

FedMobileNet: A Resource-Conscious Federated Learning Framework for Imbalanced Lung Cancer CT Image Classification

Dr. Rituparna Bora 

Department of Computer Science,

Sibsagar University, India



r264porna@gmail.com

<https://orcid.org/0009-0009-6453-9367>

Dr. Ranjan Sarmah* 

Department of Computer Science,

Sibsagar University, India



rsarmah@sibsagaruniversity.ac.in

<https://orcid.org/0000-0003-4194-2822>

Kuntala Boruah 

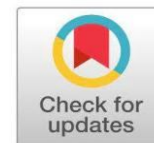
Department of Computer Science,

Sibsagar University, India



kuntalab@sibsagaruniversity.ac.in

<https://orcid.org/0000-0002-8936-7145>



Publication History: Manuscript Reference No: IRJCS/RS/Vol.13/Issue10/OCCS10080

Research Article | Open Access | Double-Blind Peer-Reviewed | Article ID: IRJCS/RS/Vol.13/Issue 10/CS.OCCS10080

Received: 11, September, Revised: 26, September 2026, Accepted: 02, October 2026, Published Online: 15, October 2026.

<https://www.irjcs.com/volumes/Vol13/iss-10/01.OCCS10080.pdf>

Article Citation: Rituparna, Ranjan*, Kuntala (2026), FedMobileNet: A Resource-Conscious Federated Learning Framework for Imbalanced Lung Cancer CT Image Classification, IRJCS: International Research Journal of Computer Science, Volume 13, Issue 10 of 2026 pages 652-656 **Doi:** <https://doi.org/10.26562/irjcs.2026.v1310.01>

BibTeX Key: Ranjan@FedMobileNet2026.

IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2026, ISSN 2393-9842, <https://doi.org/10.26562/irjcs.2026.v1310.01> The journal's official abbreviation is IRJCS.

ORCID: <https://orcid.org/0009-0004-9398-7488> About the License: Copyright©2026 copyright by the authors. This article is an open access and license under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Federated learning enables distributed training without exchanging original images, but non-IID data and class imbalance can degrade model performance. In this study, FedMobileNet is presented as a decentralized machine learning system based on MobileNetV2 for multiclass classification. The training instances were spread among four simulated clients using a Dirichlet-based non-IID partitioning method (with $\alpha = 1.0$) so that each client had at least five samples per class. The model used a multiclass focal loss with class weights set to $\alpha = [2.0, 1.0, 1.0, 1.0]$ and $\gamma = 2.0$; minority oversampling was not applied. The model that achieved the highest validation AUC at the checkpoint was selected and then evaluated once on the separate test dataset. In the test evaluation described here, FedMobileNet reached an accuracy of 89.4% and a macro-AUC of 0.97; the class-wise F1-scores were 0.81 for adenocarcinoma, 0.77 for large cell carcinoma, 0.97 for normal tissue, and 0.83 for squamous cell carcinoma. These results correspond directly to findings from a specific simulated federated configuration. They should not be interpreted as proof of clinical effectiveness or as evidence that the method is ready for deployment. The study uses a single public CT dataset, a single fixed random seed, and a specific client partition. It did not include analysis with repeated seeds and different partitions, external clinical validation, or direct measurements of memory, energy, communication cost, and inference latency. For this reason, the framework is described as resource-conscious rather than having been experimentally validated as resource-efficient for clinical or edge deployment.

Keywords: Federated Learning, Medical Image Classification, Class Imbalance, MobileNetV2, Lung CT Imaging.

I. INTRODUCTION

It is common in medical imaging to store data in a central repository to develop imaging models [1]. When sharing is not possible, medical images can be exchanged directly between different institutions as an alternative. In this approach, the shared model is optimized locally by the various clients, and only the model updates are sent to the central server for aggregation. Even though the first medical images remain at each institution, data confidentiality and information security are maintained. Risks may still arise because model updates are sent to the server [2]–[4]. The total count of images, the methods used to obtain them, and the disease ratios within medical image datasets can differ; the datasets may be non-IID or imbalanced, and when both occur simultaneously, the overall accuracy becomes difficult to interpret [3], [5]–[7]. Lung cancer remains a critical global health concern worldwide. Even though CT scans can be useful for evaluation once properly validated, automated classification remains difficult because the categories can vary widely and the scanned images do not consistently reflect the patient's true condition.

Given limited computing resources, compact models such as MobileNetV2 [8], [9], which have also been used in other image-classification and recognition systems [10], [11], are a reasonable option, and focal loss can make training more effective [12]. This problem affects people worldwide and can be evaluated using CT scans, provided the method is thoroughly evaluated [13]. Yet automated classification remains challenging because it must distinguish among categories, even when medical images may not fully reflect the patient's clinical condition.

Research has previously examined lung nodule detection, cancer subtype classification, attention-based models, and the application of federated learning to medical imaging [14]–[19]. Other work has examined class imbalance in the context of federated learning [3], [6] and in lung CT classification [20], as well as the resource constraints of edge and IoT devices [21]. While studies back these findings with a general approach, they cannot guarantee reliable performance across several institutions and devices [22]. In the present study, FedMobileNet is assessed using a consistent experimental setup when classifying CT images into four categories. The system builds upon an ImageNet-pretrained MobileNetV2 backbone. It uses multiclass focal loss, a Dirichlet-based method for distributing non-IID data among four simulated clients, sample-weighted FedAvg aggregation, and checkpoint selection based on validation AUC.

This work presents a detailed account of the proposed procedure and assesses its forecasting performance on an independent test dataset. It is not asserted that federated learning provides complete privacy, that the model is suitable for clinical deployment, or that it lowers computational or communication costs. The evaluation was conducted solely on the selected dataset, which comprised 613 training images, 72 validation images, and 315 test images. Only the training data were distributed to the client nodes. The experiment used a single random seed (42) and a single client-data split obtained from a Dirichlet distribution. The test results cited here pertain only to this particular dataset split and arrangement of clients.

II. METHODS

A. Dataset and Data Partitioning

The research used the Lung Cancer CT Scan Dataset [23], which included 613 images for training, 72 for validation, and 315 for testing. The four categories were ordered by the source directories: adenocarcinoma, large cell carcinoma, normal tissue, and squamous cell carcinoma. Among the training examples, only the ones were distributed to the four simulated clients. The validation and test sets were kept separate from the client shards and employed exclusively for model selection and final evaluation, respectively. The training images were resized to 192 by 192 pixels [24] and then scaled to the [0, 1] range. The client shards were generated using a Dirichlet-based non-IID allocation with a concentration parameter $\alpha = 1.0$, making sure that each client received at least 5 examples from each class. A fixed random seed of 42 was employed. Each of the original set of training images was assigned to exactly one client, with no duplicate file paths or oversampling of the less frequent classes. This controlled simulation does not do justice to the varied nature of data ownership but does not capture the full range of variability seen across independent hospitals.

B. Model Architecture

The feature extractor used a MobileNetV2, pretrained on ImageNet, with a width multiplier of 0.75 and an input size of $192 \times 192 \times 3$ [25]. In its implementation, the standard inverted residual blocks and linear bottlenecks were kept, and no custom Squeeze-and-Excitation module or different activation function was added. Once global pooling is done, there was a classification head consisting of a dropout layer (0.25), a ReLU-activated dense layer with 256 units, another dropout layer (0.20), and finally a four-unit softmax output layer. Fig. 1 shows a schematic overview of the implemented pipeline, rather than a measured system diagram.



Fig.1 FedMobileNet training and evaluation pipeline

C. Imbalance-Aware Loss

The α vector was the only class-weighting mechanism. No separate multiplier or oversampling step was applied. The experiment did not include an ablation study; therefore, the independent contribution of class weighting, focal modulation, or alternative imbalance strategies cannot be estimated from these results.

D. Federated Model Training and Selection

The model learned over 142 federated rounds across four simulated clients. In each round, clients performed local optimization on their respective shards, after which the server aggregated model parameters by sample-count-weighted FedAvg. The pretrained backbone was initially frozen. At round 31, the last ten MobileNetV2 layers were unfrozen for fine-tuning at a lower learning rate. Validation metrics were assessed periodically, and the checkpoint with the highest validation AUC was retained. The selected checkpoint was evaluated once on the held-out test set. The final evaluation generated accuracy, a classification report, a confusion matrix, and ROC/AUC summaries. We used a validation set for checkpoint selection rather than for final performance reporting. No repeated test evaluations or independent external test cohorts were reported.

E. Evaluation Scope

The study delivers an account of the predictive performance for the specified dataset and the simulated data partition. It did not include measurements of device memory usage, energy consumption, inference latency, training time, communication volume, or bandwidth. Even though MobileNetV2 is fairly compact, architectural efficiency alone is not sufficient to replace profiling the entire training and inference pipeline on the target hardware. Similarly, keeping images on local devices means there is no direct image transfer, but this does not constitute a proper privacy-preserving mechanism.

III. RESULTS

A. Held-Out Test Performance

The checkpoint that provided the best validation AUC, as shown in Fig. 2, achieved an accuracy of 89.4% and a macro AUC of 0.97 on the 315-image test set; the ROC analysis reported a micro-AUC of 0.96. The class-specific F1-scores were 0.81 for adenocarcinoma, 0.77 for large cell carcinoma, 0.97 for normal tissue, and 0.83 for squamous cell carcinoma. These figures show an uneven performance at the class level, normal tissue having the highest F1 score and large cell carcinoma the lowest. The test set consisted of 51 cases of adenocarcinoma, 54 of large cell carcinoma, 120 of normal tissue, and 90 of squamous cell carcinoma. Since the test set was assessed only once and no confidence intervals or repeated partitions were given, the point estimates do not provide a measure of either sampling or partition uncertainty. While the macro-AUC gives an overview of the ability to discriminate between the different classes, it should not be taken as evidence of calibration or clinical usefulness.

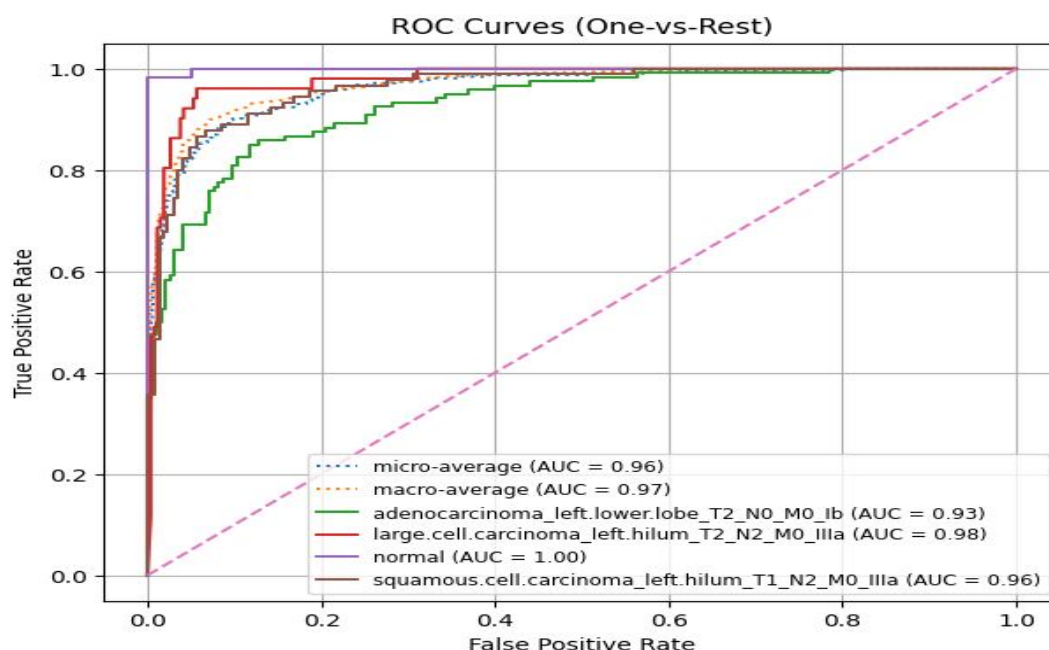


Fig.2 ROC curve.

B. Baseline Comparison

Table I gives the centralized baseline values for MobileNetV2 [9] and ResNet50 [26]. These values have not been retrained or recalculated independently in the present experimental run. Consequently, the comparison is merely descriptive and cannot be regarded as a controlled head-to-head benchmark. It is possible that differences in the data split, preprocessing, training budget, and selection procedures could affect the comparisons unless the baselines are rerun using the same protocol.

Table I - Baseline Model Comparison

Model	Accuracy	AUC (Macro)	F1 (Class 0)	Parameters
FedMobileNet	89%	97%	0.82	3.2 M
Centralized MobileNetV2	86.2%	87%	0.71	3.2 M
Centralized ResNet50	87.8%	89.1%	0.76	25.6 M

C. Training Configuration Summary

Table II summarizes the experimental configuration used in this study.

Table II - Experimental Configuration

Dataset split	613 train / 72 validation / 315 test images
Clients	4 simulated clients; Dirichlet $\alpha = 1.0$; minimum 5 per class/client
Input and backbone	$192 \times 192 \times 3$; Image Net-pretrained MobileNetV2, width multiplier 0.75
Loss	Multiclass focal loss; $\alpha = [2.0, 1.0, 1.0, 1.0]$, $\gamma = 2.0$
Federated optimization	Sample-count-weighted FedAvg; 142 rounds
Fine-tuning	Initially frozen backbone; last 10 layers unfrozen at round 31
Selection and test	Highest validation-AUC checkpoint; one held-out test evaluation
Reproducibility	Fixed seed 42; one predefined client partition

IV. DISCUSSION

The experiment demonstrates that a MobileNetV2-based model can learn in a simulated federated environment and achieves strong results on the designated test subset. Yet the findings support only this particular conclusion. The four clients correspond to simulated portions of one single public dataset and are not based on independently gathered datasets from different institutions. The experiment examines only the controlled differences across the client datasets and doesn't consider institutional differences in image acquisition, annotation procedures, or disease prevalence. Even if a model performs reliably in most cases, it may still perform badly on some less frequent classes [27]. Although it consistently achieves high accuracy, it still performs poorly compared with normal tissue for large-cell carcinoma. Although the method was meant to reduce class imbalance, its individual components were not evaluated separately. Further experiments are needed to determine which technique actually improves the model's performance. The class-specific results and the general figures (89.4% accuracy, 0.97 macro-AUC) are highly dependent on the test split and model checkpoint used, so changes to the test samples or the data distribution could affect them. Although the four classes have different numbers of test images (ranging from 51 to 120), the class-level outcomes have to account for these sample sizes, as metrics based on smaller test sets are less reliable. Future research ought to measure resource usage and compare the model with both centralized and federated models when using the same hardware. Federated learning reduces the need to transfer raw images but does not essentially prevent inference from model updates or defend against malicious participants [1], [3]. This study did not evaluate differential privacy, private aggregation, or poisoning defences. Its privacy contribution should therefore be described narrowly as data locality within the simulated training design, not as a formal privacy guarantee. Similarly, because the study did not include clinical users or prospective data, no diagnostic or treatment decision should be inferred from the reported classification metrics.

A. Limitations

The study has several important limitations: first, it uses a single public dataset with a single fixed train/validation/test split; second, it employs one random seed and one client partition; third, the clients are simulated rather than actual institutional clients; fourth, there is no component ablation nor any repeated-run uncertainty analysis; fifth, the baseline values are taken from earlier reporting in the manuscript rather than being rerun; sixth, there is no external clinical validation; and seventh, there is no direct privacy, security, or resource profiling.

B. Future Work

Use diverse data splits and metrics, compare established baselines, and evaluate each component's contribution.

V. CONCLUSION

The study introduced FedMobileNet, a classifier based on MobileNetV2 for analyzing four-class CT images, which was trained using sample-count-weighted federated averaging combined with focal loss in a simulated four-client non-IID environment. On the test set, the model achieved 89.4% accuracy and a macro-AUC of 0.97. Its F1-scores ranged from 0.77 to 0.97 across classes. These results apply only to this experiment and do not prove that the model is resource-efficient, clinically effective, privacy-protected, or ready for deployment.

VI. ACKNOWLEDGEMENT

All the authors would like to acknowledge Sibsagar University for providing the necessary facilities.

Author contribution Statement

Conceptualization: Dr.Rituparna Borah and Dr.Ranjan Sarmah
Literature Review and Methodology design: Dr.Rituparna Borah and Dr.Ranjan Sarmah
Formal analysis: Dr.Ranjan Sarmah
Investigation: Dr.Kuntala Boruah
Resources: Dr.Rituparna Borah
Data Curation: Dr.Rituparna Borah
Writing original draft preparation: Dr.Kuntala Boruah
Writing review and editing: Dr.Kuntala Boruah
Visualization: Dr.Ranjan Sarmah
Supervision: Dr.Kuntala Boruah
Project Administration: Dr.Ranjan Sarmah
All authors have read and agreed to the published version of the manuscript

Conflict of interest

The authors declare no conflicts of interest.

REFERENCES

1. J.Li et al., "A Systematic Collection of Medical Image Datasets for Deep Learning," ACM Comput. Surv., vol. 56, no. 5, pp. 1–51, 2023, <https://doi.org/10.1145/3615862>.
2. A.Brauneck et al., "Federated Machine Learning, Privacy-Enhancing Technologies, and Data Protection Laws in Medical Research: Scoping Review," J. Med. Internet Res., vol. 25, p. e41588, 2023, <https://doi.org/10.2196/41588>.
3. J.Zhang, C.Li, J.Qi, and J. He, "A Survey on Class Imbalance in Federated Learning," arXiv preprint arXiv:2303.11673, 2023, <https://doi.org/10.48550/arXiv.2303.11673>.
4. K.A.Kumar, J.S.A.Kumari, and A.K., "Secure Federated Learning Framework for Privacy-Preserving AI Models," Int. Res. J. Comput. Sci. (IRJCS), vol. 13, no. 3, 2026, <https://doi.org/10.26562/irjcs.2026.v1303.16>.

5. F.Zhang et al., "Recent methodological advances in federated learning for healthcare," *Patterns*, vol. 5, no. 6, Art. no. 101006, 2024, <https://doi.org/10.1016/j.patter.2024.101006>.
6. L.Weber, A. Hauschild, M. Altenbuchinger, U. Sax, and J. Hgel, "Effects of Non-IID Distributions in Lung Cancer Data on Survival Prediction with Federated Ensemble Learning," *Stud. Health Technol. Inform.*, vol. 290, pp. 116–121, 2026, <https://doi.org/10.3233/SHTI260116>.
7. A.Yogi and R. Dey, "Class Imbalance Problem in Data Science: Review," *Int. Res. J. Comput. Sci. (IRJCS)*, vol. 9, no. 4, Apr. 2022, <https://doi.org/10.26562/irjcs.2021.v0904.002>.
8. V.S.Negi and S.Chinara, "Study of MobileNets Model in Federated Learning," in *Proc. 2023 IEEE Int. Conf. Dependable, Autonomic and Secure Computing*, 2023, pp. 1040–1044, <https://doi.org/10.1109/DASC/PiCom/CBDDCom/Cy59711.2023.10361356>.
9. M.Sandler et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 4510–4520, <https://doi.org/10.1109/CVPR.2018.00474>.
10. C. Saranya, D. Anusha, H. K. Ganga, K. Brunda, and C. S. Bhagya, "Smart Waste Segregation System," *Int. Res. J. Comput. Sci. (IRJCS)*, vol. 12, no. 11, pp. 643–647, Nov. 2025, <https://doi.org/10.26562/irjcs.2025.v1211.02>.
11. K.Srikanth, B. Manasa, and M. Chanda, "A Hybrid Deep Learning Framework for Isolated Sign Language Recognition," *Int. Res. J. Comput. Sci. (IRJCS)*, vol. 13, no. 5, pp. 570–576, May 2026, <https://doi.org/10.26562/irjcs.2026.v1305.02>.
12. T.Y. Lin et al., "Focal Loss for Dense Object Detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2999–3007, <https://doi.org/10.1109/ICCV.2017.324>.
13. F.Bray et al., "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer J. Clin.*, vol. 74, no. 3, pp. 229–263, 2024, <https://doi.org/10.3322/caac.21834>.
14. S.El-Bana, A.Al-Kabbany, and M.Sharkas, "A Two-Stage Framework for Automated Malignant Pulmonary Nodule Detection in CT Scans," *Diagnostics*, vol. 10, no. 3, Art. no. 131, 2020, <https://doi.org/10.3390/diagnostics10030131>.
15. Y.Lu et al., "A hybrid CNN-RNN approach for survival analysis in a Lung Cancer Screening study," *Heliyon*, vol. 9, no. 8, Art. no. e18695, 2023, <https://doi.org/10.1016/j.heliyon.2023.e18695>.
16. P.Afshar et al., "3D-MCN: A 3D Multi-scale Capsule Network for Lung Nodule Malignancy Prediction," *Sci. Rep.*, vol. 10, Art. no. 7948, 2020, <https://doi.org/10.1038/s41598-020-64824-5>.
17. C.Saha et al., "Lung-AttNet: An Attention Mechanism-Based CNN Architecture for Lung Cancer Detection with Federated Learning," *IEEE Access*, vol. 13, pp. 57369–57386, 2025, <https://doi.org/10.1109/ACCESS.2025.3554744>.
18. K.A.S.Kumar, L.Nelson, and B.R.Jibinsingh, "Systematic review of privacy-preserving Federated Learning in decentralized healthcare systems," *Franklin Open*, vol. 13, Art. no. 100440, 2025, <https://doi.org/10.1016/j.fraope.2025.100440>.
19. M.A.Javed et al., "Abnormalities Detection in Chest CT Scan Images using Federated and Deep Learning," in *Proc. 19th Int. Conf. Emerging Technologies (ICET)*, 2024, <https://doi.org/10.1109/ICET63392.2024.10935071>.
20. M.Tabibian, T.Razmpour, and R.Saha, "Diffusion Models vs. DCGANs for Class-Imbalanced Lung Cancer CT Classification: A Comparative Study," preprint, 2025, <https://doi.org/10.1101/2025.11.19.689331>.
21. R.M.Rajendran, "Federated Learning for Edge Computing in REST APIs," *Int. Res. J. Comput. Sci. (IRJCS)*, vol. 11, no. 6, 2024, <https://doi.org/10.26562/irjcs.2024.v1106.02>.
22. I.Shiri et al., "Multi-institutional PET/CT image segmentation using federated deep transformer learning," *Comput. Methods Programs Biomed.*, vol. 240, Art. no. 107706, 2023, <https://doi.org/10.1016/j.cmpb.2023.107706>.
23. M.Hany, "Chest CT-Scan images dataset," Kaggle, 2020. <https://www.kaggle.com/datasets/mohamedhany/chest-ctscan-images>
24. A.Heidari et al., "A new lung cancer detection method based on the chest CT images using Federated Learning and blockchain systems," *Artif. Intell. Med.*, vol. 141, Art. no. 102572, 2023, <https://doi.org/10.1016/j.artmed.2023.102572>.
25. G.Inbasakaran and J.Anitha Ruth, "Clinical-ready CNN framework for lung cancer classification: Systematic optimization for healthcare deployment with enhanced computational efficiency," *Intell.-Based Med.*, vol. 12, Art. no. 100292, 2025, <https://doi.org/10.1016/j.ibmed.2025.100292>.
26. K.He, X.Zhang, S.Ren, and J.Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
27. E.Ferrante et al., "Class imbalance on medical image classification: Towards better evaluation practices for discrimination and calibration performance," *Med. Image Anal.*, 2024. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/38861161/>