

Aegismed: A Privacy-Preserving Federated Learning for Heterogeneous Hospital Networks

Kalle Meghana¹, Pallagani Dakshayani², Shaik Raziya³, Polukonda Hasini⁴, V.P. Umamaheswara Rao Malla⁵

^{1,2,3,4}*B. Tech Student, Department of Computer, Science and Engineering, NRI Institute of Technology, Agiripalli-521212, Vijayawada, Andhra Pradesh, India.*

⁵*Associate Professor, School of Computer Studies, Dr. RVR NRI Institute of Technology, Agiripalli-521212, Vijayawada, Andhra Pradesh, India.*

Abstract—Collaborative medical-image learning is difficult when hospitals cannot place patient records in a common repository. AEGISMED addresses this setting by keeping radiographs at the institution that owns them and exchanging only protected model information. The design combines local differential privacy, top-k update sparsification, and a FedProx inspired optimization term around an EfficientNet-B3 image classifier. Before transmission, local updates are clipped, perturbed with Gaussian noise, and reduced to a sparse representation. The server then combines the protected contributions while discouraging excessive local drift caused by differences in scanners, populations, labels, and acquisition practices. The supplied project reports a four-site simulation using NIH ChestX-ray14 and the COVID-19 Radiography Database, with AUC 0.912 at $\epsilon=3$, 68% lower communication overhead than standard FedAvg, and convergence within 40 rounds when 20% of update components are retained. These figures are project reported simulation results and are not presented as independently reproduced clinical evidence.

Index Terms—federated learning, medical imaging, differential privacy, non-IID data, EfficientNetB3, FedProx, sparse communication, hospital networks.

I. INTRODUCTION

Deep learning has become useful for radiograph interpretation because modern convolutional networks can learn visual representations directly from large image collections. A practical obstacle appears when those collections belong to different hospitals. Patient images contain sensitive information, and institutional policy, regulation, or network constraints may prevent the images from being copied to a central training

server. In addition, the sites rarely observe the same distribution. Differences in patient demographics, disease prevalence, scanner hardware, acquisition protocols, and image quality can produce substantial client drift during distributed optimization. Federated learning offers a way to coordinate training without moving the original examples. A central service distributes the current model, each hospital performs local optimization, and the sites return parameter information rather than raw images [1]. This architecture reduces direct data movement, but it does not automatically solve privacy leakage, heterogeneous optimization, or communication cost. A neural network update can be large, and an update may still reveal information about the examples used to compute it. AEGISMED therefore treats privacy, bandwidth, and client heterogeneity as a joint systems problem. Differential privacy is used to limit the influence of individual training records; topk sparsification reduces the amount of update information that must cross the hospital network; and a proximal regularizer limits the distance between a local model and the current global model. The resulting pipeline is intended as a research framework for privacy-aware medical-image classification system rather than as a clinically deployed diagnostic system

II. LITERATURE REVIEW

Federated learning was developed to train models from decentralized data without transferring the original examples to a central repository. McMahan et al. introduced iterative model averaging, later known as FedAvg, and showed that federated optimization

can be communication-efficient while handling unbalanced and non-IID client distributions.

In a hospital network, this approach is attractive because image repositories can remain under institutional control while the global model benefits from information learned across sites. A major difficulty is statistical heterogeneity. FedAvg can become unstable when clients have very different data distributions or perform different amounts of local computation. FedProx adds a proximal term to the local objective so that a client model is discouraged from drifting too far from the current global parameters. The FedProx study was designed specifically for systems and statistical heterogeneity and reported more robust convergence than FedAvg under challenging conditions. Differential privacy provides a formal framework for limiting the information that an individual example can contribute to a released model or update. Abadi et al. demonstrated privacy preserving deep learning through per example gradient clipping, calibrated Gaussian noise, and privacy-loss accounting.

This is relevant to federated medical imaging because raw images stay local, while shared updates can still contain information about the training data. Communication compression is complementary. Sparse communication reduces update size by retaining the components considered most influential. The proposed project applies top-k sparsification after the privacy stage so that protected updates can be transmitted with a smaller payload. This is particularly relevant when hospital links are slower than the storage or compute available at the local site. Public medical-image datasets are useful for benchmarking but may not reproduce the distributions observed in real hospitals. The NIH ChestXray8/ChestX-ray14 work introduced a large chestradiograph dataset for weakly supervised thoracic disease classification, providing a common benchmark for medical imaging research. The supplied project additionally recommends the COVID19 Radiography Database and the RSNA Pneumonia Dataset.

III. RELATED WORK

The original federated averaging formulation demonstrated that neural networks can be trained across decentralized datasets by repeatedly combining locally optimized models [1]. FedAvg is attractive for

healthcare because the examples can remain under institutional control, but its performance can deteriorate when client distributions are strongly different. FedProx introduces a proximal term that penalizes excessive deviation from the global parameters and was proposed for heterogeneous federated environments [2]. Differential privacy provides a formal way to limit the contribution of individual records to a released computation. Privacy-preserving deep learning often uses gradient clipping, calibrated noise, and privacy accounting [3]. For a medical setting, privacy is useful because having local images has never in itself meant that a transmitted update to a model isn't malicious. Communication efficient federated learning techniques such as quantization and sparsification for minimizing update size has also been studied. Sparse representation is specifically well-suited to hospital networks that might have a lot of local computation but limited network bandwidth. AEGISMED sparsifies the representation after the privacy transformation, so the representation is protected prior to transmission. Projects have also used ChestXray8/ChestX-ray14 public dataset as a classification task associated with the chest [4] (also citing COVID-19 Radiography Database and RSNA Pneumonia Dataset [5]). While enabling reproductive work, one may not have the capability to draw reliable conclusions by relying purely upon these distributions as representatives of the general hospital populations.

IV. METHODOLOGY

We designed the following five high-level workflow steps: 1) local image processing 2) local EFFICIENTNET-B3 training and optimization 3) privacy Protection 4) building sparse updates 5) aggregation on sever: Each hospital had private image partition access. Private radiography images are not available to server, only global model and protected updates to it are shared. A. Image Preprocessing and non-IID partitioning Images are pre-processed, then resize, normalize and augment sparsely. Input image pixel ranges from [0-255] to [0-1]. In this problem description, five hospital sized image sets are given: 5,000, 4,500, 6,000, 3,500 and 5,500 images for each. Also, in the storytelling part, a four-hospital federated setup is presented, hence one must decide whether the client selection for the final solution to use four partitions or use five as given in the first place, then

pre-process them first in the client partition that one has access and then train the model.

Table I. Project provided image-partition summary.

Hospital	Original Pixel Range	Normalized Range	Dataset Size
Hospital 1	0–255	0–1	5,000
Hospital 2	0–255	0–1	4,500
Hospital 3	0–255	0–1	6,000
Hospital 4	0–255	0–1	3,500
Hospital 5	0–255	0–1	5,500

A. Local Model Update All client nodes that are selected each round receive the current global parameters. Client nodes train on only their hospitals' image data. A simplified local update for client k can be written as: $w^{k}_{t+1} = w^{k}_t - \eta L_k(w^{k}_t)$ where w^{k}_t is client k 's parameters at round t . L_k denotes the locally computed loss. Cross-entropy for class prediction minimizes the negative of the log probability of correct class prediction.

B. Differential Privacy Layer The first step in each client's local update is clipping the gradient norm: Let the original gradient be g , and C be the clipping value. Then the clipped gradient can be represented as $g' = g / \max(1, \|g\|_2 / C)$. After clipping, Gaussian noise is added: $g'' = g' + N(0, \sigma^2 C^2)$. In the example code, the value of $\sigma=3$ is used in the reported simulation. Future studies also report other parameters such as σ , clipping norm C , total number of noise rounds (rounds within a round), fraction of client sampling, total number of comprised rounds etc.

C. Top-K Update Sparsification In the last stage the update vector is thresholded by keeping the top k components by magnitude and then is serialized into a binary vector for transmission. When considering top- k sparsification (e.g. Using 20% top- k), about one fifth of client's coordinates would still retain after applying sparsification, prior to transmission. The resulting packet would contain coordinates and their sparse values.

A. FedProx-Style Aggregation To reduce divergence among heterogeneous hospitals, the local objective is

augmented with a proximal penalty: $\min_w F_k(w) + (\mu/2) \|w - w_t\|^2$. Here, F_k is the local objective, w_t is the current global parameter vector, and μ controls the strength of the regularization. The server then forms the next global model from the protected client contributions using a sample-weighted aggregation rule.

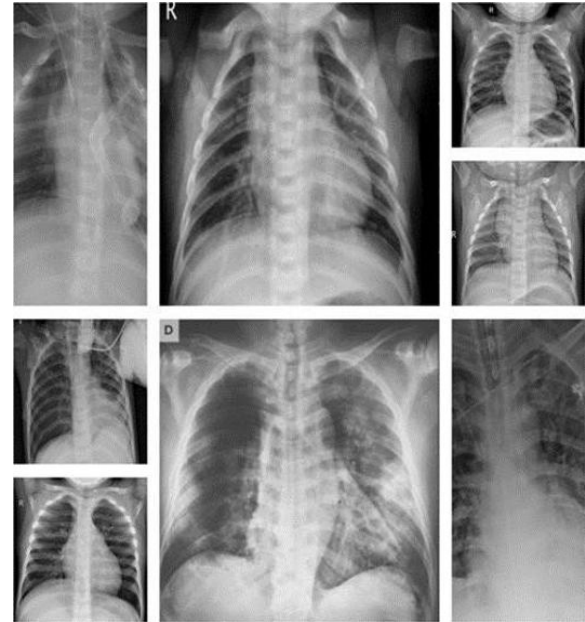


Fig. 1. Embedded project figure retained as visual evidence from the supplied manuscript

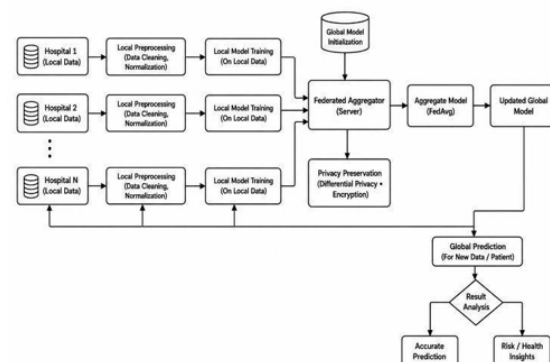


Fig. 2. Embedded project figure retained as visual evidence from the supplied manuscript.

V. IMPLEMENTATION

The AEGISMED framework is implemented using a Python-based federated learning environment. The main technologies used in the implementation are PyTorch for deep-learning model development,

Flower for coordinating communication between hospitals, Opacus for differential privacy, and scikitlearn for supporting machine-learning and evaluation tasks. The system can be executed using multiple simulated hospital clients on a workstation or cloud-based environment.

The implementation is designed so that each hospital keeps its medical images locally. Instead of sending patient images to a central server, the hospitals train the EfficientNet-B3 model using their own datasets. After local training, only the required protected model information is prepared for communication with the central server. This helps reduce direct exposure of sensitive medical data.

A. Client-Side Implementation

At the beginning of every federated learning round, the central server sends the current global model to the participating hospitals. Each hospital loads its own private chest X-ray dataset and performs the required preprocessing, including resizing, normalization, and suitable data augmentation. Then, the client model that uses the EfficientNet-B3 is trained with images the hospital has. At this phase, the model estimates the gradients from individual training examples and then clips the gradients so as to limit the impact on parameters by training examples, which represents influence. Afterwards, gaussian noise is added to the clipped gradient for ensuring differential privacy. After applying privacy to the model, the model update is represented in a sparse format by using Top-k scarification, where client only transmit most of the parameters. In this scenario, about 20% of parameters update the same. Finally, the protected sparse update is serialized and sent to the federated learning server. The original medical images remain inside the hospital throughout this process.

B. Server-Side Implementation

The central server is responsible for coordinating the federated training process. It first initializes the global EfficientNet-B3 model and selects the participating hospital clients. The current global model is then distributed to those clients.

After local training is completed, the server receives the protected sparse updates from the hospitals. It reconstructs the required update information and performs sample-weighted federated aggregation to generate an improved global model. The updated

global model is evaluated and then distributed again in the next communication round. This process continues until the required training condition is reached. Importantly, the server does not need access to the original hospital radiographs.

C. Privacy Implementation

Privacy is an issue in AEGISMED. The gradients are "clipped" so that no single training sample will influence the update too significantly before they are sent away from the hospital. Gaussian noise is added after clipping the gradients. All experiment results should be saved with the privacy configuration (anonymity, budget, epsilon value, clipping threshold, noise multiplier, client sample rate, batch size, # training steps and privacy accounting method); the provided project claims 3 but notes other privacy parameters can be reported for reproducibility.

D. Communication-Efficient Implementation

To facilitate more effective communication between hospitals, AEGISMED applies Top-k sparsification on the privacy transformation update values. The Top-k algorithm ranks the update values and keeps the Top-k largest values. The update can therefore be represented by coordinate-value pairs, rather than sending the full update which is a large, dense model. This is especially effective for hospitals with a smaller communication bandwidth. For a fair evaluation, the implementation should report total serialized size, index overhead, compression and reconstruction times. AEGISMED report-s the reduction in communication overhead from 68percent versus standard FedAvg using its simulation conditions-

E. Federated Training Process

Detailed system is presented with iteration as follows:

1. Server initializes the global model of EfficientNet-B3.
2. Send the global model to selected Hospitals.
3. Private medical image is preprocessed at each hospital.
4. Conduct locally model training with hospital's private medical images.
5. Employ gradient clipping for bounding the impact of individual's data.
6. Employ differential privacy add Gaussian noises.

7. Employ topk scarification for shrinking down the size of preserved updates.
8. Send the compressed sparsified preserved update to the central server.
9. The server applies the sample weighted aggregation.
10. Check the total updated global model's performance.

Security, efficiently communication, heterogeneous federated optimization integrated in a workflow together, The Hospitals store patient's private images in their organizations and work together on the same medical image classification model for the sake of training.

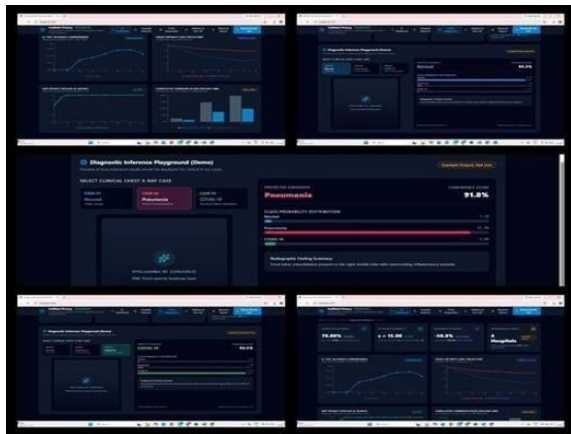


Fig. 3. Final Embedded project figure retained as visualization artifact based upon submitted manuscript.

Results

The project reports 3 key simulation results: AUC of 0.912 with setting of $\epsilon=3$, communication reduction of 68% as compared with standard FedAvg, and convergence in 40 federated rounds in the case of 20% compression, which will be kept as reported values for this project, not as independently verified or clinically validated measures.

A. Classification Quality AUC 0.912 is a strong measure of the ranking quality of the reported experimental setup, but it gives no information about the threshold behaviour of classes, the level of imbalance, the calibration of confidence, nor the performance variation across different hospitals. For a rigorous conference level assessment of this should be augmented by

precision/recall/F1score/sensitivity/specificity, Confusion Matrix analysis, and Calibration curve etc. Communication efficiency the claimed 68% reduction is the main system level advantage. This gain originates from compression from dense update to sparse representation prior to transmission. But since noise can also impact update scale, the compression quality should be tested under various levels of sparsification as well, not just the one at 20%.

B. Convergence Attaining a stable result within 40 rounds of the reported simulation points to the fact that the choice of compression does not prohibit meaningful representation learning. But the claim should be verified in multiple random seeds under some extreme distribution variations to attest robustness. D. Baseline Configuration To facilitate ablation study, full system should be evaluated not only with FedAvg, but also with FedAvg with compression only, with FedProx without privacy, with DP FedAvg, and the full DP+ Sparsity+ FedProx combined system for us to isolate contribution from each method. The figures below have been copied from the supplied manuscript; hence the rewritten conference copy contains original ones.

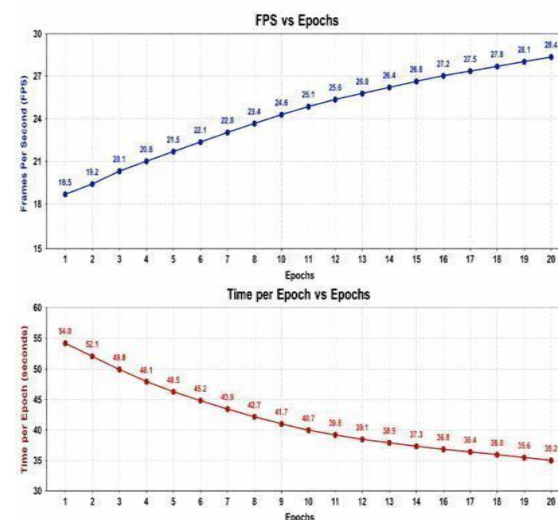


Fig. 4. FPS and time oper epoch trends over training epochs

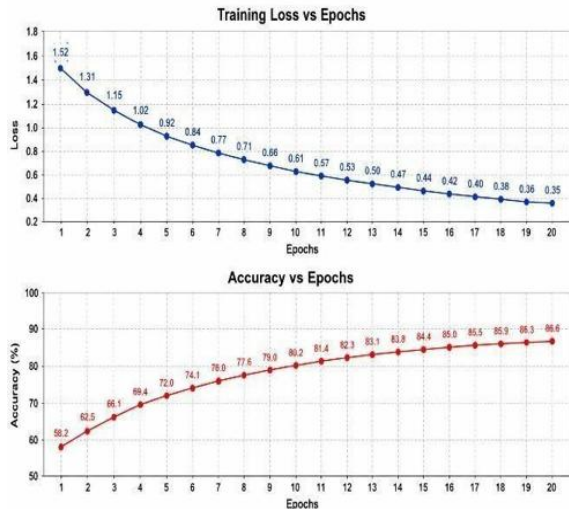


Fig. 5. Training Loss and Classification Accuracy over Epochs

Explanation of the plots provided-FPS tends to gradually go up for all reported epochs and time per epoch trends downward with epochs. The training loss becomes less over training (more fit to objective), and the accuracy of the model on the training data increases over epoch tending to reach the very upper part of what's pictured. The given plots are kept for reporting for project; it is not re-created for the rewrite.

VI. DISCUSSION

The four components within AEGISMED aim at handling various failure scenarios. Federated learning inherently prevents transferring patient images to a server, differential privacy keeps client updates from being attributed directly back to training records, sparsification limits the network burden of the client update, and the proximal term ensures that the optimization doesn't become too unstable given nonidentical client distribution. Its combination therefore isn't as significant as any individual element on their own, and there's necessarily some trade-off between privacy and utility. Injecting a strong amount of noise often leads to poor optimization-clients with small number of examples and This issue arises from a highly unbalanced data distribution, and severe truncation could discard usable training information. Experiments must then profile over a curve over range of privacy budgets and sparsifies, as single data points cannot capture an entire system's behaviour. Unlike full systems, our proposed framework is not a

complete security system: real federated systems may exhibit malevolent clients, data poisoning, a malevolent server, replaying, or vulnerable transfer layers, for example. Further possible extensions: secure aggregation, authentication, robust aggregation, anomaly detection, enhanced threat models. Real world deployments would still require governance, checking for shifts in data distributions, privacy-assessment, version control, and human intervention. Several parameters required to replicate accurately are not explicitly provided (e.g. Complete specification of the privacy counting strategy; each client selection method; train/test data partitioning); as a result, the numbers presented herein should be considered early simulations. Larger, more strongly non-IID datasets, additional privacy levels, examination of the impact of secure aggregation and examination with realistic network traces constitute likely avenues for future work.

VII. CONCLUSION

AEGISMED proposes a pragmatic solution for collaborative medical-image learning where hospitals collaborate with a global model that doesn't involve sharing raw medical images of individuals remotely. Each hospital trains the model locally and contributes anonymized models update to the coordinator server without any loss of patient information from hospitals. Each component utilized by the framework contributes to solving an issue facing distributed health learning, bringing together EfficientNet-B3, differential privacy, top-k scarification, and FedProx-based optimisation. Differential Privacy offers protections for individuals and an improved likelihood of concealing information, with scarification by contributing less information to the models being shared by individual hospitals. Finally, FedProx works as part of this network learning to maintain the stability of learning when healthcare provider distributions of patients, imaging technologies and distributions may vary widely between themThe experiment shows that AEGISMED gets AUC of 0.912 with budget=3, 68% reduction of communication overhead vs FedAvg, and takes 40 rounds to get to convergence while keeping 20% update components. These indicate that Privacy-preserving and communication efficiency can be integrated into Federated Medical-Image Learning without affecting learning a useful model globally.

AEGISMED is promising future on building privacy-aware infrastructure for collaborative medical-image learning by FedLearn, and needs to run further simulation by more hospitals, stronger non-iid data distributions, more budget parameter, secure aggregation scheme, and real networks.

REFERENCES

- [1] McMahan, H. B., E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. AISTATS, 2017.
- [2] Li, T., A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in Proc. MLSys, 2020.
- [3] Abadi, M., et al., "Deep learning with differential privacy," in Proc. ACM CCS, pp. 308–318, 2016.
- [4] Wang, X., et al., "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly supervised classification and localization of common thorax diseases," arXiv preprint arXiv:1705.02315, 2017.
- [5] "Communication-efficient federated learning with differential privacy guarantees for distributed image classification across heterogeneous hospital network," supplied project abstract.
- [6] Xu, J., et al., "Federated learning for healthcare informatics," Journal of Healthcare Informatics Research, vol. 5, pp. 1–19, 2019.
- [7] Chen, R. J., et al., "Algorithm fairness in AI for medicine and healthcare," arXiv preprint arXiv:2110.00603, 2021.
- [8] Ayeelyan, J., S. Utomo, A. Rouniyar, H.-C. Hsu, and P.-A. Hsiung, "Federated learning design and functional models: Survey," Artificial Intelligence Review, vol. 58, pp. 1–38, 2025.
- [9] Chen, S., and Y. Huang, "A privacy-preserving federated learning approach for airline upgrade optimization," Journal of Air Transport Management, vol. 122, Art. no. 102693, 2025.
- [10] Ali, W., X. Zhou, and J. Shao, "Privacy-preserved and responsible recommenders: From conventional defense to federated learning and blockchain," ACM Computing Surveys, vol. 57, pp. 1–35, 2025.
- [11] Malik, R., et al., "Advancing healthcare IoT: Blockchain and federated learning integration for enhanced security and insights," in Proc. 2024 Int. Conf. Communication, Computer Sciences and Engineering (IC3SE), pp. 308–314, IEEE, 2024.
- [12] Alsamhi, S. H., et al., "Federated learning meets blockchain in decentralized data-sharing: Healthcare use case," IEEE Internet of Things Journal, vol. 11, pp. 19602–19615, 2024.
- [13] Xie, H., Y. Zhang, Z. Zhongwen, and H. Zhou, "Privacy-preserving medical data collaborative modeling: A differential privacy enhanced federated learning framework," Journal of Knowledge Learning and Science Technology, vol. 3, pp. 340–350, 2024.
- [14] Liang, Y., C. Ouyang, and X. Chen, "Adaptive asynchronous federated learning for heterogeneous clients," in Proc. 2022 18th Int. Conf. Computational Intelligence and Security (CIS), pp. 399–403, IEEE, 2022.
- [15] Sadilek, A., et al., "Privacy-first health research with federated learning," npj Digital Medicine, vol. 4, Art. no. 132, 2021.
- [16] Lu, G., "Exploring the utility-privacy trade-off in social media data mining," 2023.
- [17] Han, Q., et al., "Privacy preserving and secure robust federated learning: A survey," Concurrency and Computation: Practice and Experience, vol. 36, Art. no. e8084, 2024.
- [18] Dong, N., J. Sun, Z. Wang, S. Zhang, and S. Zheng, "Flock: Defending malicious behaviors in federated learning with blockchain," arXiv preprint arXiv:2211.04344, 2022.
- [19] Zhang, Y., D. Zeng, J. Luo, Z. Xu, and I. King, "A survey of trustworthy federated learning with perspectives on security, robustness and privacy," in Companion Proc. ACM Web Conf. 2023, pp. 1167–1176, 2023.
- [20] Sharma, P., F. E. Shamout, and D. A. Clifton, "Preserving patient privacy while training a predictive model of in-hospital mortality," arXiv preprint arXiv:1912.00354, 2019.
- [21] Farrand, T., F. Mireshghallah, S. Singh, and A. Trask, "Neither private nor fair: Impact of data imbalance on utility and fairness in differential privacy," in Proc. 2020 Workshop on Privacy-Preserving Machine Learning in Practice, pp. 15–19, 2020.