

Dual Attention Heads for Personalized Federated Learning in Multi-Center ECG Classification

Joseph Lindley, Kien Le, Quoc Bao Phan, Tuy Nguyen

I. Background Information

- Massive amounts of ECG data overwhelm doctors, leading to diagnostic errors [1].
- AI can automate ECG analysis, but training accurate AI requires diverse datasets [2], [3].
- We propose a novel Dual-Head Transformer to personalize the AI for each dataset coupled with federated learning to protect patient data (FedDualAtt)

II. Methods

Setup - 4-client cross-silo federated learning (FedCVD benchmark [4])

Proposed Architecture:

ResNet1D-34 - extracts temporal features from ECG
Dual Attention Mechanism -

- Global Heads: Shared with server to learn broad patterns
- Local Heads: Kept private and learns site-specific patterns

Federated Learning:

1. Server broadcasts global parameters
2. Clients merge global and local parameters
3. Client trains on private ECG dataset
4. Clients only upload updated global parameters
5. Server aggregates global heads using weighted average

Metrics:

Micro-F1: Measures overall reliability by aggregating true/false positives and false negatives.

Mean Average Precision (mAP): Evaluates precision across all 20 disease types equally

III. System Overview and Experimental Results

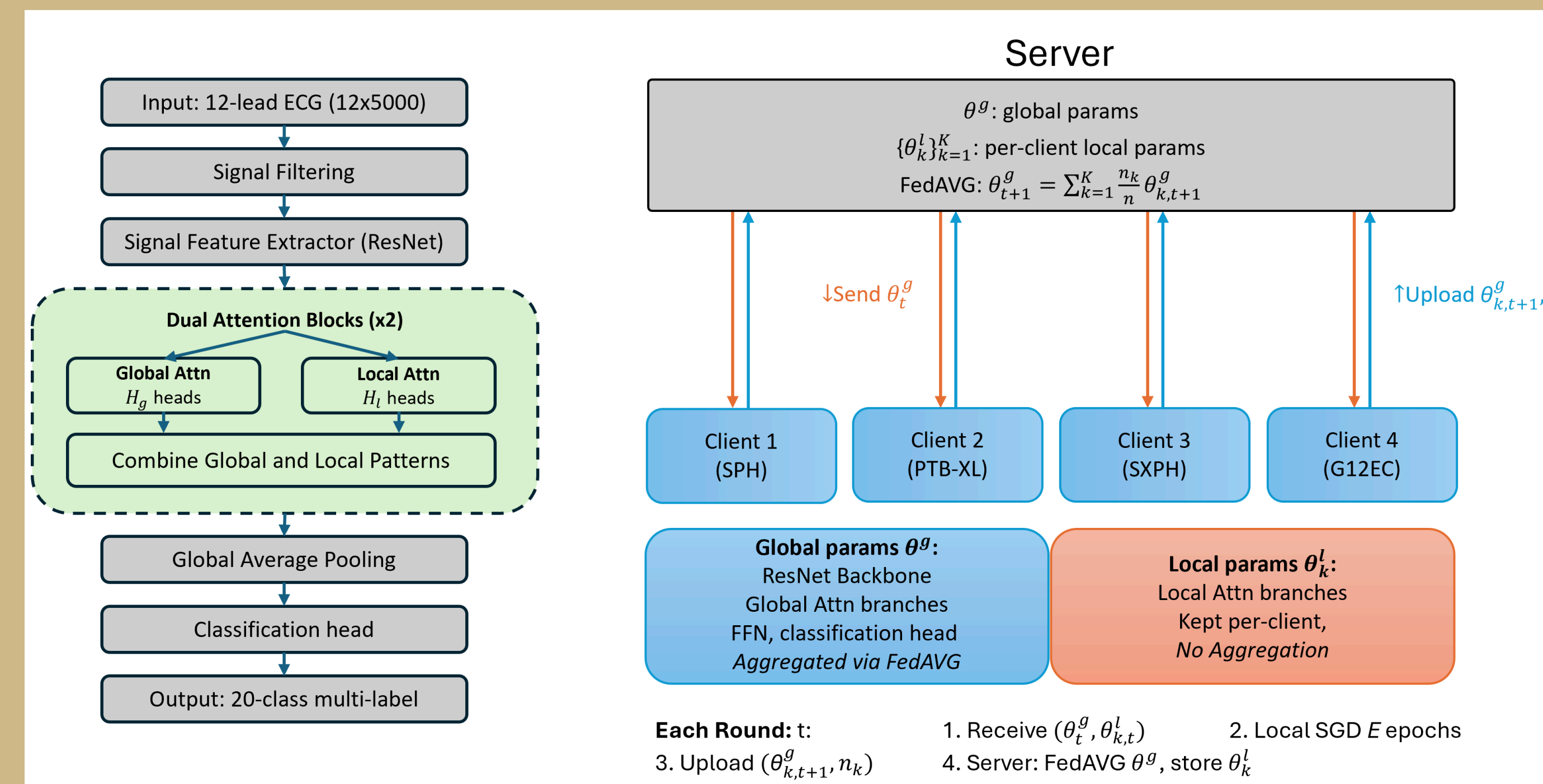


Figure 1. Overview of the FedDualAtt architecture

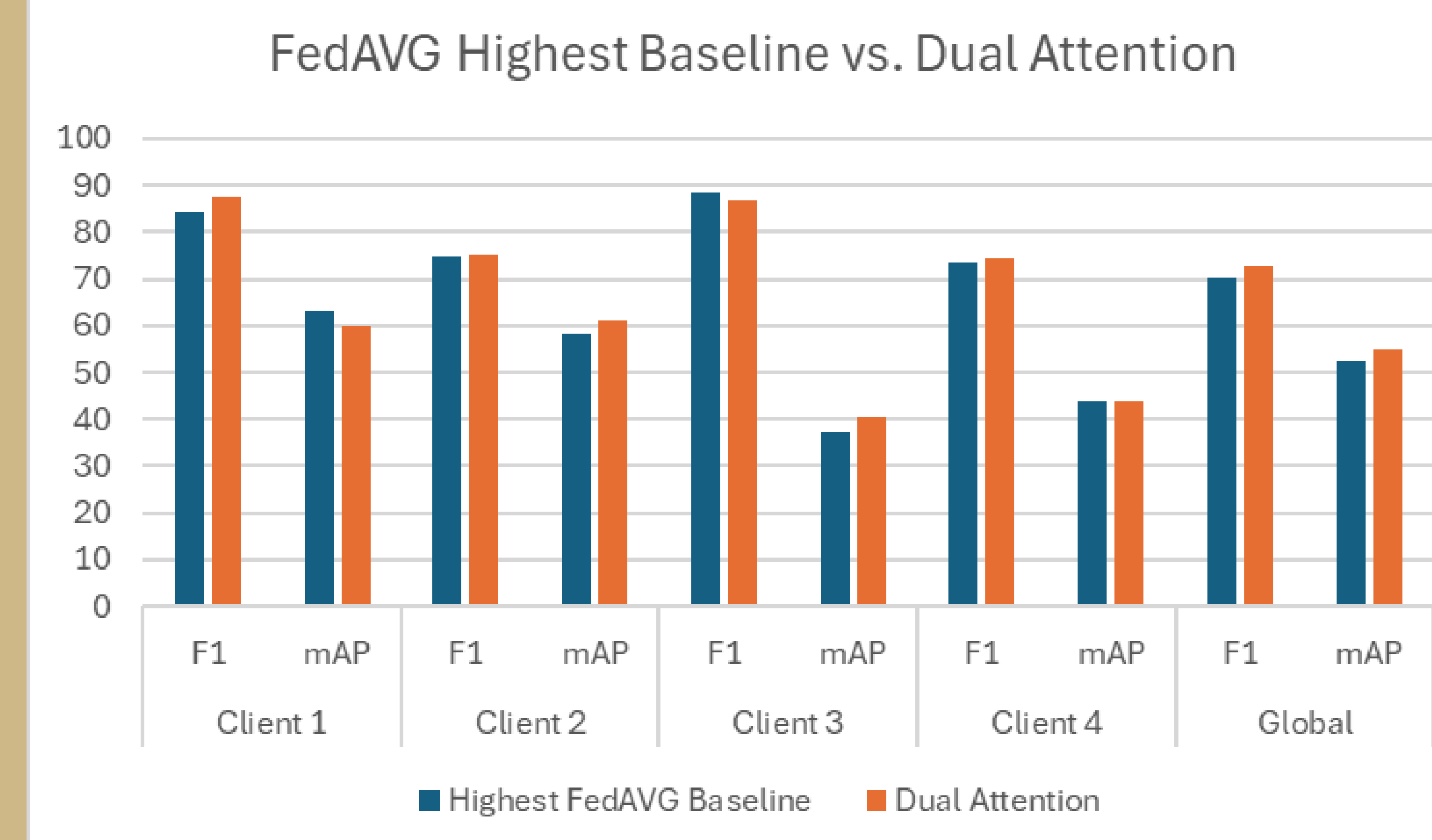


Figure 2 - Comparison showing FedDualAtt outperforming the FedAVG baseline [4]

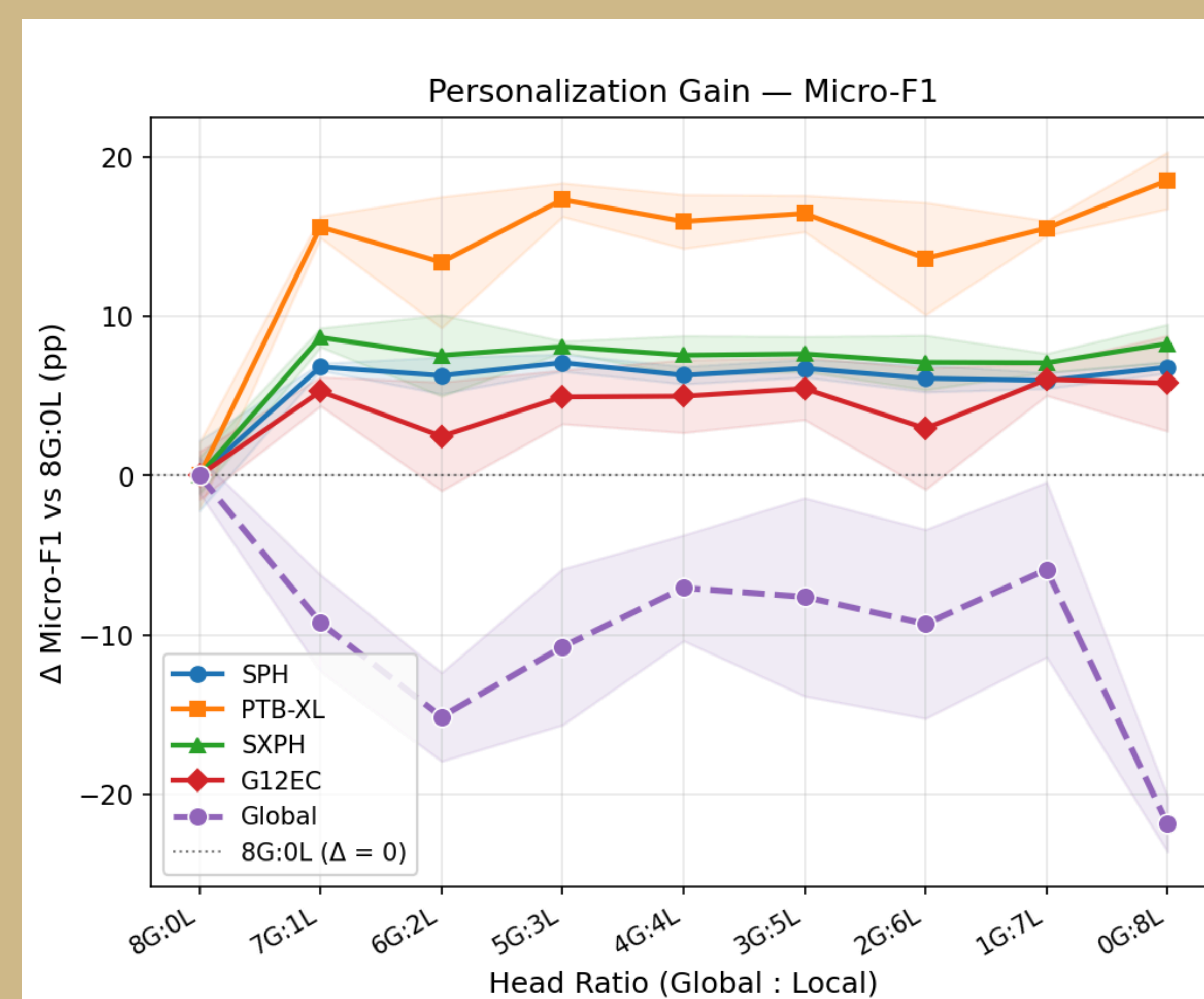


Figure 3 - Global-to-local head ratios and performance on each dataset

- FedDualAtt model outperformed standard Federated Averaging (FedAVG) baselines across almost all metrics
- Use of local heads showed increase of up to 15% in F1 score and 5% in mAP.
- Some datasets prefer more personalization than others

IV. Discussion

- Architecture allows AI to train from large, diverse datasets while personalizing to each client
- Performance from personalization is not scalable
- Ideal split between heads must be determined manually by observing performance
- Future research may investigate the optimal global vs local heads needed for a specific type of dataset

References

1. Drew, B. J., Harris, P., Zègre-Hemsey, J. K., Mammone, T., Schindler, D., Salas-Boni, R., Bai, Y., Tinoco, A., Ding, Q., & Hu, X. (2014). Insights into the Problem of Alarm Fatigue with Physiologic Monitor Devices: A Comprehensive Observational Study of Consecutive Intensive Care Unit Patients. *PLoS ONE*, 9(10). <https://doi.org/10.1371/journal.pone.0110274>
2. Ebrahimi, Z., Loni, M., Daneshlab, M., & Gharehbaghi, A. (2020). A review on deep learning methods for ECG arrhythmia classification. *Expert Systems with Applications: X*, 7(100033). <https://doi.org/10.1016/j.eswax.2020.100033>
3. Nguyen, D. C., Pham, Q.-V., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., Dobre, O. A., & Hwang, W.-J. (2021). Federated Learning for Smart Healthcare: A Survey. *ArXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2111.08834>
4. Zhang, Y., Chen, G., Xu, Z., Wang, J., Zeng, D., Li, J., Wang, J., Qi, Y., & King, I. (2024). FedCVD: The First Real-World Federated Learning Benchmark on Cardiovascular Disease Data. *ArXiv.org*. <https://arxiv.org/abs/2411.07050>