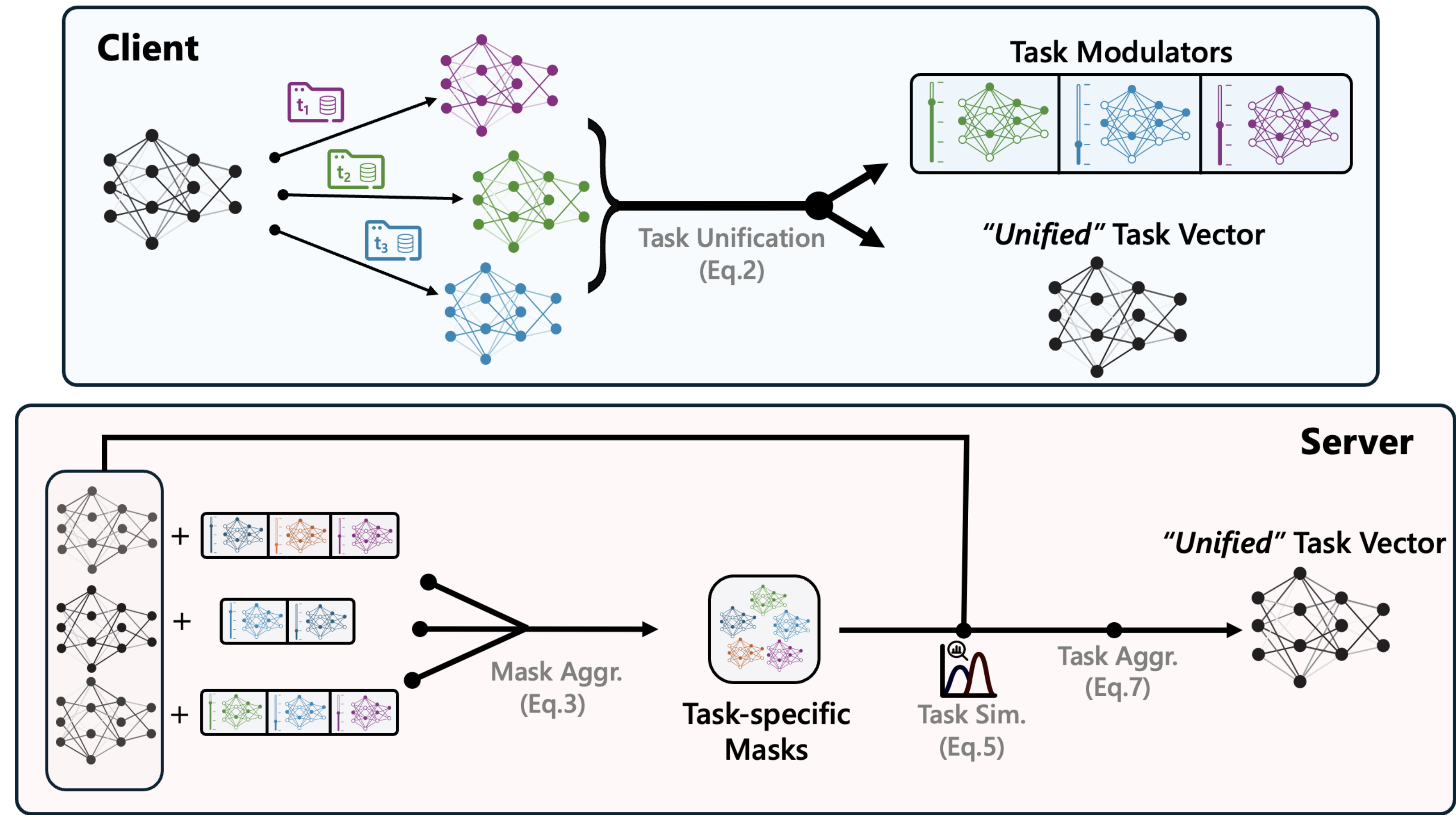
Eight-task comparisons use unequal round budgets; repeated-run uncertainty is unreported. Evidence is limited to vision classification with vision transformer (ViT)-B/32 and low-rank adaptation (LoRA).

### Notation

megabits (Mbit)

## 02 MaTU

One vector per client, distinct updates per task



### Technical core

$$\tau = \sigma \odot \mu$$

$\tau$ : unified task vector;  $\sigma$ : sign of the summed task vectors;  $\mu$ : largest magnitude agreeing with that sign, entry by entry.

$$\dot{\tau}_t = \lambda_t m_t \odot \tau, \quad \theta_t = \theta_p + \dot{\tau}_t$$

Mask  $m_t$  keeps task-aligned entries; scalar  $\lambda_t$  rescales them. Add the recovered task vector to pretrained weights  $\theta_p$ .



Paper



Code

## 03 Evaluation

79.47%

average task accuracy

8 tasks; multiple tasks per client.

8.04 Mbit/task

communication per round

MaT-FL: 20.43 Mbit/task/round.

### Multiple tasks per client: accuracy and communication

| Method                 | Avg. accuracy (%) | Mbit/task/round |
|------------------------|-------------------|-----------------|
| Individual centralized | 90.21             | —               |
| FedAvg                 | 51.63             | 20.43           |
| FedPer                 | 48.37             | 17.16           |
| MaT-FL                 | 67.60             | 20.43           |
| <b>MaTU (Ours)</b>     | <b>79.47</b>      | <b>8.04</b>     |

8 tasks; ViT-B/32, LoRA rank 16. 30 clients, 1 local epoch, task concentration  $\zeta_t=0.5$ , 20% participation. 100 rounds; MaT-FL: 300. No uncertainty reported.

### Scaling to the 30-task benchmark

| Method             | Normalized accuracy (%) |
|--------------------|-------------------------|
| MaT-FL             | 52.62                   |
| <b>MaTU (Ours)</b> | <b>77.40</b>            |

30 tasks; accuracy normalized to individual task fine-tuning. ViT-B/32, 30 clients, 300 rounds, 1 local epoch, task concentration  $\zeta_t=0.2$ , full participation.

### Takeaway

**MaTU reaches 79.47% at 8.04 Mbit/task/round** across eight tasks. Its 30-task normalized accuracy is **77.40% versus MaT-FL's 52.62%**, below individual task fine-tuning.