Unsupervised Multi-Source Federated Domain Adaptation under Domain Diversity through Group-Wise Discrepancy Minimization

Larissa Reichart, Cem Ata Baykara, Ali Burak Ünal, Harlin Lee, Mete Akgün

Medical Data Privacy and Privacy-preserving Machine Learning (MDPPML), Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, Germany

School of Data Science and Society, University of North Carolina at Chapel Hill, NC, USA

Problem Statement

Unsupervised multi-source domain adaptation seeks to generalize to an unlabeled target domain using labeled data from many diverse sources, but current distributed methods struggle to scale beyond a few domains. As the number of **heterogeneous sources** grows, these approaches become impractical due to high computational costs and unstable performance.

Inter-Group Discrepancy (IGD)

- We align source feature extractor by minimizing their prediction discrepancies on target data
- Approximate the full pairwise discrepancies with a **group-wise** formulation for scalability:

$$\mathcal{L}_{IGD} = \mathbb{E}_{x\epsilon\mathbb{D}_T}\left[\left|F_{\mathcal{G}_1}\big(G(x)\big) - F_{\mathcal{G}_2}\big(G(x)\big)\right|\right]$$

- IGD reduces variance compared to approximation by random domain pair sampling applied in [1]

Domain Weighting (MDMGB+)

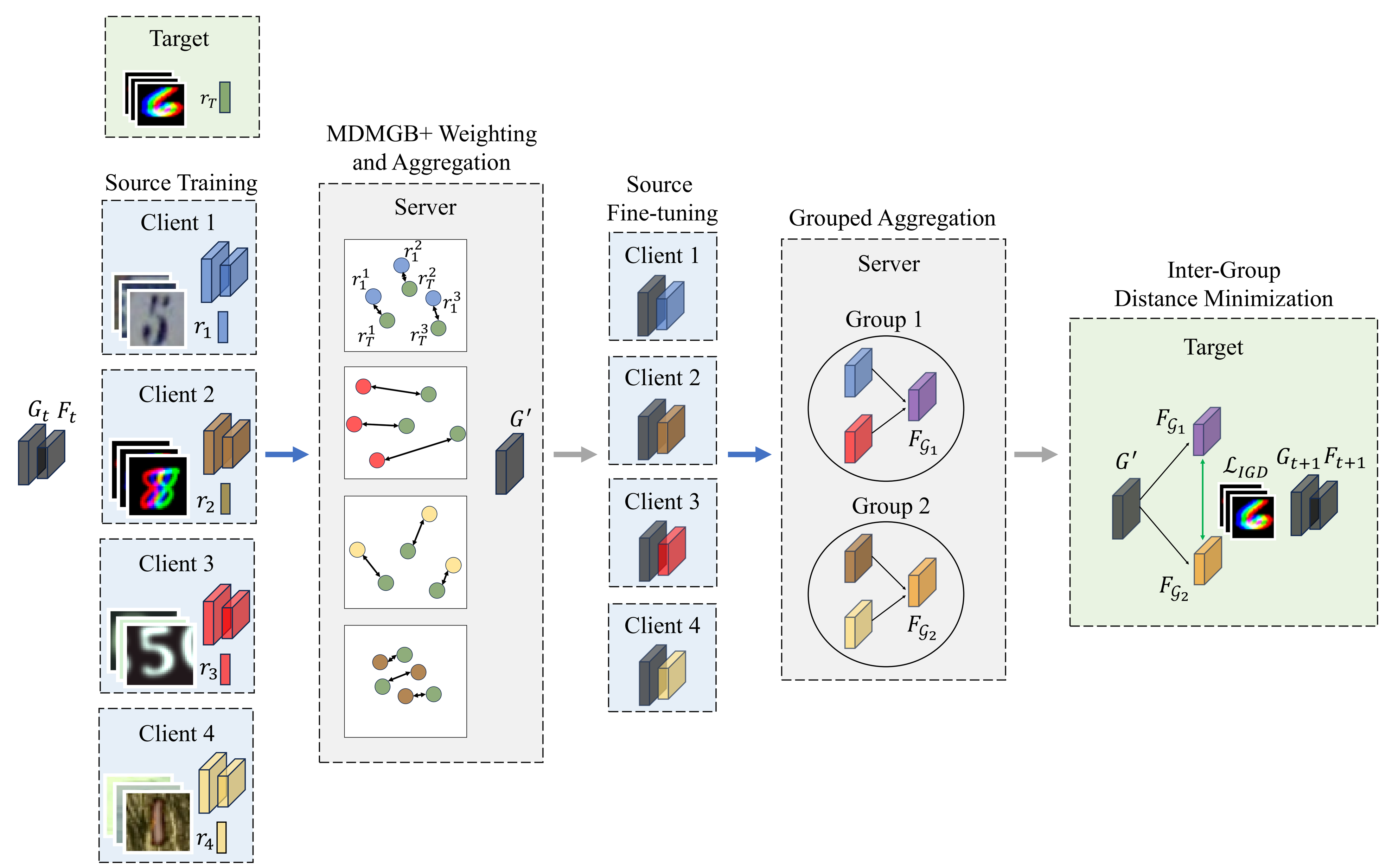
- We use centroid-based weighting of sources to identify target similar sources based on MDMGB [2]. Class-wise feature centroids:

$$r_n^c = \frac{\sum_{x\epsilon\mathbb{D}_S^n} \delta_c(x)\cdot G(x)}{\sum_{x\epsilon\mathbb{D}_S^n} \delta_c(x)}$$

- Compute cosine similarity $S(r_T, r_n)$ between source and target based on centroids
- We extend this by temperature-controlled softmax to MDMGB+

$$w_n = \frac{\exp\big(\tau\cdot S(r_T,r_n)\big)}{\sum_{j=1}^{N}\exp\big(\tau\cdot S(r_T,r_j)\big)}$$

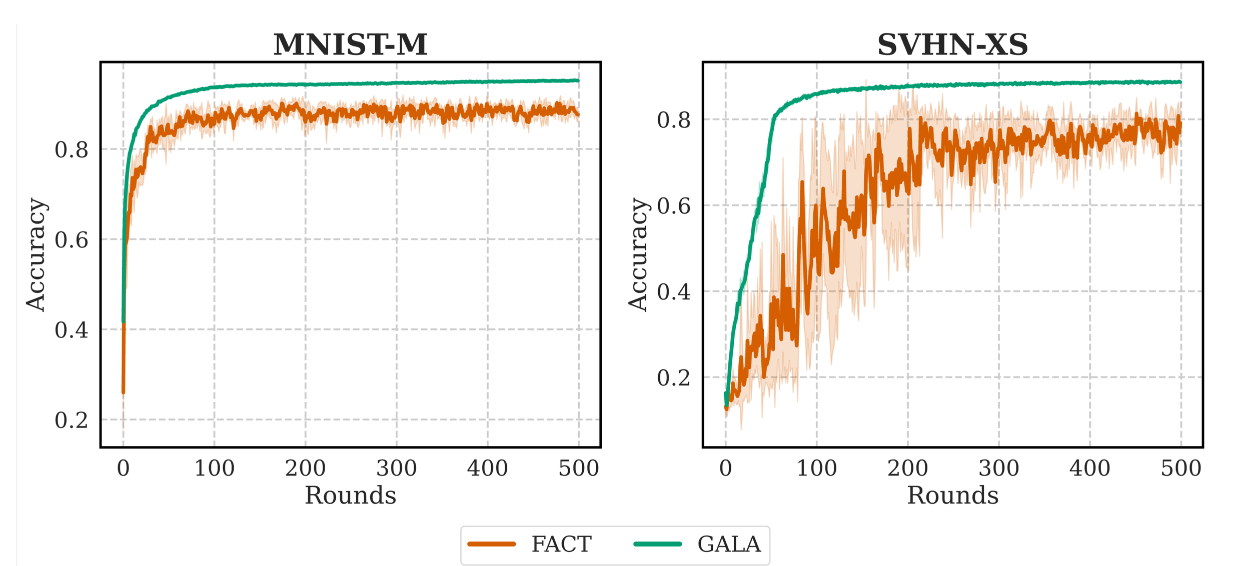
The GALA Method



Digit-18 Benchmark

Results on High-source Setting

Method	*mnist*	*mnistm*	*svhn*	*syn*	*usps*	*synm*	*svhn-xs*	*svhnstack*	*usps-m*	*ardis-xs*	Avg
Oracle	$99.0_{\pm0.03}$	$95.6_{\pm0.26}$	$88.4_{\pm0.15}$	$97.0_{\pm0.12}$	$98.9_{\pm0.12}$	$83.8_{\pm0.32}$	$84.7_{\pm0.37}$	$86.5_{\pm0.04}$	$92.0_{\pm0.51}$	$97.5_{\pm0.32}$	92.3
FACT	$98.6_{\pm0.20}$	$87.6_{\pm1.58}$	$92.5_{\pm0.41}$	$97.4_{\pm0.46}$	$98.3_{\pm0.26}$	$79.4_{\pm1.51}$	$79.1_{\pm4.75}$	$91.9_{\pm1.63}$	$87.6_{\pm1.06}$	$93.9_{\pm1.21}$	90.6
GALA	$\mathbf{99.3}_{\pm0.07}$	$\mathbf{95.2}_{\pm0.10}$	$\mathbf{95.4}_{\pm0.17}$	$\mathbf{98.4}_{\pm0.05}$	$\mathbf{99.0}_{\pm0.10}$	$\mathbf{85.8}_{\pm0.23}$	$\mathbf{88.6}_{\pm0.29}$	$\mathbf{94.8}_{\pm0.10}$	$\mathbf{92.9}_{\pm0.34}$	$\mathbf{97.1}_{\pm0.21}$	**94.6**
	↑ 0.7	↑ 7.6	↑ 2.9	↑ 1.0	↑ 0.7	↑ 6.4	↑ 9.5	↑ 2.9	↑ 5.3	↑ 3.2	↑ 4.0



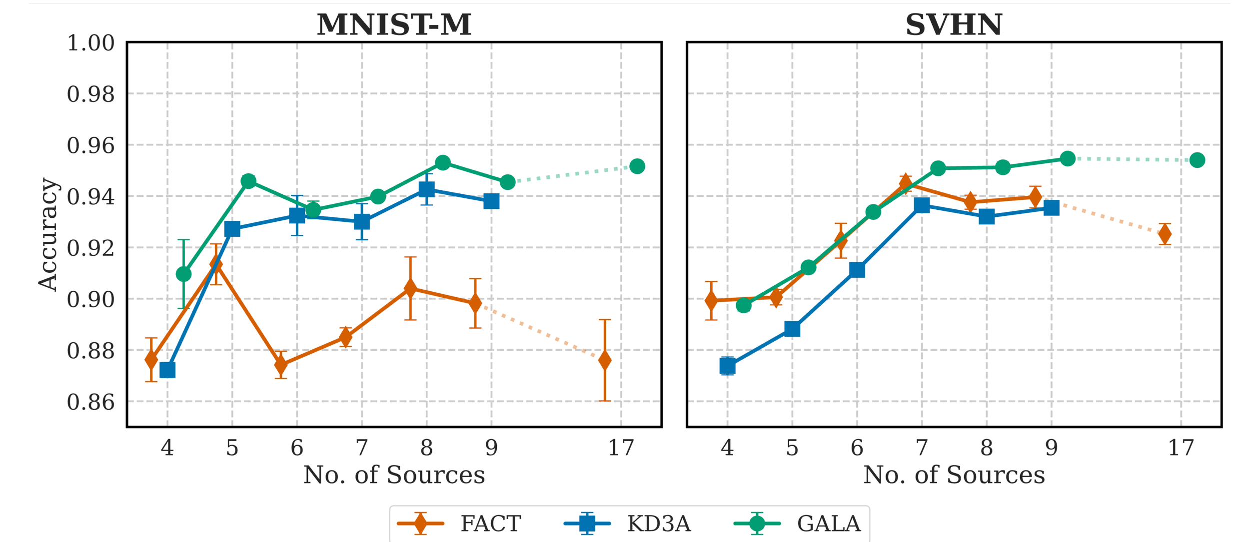
# Sources	3	5	7	9
KD3A	$50.73_{\pm0.3}$	$216.03_{\pm2.7}$	$1029.84_{\pm10.6}$	$5600.48_{\pm32.0}$
FACT	$3.65_{\pm0.3}$	$3.22_{\pm0.4}$	$3.39_{\pm0.5}$	$4.05_{\pm0.5}$
GALA (ours)	$17.27_{\pm0.1}$	$17.79_{\pm0.7}$	$22.21_{\pm0.1}$	$22.37_{\pm0.1}$

Per-round training time (in seconds)

With growing number of sources FACT [1] struggles with performance and stability, while KD3A [3] becomes prohibitively slow.

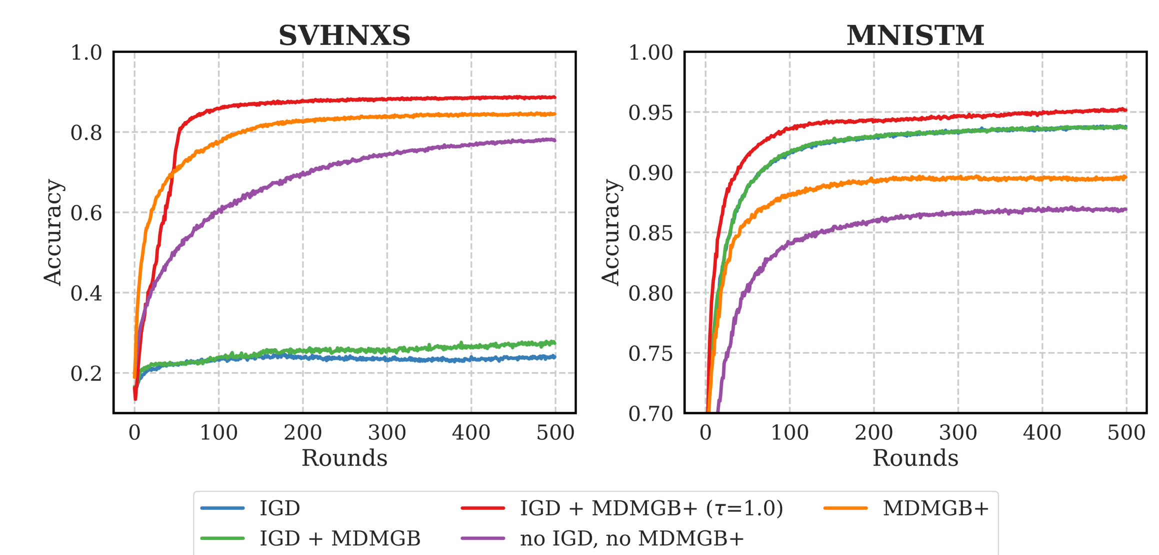
Increasing Number of Sources

GALA provides robust performance and prevents negative transfer, while still being practical



Ablation

All components of GALA contribute to high performance, in particular for challenging targets



Conclusions

- ✓We propose **GALA** to address the scalability challenges of diverse distributed UMDA
- ✓We introduce **Digit-18** as a new multi-source benchmark
- ✓Temperature-scaled centroid **weighting** and **inter-group discrepancy minimization** achieves stable, efficient alignment to an unlabeled target domain.
- ✓GALA outperforms existing methods on standard UMDA benchmarks and the new high-diversity **Digit-18** dataset

[1] "FACT: Federated Adaptive Cross Training." Knowledge-Based Systems (2025): 113655.

[2] "A framework for self-supervised federated domain adaptation." EURASIP Journal on Wireless Communications and Networking 2022.1 (2022): 37.

[3] "KD3A: Unsupervised Multi-Source Decentralized Domain Adaptation via Knowledge Distillation." ICML. Vol. 4. 2021.

Larissa Reichart:
larissa.reichart@uni-tuebingen.de

Cem Ata Baykara:
cem.baykara@uni-tuebingen.de