Supplement Data and Results

for the following study

Azghadi, Seyed Alireza Rahimi, et al. "SF2D: Semi-supervised Federated Learning for Fall Detection using (Un) labelled Data in Edge-Cloud." Proceedings of the Canadian Conference on Artificial Intelligence. 2025.

Context

As part of the above-mentioned research work, additional results are reported here to support reproducibility and transparency.

During the course of evaluation, an initial analysis [1] was conducted on the SiSFall dataset [2] without any adjustment in the original data. This approach was chosen to ensure that the baseline performance of the proposed method could be assessed under the raw distribution of activities and falls present in the dataset. For completeness, we have applied the oversampling technique to make the dataset more balanced.

The model, threshold, and evaluation code remained unchanged. We carried out additional experiments on the adjusted dataset. Both evaluation results will be reported as follows.

Experiment 1: Original Evaluation

Confusion matrix (rows = actual, columns = predicted):

$$\begin{bmatrix} \text{TN} = 49620 & \text{FP} = 264 \\ \text{FN} = 153 & \text{TP} = 1953 \end{bmatrix}$$

Per-class metrics (positive class = minority):

Class	Precision (%)	Recall (%)	F1 (%)
Negative	99.69	99.47	99.58
Positive	88.09	92.74	90.35

Aggregate metrics:

Metric	Value	Notes
Accuracy (%)	99.19	$(\text{TP} + \text{TN})/N$
Specificity (%)	99.47	$\text{TN}/(\text{TN} + \text{FP})$
Balanced Accuracy (%)	96.10	$\frac{1}{2}(\text{TPR} + \text{TNR})$
Matthews Corr. Coef. (MCC)	0.900	$\frac{\text{TP}\cdot\text{TN} - \text{FP}\cdot\text{FN}}{\sqrt{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})}}$
Macro F1 (%)	94.98	mean of classwise F1
Weighted F1 (%)	99.21	support-weighted mean of classwise F1

Experiment 2: Evaluation on Adjusted Dataset (Additional Experiment Results)

Confusion matrix (rows = actual, columns = predicted):

$$\begin{bmatrix} \text{TN} = 37649 & \text{FP} = 710 \\ \text{FN} = 688 & \text{TP} = 20012 \end{bmatrix}$$

Per-class metrics (positive class = minority):

Class	Precision (%)	Recall (%)	F1 (%)
Negative	98.21	98.15	98.18
Positive	96.57	96.68	96.62

Aggregate metrics:

Metric	Value	Notes
Accuracy (%)	97.63	$(\text{TP} + \text{TN})/N$
Specificity (%)	98.15	$\text{TN}/(\text{TN} + \text{FP})$
Balanced Accuracy (%)	97.41	$\frac{1}{2}(\text{TPR} + \text{TNR})$
Matthews Corr. Coef. (MCC)	0.948	definition as above
Macro F1 (%)	97.40	mean of classwise F1
Weighted F1 (%)	97.63	support-weighted mean of classwise F1

Summary and Interpretation

- The Accuracy from the original SiSFall dataset obtained from our proposed method is 99.19%, while the Accuracy from the additional experiment is 97.63%, reflecting a more realistic estimate of general performance.
- Minority-class performance from the original evaluation and additional experiments for positive-class F1 are 90.35% and 96.62%, respectively. This proves further that our proposed method is transparent and unbiased in its decision.
- The balanced accuracy of the original experiment was 96.10%, while the additional experiment shows that the mentioned balanced accuracy is 97.41%.
- The MCC scores from both experiments are as follows: 0.900 and 0.948, confirming overall reliability.

Reproducibility Note. These results are provided as supplementary material to ensure transparency and reproducibility. The only modification was the extra experiment on the adjusted dataset; all model settings, thresholds, and evaluation code remained unchanged. This project is available on GitHub: https://github.com/Analytics-Everywhere-Lab/federated_fall_detection.

References

[1] Seyed Alireza Rahimi Azghadi, Hung Nguyen, Irina Kondratova, Hélène Fournier, Monica Wachowicz, Francis Palma, René Richard, and Hung Cao. Sf2d: Semi-supervised federated learning for fall detection using (un) labelled data in edge-cloud. *The 38th Canadian Conference on Artificial Intelligence*, 2024.

[2] Angela Sucerquia et al. Sisfall: A fall and movement dataset. *Sensors*, 17(1):198, 2017.