

Federated deep learning for early detection and progression modelling of diabetic nephropathy

Aravind Kumar Kyrika, Sumit Kumar Banchhor

Amity Institute of Information Technology, Amity University Chhattisgarh, Raipur, India

Article Info

Article history:

Received Dec 1, 2025

Revised May 26, 2026

Accepted Jul 5, 2026

Keywords:

Diabetic nephropathy

Differential privacy

Federated learning

Healthcare AI

Multi-modal deep learning

Secure aggregation

ABSTRACT

One of the main causes of end-stage renal disease is diabetic nephropathy (DN). Patient privacy-sensitive data is a restriction in centralized machine-learning approaches for the healthcare domain. This research introduces a privacy-preserving federated deep-learning (Fed-DN) method for early diagnosis of DN and modelling disease progression and its impact. Fed-DN combines safe aggregation with differential privacy (DP) using Opacus and multi-modal data sources (electronic health records, fundus imaging, and histopathology). Both simulation tests with multi-modal data sets distributed across hospitals show an efficient privacy-utility trade-off, with Fed-DN achieving near-centralized performance ($AUC \approx 0.90$) (Accuracy = 87%) while maintaining a privacy budget of $\epsilon \approx 3.1$. This is a good example of the trade-off achieved between model accuracy and privacy protection. The technique does not reveal patient information, which enables collaborative, scalable healthcare analytics. Furthermore, it enhances trust among patients and institutions.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Sumit Kumar Banchhor

Amity Institute of Information Technology, Amity University Chhattisgarh

Raipur, Chhattisgarh, India

Email: skbanchhor@rpr.amity.edu

1. INTRODUCTION

Diabetic nephropathy (DN) is a prevalent and severe diabetes mellitus complication and a major cause of chronic kidney disease (CKD) and end-stage renal disease (ESRD) [1], [2]. Machine learning (ML) holds out the potential for earlier diagnosis and more precise progression modeling, but centralized training is usually impractical due to privacy laws like the health insurance portability and accountability act (HIPAA) and general data protection regulation (GDPR) [3], [4]. Federated learning (FL) allows collaborative model training across hospitals without sharing raw patient records [5], [6]. However, real-world deployment is confronted with challenges like non-independent and identically distributed (IID) local data, communication limitations, and privacy leakage via model updates [7], [8]. This paper introduces federated deep-learning (Fed-D, which is a federated system that mixes differential privacy (DP) stochastic gradient descent (DP-SGD) with secure aggregation [9]-[11] to limit leakage, and personalization methods to address heterogeneity, used for multi-modal DN prediction and progression modeling.

2. THE PROPOSED WORK

FedAvg and communication-aware FL algorithms are the backbone of cross-institutional ML applications [4]-[6]. For instance, FL has successfully been applied to imaging tasks such as brain tumor segmentation [7], medical diagnosis [12], [13] and modelling of electronic health records (EHR) [14].

Other privacy-preserving techniques, like secure aggregation [9], [11] and DP [10], [15], [16] have been applied to FL to protect against membership inference [17], [18] and reconstruction attacks. While the integration of tabular EHR with imaging (fundus, histopathology) increases diagnostic accuracy [13], [19], [20], multi-modal FL is still not fully explored and is the gap that Fed-DN will address.

Algorithm problem formulation:

- Input: Clinical text data from various hospitals without compromising patient data.
- Pre-processing: Clean clinical texts and tokenize them to obtain relevant clinical features.
- Local training: The model is trained locally at each hospital with DP-SGD, gradient clipping, and Gaussian noise for privacy.
- Federated aggregation: Securely merges the noisy model updates from the participating hospitals
- Personalization: A proximal term takes care of non-IID data and generates a personalized model for every hospital, while retaining global knowledge.

3. METHODS

Let S hospitals participate, each with a private dataset

$$D_s = \{(x_i^{(s)}, y_i^{(s)})\}_{i=1}^{n_s}$$

The global empirical objective is

$$F(w) = \sum_{s=1}^S \frac{n_s}{N} F_s(w)$$

$$F_s(w) = \frac{1}{n_s} \sum_{i=1}^{n_s} \ell(f_w(x_i^{(s)}), y_i^{(s)})$$

where w denotes model parameters and ℓ the per-example loss (e.g., cross-entropy).

Clients perform DP-SGD with gradient clipping norm C and Gaussian noise:

$$\widetilde{\Delta w}_s = \Delta w_s + \mathcal{N}(0, \sigma^2 C^2 I)$$

The server aggregates noisy updates from participating clients S_t at round t :

$$w_{t+1} = w_t + \eta \frac{1}{|S_t|} \sum_{s \in S_t} \widetilde{\Delta w}_s$$

To improve personalization under heterogeneous distributions, we include a proximal personalization term:

$$\min_{w_s} F_s(w_s) + \lambda \|w_s - w\|_2^2$$

3.1. System overview

As shown in Figure 1, Fed-DN has three major components: First, a central server that coordinates communication rounds, sets up secure aggregation, and maintains the global model; Second, clients, namely, the hospitals that train a local DP-SGD model on their own multi-modal patient data and optionally fine-tune a personalised model head to reflect the data heterogeneity of each hospital; Third, a common data catalogue that provides consistent feature mappings, data representation, and data preprocessing across all clients.

3.2. Multi-modal model architecture

Figure 2 illustrates a multimodal healthcare system integrating EHR data from clinical text, medical imaging, and patient lab information. The outputs of this layer include imaging and diagnosis processed from these inputs and sent to a fusion layer. The fusion layer combines all of the information to create an integrated diagnosis and summary.

Encoders:

- EHR encoder: Dense or transformer-based encoder for tabular/time-series features of clinical text data, medical imaging, and patient lab data [14].
- Fundus encoder: convolutional neural network (CNN) backbone (e.g., ResNet) for retinal fundus images [19].
- Histopathology encoder: CNN or patch-transformer for tissue patches [13].

Latent vectors $z_{\text{ehr}}, z_{\text{fundus}}, z_{\text{histo}}$ are concatenated and fused:

$$z_{\text{fuse}} = \text{ReLU}(W_f[z_{\text{chr}}; z_{\text{fundus}}; z_{\text{histo}}] + b_f)$$

Final prediction:

$$\hat{y} = \sigma(W_s z_{\text{fuse}} + b_s)$$

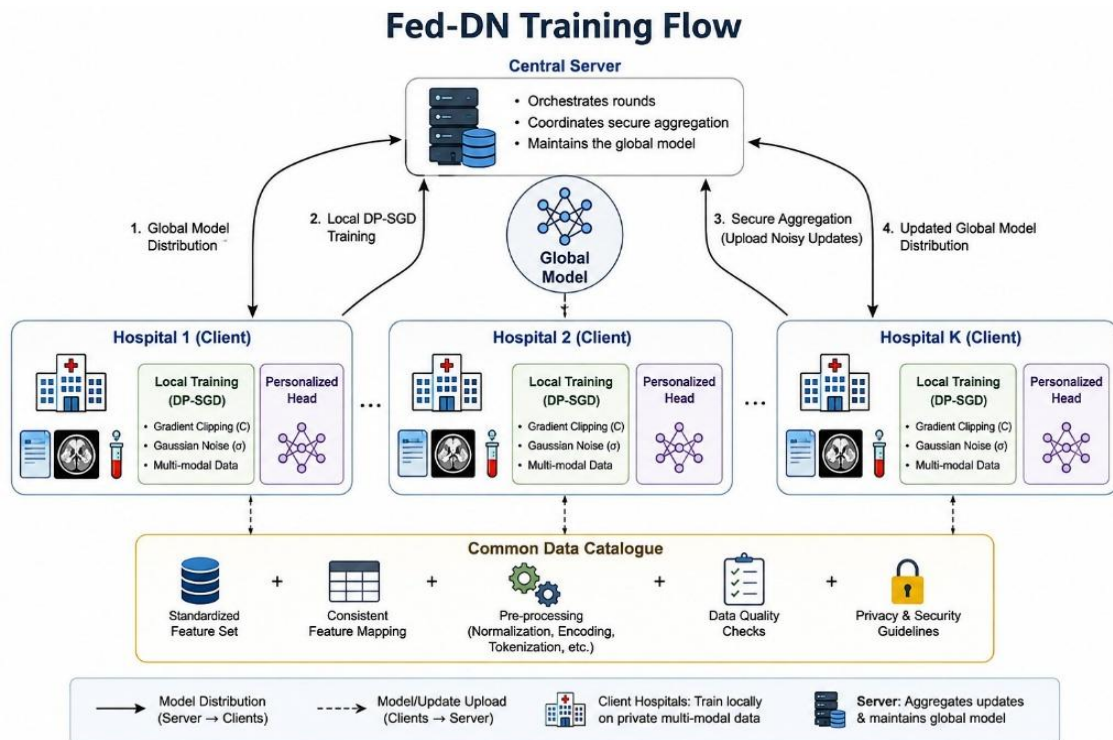


Figure 1. FL architecture

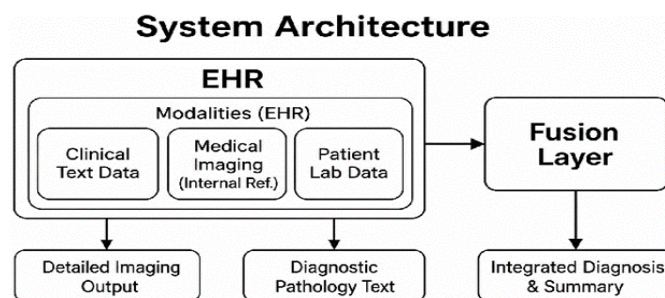


Figure 2. The complete system architecture diagram

3.3. Privacy and security mechanisms

Figure 3 displays the privacy threats and mitigation strategies in ML. Model inversion attempts to extract sensitive data from a trained model. Membership inference refers to whether a specific data record is used for training. DP helps safeguard individual data, and secure aggregation helps ensure that data remains private during model training.

- DP: Per-example or per-sample gradient clipping and Gaussian noise addition (DP-SGD) with Opacus for privacy accounting [5], [21], [22].
- Secure Aggregation: Cryptographic protocols to aggregate encrypted client updates so the server cannot observe individual updates [9], [11].

Combined, DP and secure aggregation mitigate gradient-inversion and membership-inference attacks [17], [18].

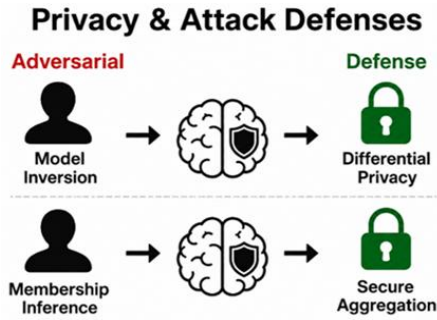


Figure 3. Privacy attack and defense diagram

3.4. High-level algorithm

This algorithm is called Fed-DN and is composed of a central server and participating hospitals. The server initializes the global model, picks a set of clients for each communication round, sends out the global model, gets back noisy updates that have been encrypted, safely combines them, and updates the global model. Local DP-SGD is implemented on each hospital's private patient data. The norm C is used to clip gradients, and Gaussian noise with the multiplier σ is added for privacy. It is possible to fine-tune optional personalization to local data with the model. Lastly, the noisy update is sent to the server, but it is encrypted for secure aggregation.

Server: initialize w_0 . For $t = 0, \dots, T - 1$: sample subset S_t , send w_t , receive encrypted noisy updates Δw_s , aggregate and update w_{t+1} .

Client s : receive w_t , perform DP-SGD with clipping norm C and noise multiplier σ , optional personalization fine-tuning, send encrypted Δw_s .

4. RESULTS

The results shown in Table 1 demonstrate that Fed-DN outperforms the local-only and FedAvg models, and performs close to the centralized CNN. Fed-DN achieves near-centralized performance while providing formal DP guarantees and improved adaptation to non-IID client distributions via personalization [23], [24]. The small accuracy gap vs centralized models arises from DP noise and simulated data limitations. Algorithm 1 demonstrate Combined DP + Personalization is observed to produce near-centralized performance (AUC ≈ 0.90) (Accuracy = 87%) while maintaining a privacy budget of $\epsilon \approx 3.1$. In general, the performance of the models can be enhanced by combining modalities.

Table 1. Summary of experimental performance (simulated data)

Method	DP	Personalization	Privacy Budget (ϵ)	AUC	Accuracy
Centralized CNN	No	No	–	0.92	89%
Centralized CNN + DP	Yes	No	3.8	0.89	86%
Local-only	No	Yes	–	0.78	75%
Local-only + DP	Yes	Yes	3.5	0.75	72%
FedAvg	No	No	–	0.88	86%
FedAvg + DP	Yes	No	3.2	0.85	83%
Fed-DN	No	Yes	–	0.92	89%
Fed-DN (DP + Personalization)	Yes	Yes	3.1	0.90	87%

Algorithm 1. Opacus DP-SGD code sketch

```

for batch in dataloader:
    optimizer.zero_grad()
    loss = criterion(model(batch x), batch y)
    loss.backward()
    torch.nn.utils.clip_grad_norm_(model.parameters(), 1.0) # clip
    add_noise(model.parameters(), sigma=1.0) # Opacus handles per-sample noise
optimizer.step()

eps = privacy_engine.accountant.get_epsilon(delta=1e-5)
# Privacy parameters: batch_size=32, epochs=2, sigma=1.0 -> epsilon  $\approx$  3.1

```

Multi-modal DN prediction can be achieved with a workable privacy-utility trade-off for Fed-DN. The use of simulated data instead of actual multi-institutional cohorts, the engineering and communication overhead for actual deployments [4], and the accuracy degradation brought on by DP noise are some of the limitations. Longitudinal modelling and time-to-event analysis, interpretability for clinical uptake [20] and pilot multi-center deployment [3], and fairness and bias assessment [25] are future directions.

5. CONCLUSION

Finally, the proposed Fed-DN framework shows the potential of a privacy-preserving federated deep learning model for the early diagnosis and modelling of the progression of diabetic nephropathy. The framework incorporates federated learning, DP based on Opacus, secure aggregation, and multi-modal healthcare data to facilitate hospitals in jointly building a model without exposing patients' sensitive data. The simulation results show that Fed-DN performs nearly centrally, with an AUC of about 0.90 and a privacy budget of $\epsilon \approx 3.1$, balancing predictive performance and privacy protection. As such, Fed-DN offers a scalable and secure solution for collaborative healthcare analytics, tackling the critical issues of data privacy and security. Future research may involve the extension of the validation to real-world clinical settings, dealing with non-IID and imbalanced data, personalisation between hospitals, and optimizing communication and privacy costs.

ACKNOWLEDGEMENT

The authors are thankful to the Department of Information Technology for providing the tools and facilities to successfully conduct this research. We thank clinical nutrition and nephrology experts for domain guidance and the FL community for foundational methods.

FUNDING INFORMATION

The authors state that no funding was involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Aravind Kumar Kyrika		✓	✓		✓	✓		✓	✓					
Sumit Kumar Banchhor	✓						✓			✓	✓	✓	✓	

C : **C**onceptualization

M : **M**ethodology

So : **S**oftware

Va : **V**alidation

Fo : **F**ormal analysis

I : **I**nterpretation

R : **R**esources

D : **D**ata Curation

O : **O**riginal Draft

E : **E**diting

Vi : **V**isualization

Su : **S**upervision

P : **P**roject administration

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

Data availability does not apply to this paper as no new data were created or analyzed in this study.

REFERENCES

- [1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, vol. 54, pp. 1273–1282, 2017.
- [2] P. Kairouz *et al.*, "Advances and open problems in federated learning," *Foundations and trends® in machine learning*, vol. 14, no. 1–2, pp. 1–210, 2021.
- [3] K. Bonawitz *et al.*, "Practical secure aggregation for privacy-preserving machine learning," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, Oct. 2017, pp. 1175–1191, doi: 10.1145/3133956.3133982.
- [4] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–487, Aug. 2014, doi: 10.1561/04000000042.

- [5] M. Abadi *et al.*, “Deep learning with differential privacy,” in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, Oct. 2016, pp. 308–318, doi: 10.1145/2976749.2978318.
- [6] I. Mironov, “Rényi differential privacy,” in *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, Aug. 2017, pp. 263–275, doi: 10.1109/csf.2017.11.
- [7] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, “Multi-institutional deep learning modeling without sharing patient data: a feasibility study on brain tumor segmentation,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Springer International Publishing, 2019, pp. 92–104.
- [8] N. Rieke *et al.*, “The future of digital health with federated learning,” *npj Digital Medicine*, vol. 3, no. 1, 2020, doi: 10.1038/s41746-020-00323-1.
- [9] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi, “Federated learning of predictive models from federated Electronic Health Records,” *International Journal of Medical Informatics*, vol. 112, pp. 59–67, Apr. 2018, doi: 10.1016/j.ijmedinf.2018.01.007.
- [10] R. Shokri and V. Shmatikov, “Privacy-preserving deep learning,” in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, Oct. 2015, pp. 1310–1321, doi: 10.1145/2810103.2813687.
- [11] M. Fredrikson, S. Jha, and T. Ristenpart, “Model inversion attacks that exploit confidence information and basic countermeasures,” *Proceedings of the ACM Conference on Computer and Communications Security*, vol. 2015–Octob, pp. 1322–1333, 2015, doi: 10.1145/2810103.2813677.
- [12] K. Bonawitz *et al.*, “Towards federated learning at scale: System design,” *Proceedings of Machine Learning and Systems (MLSys)*, vol. 1, pp. 374–388, 2019.
- [13] A. Yousefpour *et al.*, “Opacus: User-friendly differential privacy in PyTorch,” *arXiv preprint arXiv:2109.12298*, 2021.
- [14] V. Gulshan *et al.*, “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” *JAMA*, vol. 316, no. 22, p. 2402, Dec. 2016, doi: 10.1001/jama.2016.17216.
- [15] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [16] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nature Medicine*, vol. 25, no. 1, pp. 24–29, Jan. 2019, doi: 10.1038/s41591-018-0316-z.
- [17] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated machine learning: Concept and applications,” *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, pp. 1–19, Jan. 2019, doi: 10.1145/3298981.
- [18] M. J. Sheller *et al.*, “Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data,” *Scientific Reports*, vol. 10, no. 1, 2020, doi: 10.1038/s41598-020-69250-1.
- [19] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: challenges, methods, and future directions,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, May 2020, doi: 10.1109/MSP.2020.2975749.
- [20] N. Papernot, M. Abadi, Ú. Erlingsson, I. Goodfellow, and K. Talwar, “Semi-supervised knowledge transfer for deep learning from private training data,” *International Conference on Learning Representations (ICLR)*, 2017.
- [21] A. Z. Tan, H. Yu, L. Cui, and Q. Yang, “Towards personalized federated learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 12, pp. 9587–9603, Dec. 2023, doi: 10.1109/TNNLS.2022.3160699.
- [22] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, “Federated learning for healthcare informatics,” *Journal of Healthcare Informatics Research*, vol. 5, no. 1, pp. 1–19, Mar. 2021, doi: 10.1007/s41666-020-00082-4.
- [23] M. Kitada, Y. Ogura, I. Monno, and D. Koya, “A low-protein diet for diabetic kidney disease: Its effect and molecular mechanism, an approach from animal studies,” *Nutrients*, vol. 10, no. 5, p. 544, Apr. 2018, doi: 10.3390/nu10050544.
- [24] X. F. Li *et al.*, “Efficacy of low-protein diet in diabetic nephropathy: a meta-analysis of randomized controlled trials,” *Lipids in Health and Disease*, vol. 18, no. 1, p. 82, Dec. 2019, doi: 10.1186/s12944-019-1007-6.
- [25] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, “Membership inference attacks against machine learning models,” in *2017 IEEE symposium on security and privacy (SP)*, IEEE, May 2017, pp. 3–18, doi: 10.1109/SP.2017.41.

BIOGRAPHIES OF AUTHORS



Aravind Kumar Kyrika     is working as an Assistant Professor in Computer Science and has more than 18 years of teaching experience and 2 years of Industry Experience. Ph.D. pursuing at Amity University Raipur. He can be contacted at email: aravindkumarkyrika@gmail.com.



Sumit Kumar Banchhor     received his Ph.D. from the National Institute of Technology (NIT), Raipur, India, in 2018. He is currently serving as an Assistant Professor in Amity Institute of Information Technology at Amity University Chhattisgarh, India. With more than nine years of teaching and research experience, his academic and research interests span biomedical engineering and electronics, with a focus on the Internet of Things (IoT), soft computing and machine learning applications, image processing, and speech processing. He has published more than 30 research articles in reputed international journals. In addition, he is the author of a book and holds nine intellectual property rights (IPRs), including patents and copyrights. He can be contacted at email: skbanchhor@rpr.amity.edu.