Cervical Cancer Detection Using Federated and Few-shot Learning in Rural Health Centers

Kavita R. Singh[1], and Ajay Kumar[2,]

[1] Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur

[2] Adjunct Faculty and Post Doc Supervisor, Lincoln University College, Malaysia

***Correspondence:** singhkavita19@gmail.com

Abstract: Cervical cancer remains a major global health challenge, ranking as the fourth most common cancer among women, with India bearing a disproportionate burden in rural and underserved regions. Late diagnosis, weak referral systems, and inadequate screening infrastructure contribute to high mortality rates in such settings. Artificial intelligence (AI) has demonstrated potential in improving screening and diagnostics, yet conventional centralized approaches raise concerns of privacy, security, and regulatory compliance. Federated Learning (FL) addresses these limitations by enabling decentralized model training without transferring sensitive data, thus ensuring privacy and adaptability across heterogeneous sources. This review examines recent advances in FL for cervical cancer imaging, including self-supervised and few-shot learning, explainable AI, and privacy-preserving frameworks. Key challenges and research gaps are highlighted, with future directions focused on communication efficiency, rural deployment, and end-to-end clinical validation.

Keywords: Cervical Cancer, Federated Learning, Privacy

1. Introduction

Cervical cancer is the fourth most common cancer among women worldwide, with more than 530,000 new cases and over 280,000 deaths reported annually. In India, the burden is disproportionately high in rural and underserved regions, where late-stage diagnosis, inadequate referral systems, and insufficient screening infrastructure contribute to high mortality rates. Districts such as Gadchiroli, Wardha, Yavatmal, and Gondia in Maharashtra exemplify the challenges of limited access to timely screening and treatment, leading to preventable disease progression.

Artificial intelligence (AI) has shown significant promise in enhancing medical imaging, screening, and diagnostic workflows for cervical cancer. However, conventional AI approaches often rely on centralized data aggregation, which raises concerns around patient privacy, data security, and compliance with data protection regulations such as HIPAA, GDPR, and India's DISHA framework. Furthermore, cloud-dependent models are poorly suited for rural healthcare ecosystems, where connectivity, computational resources, and specialist availability are limited.

Federated Learning (FL) offers a paradigm shift by enabling decentralized training of AI models across multiple healthcare institutions without requiring sensitive patient data to leave local servers. This approach preserves privacy, enhances data security, and facilitates compliance with legal frameworks while enabling robust model development across heterogeneous datasets. When combined with emerging techniques such as few-shot and self-supervised learning, FL can potentially overcome challenges of data scarcity and non-IID distributions. Moreover, integrating explainability into FL frameworks is essential to foster clinician trust and ensure reliable deployment in life-critical healthcare settings.

This review consolidates recent advances in federated learning for medical AI, with a particular focus on cervical cancer imaging and cytology. It highlights the state of the art across domains such as FL-based cervical imaging, few-shot and self-supervised federated learning, explainable AI, centralized cervical imaging models, large-scale clinical AI deployments, and privacy-preserving technical innovations. The paper also identifies key research gaps and challenges—including data heterogeneity, communication overhead, privacy vulnerabilities, and limited rural deployments—before outlining potential directions for future work toward building clinically validated, explainable, and equitable AI solutions for cervical cancer screening.

2. Literature Review

FL permits decentralized model training without compromising on data privacy and hence it can be a paradigm shift in ML. Patient privacy, security threats, and adherence to data protection laws like HIPAA (Health Insurance Portability and Accountability Act, USA) and GDPR (General Data Protection Regulation, EU) are the shortfalls with conventional AI approaches as they aggregate healthcare data in centralized servers for model training. FL overcomes these shortfalls by sharing encrypted model updates rather than raw data and training models locally on hospital servers. This decentralized approach is very beneficial for sensitive industries like healthcare, where data security and confidentiality are critical.

This part presents a concise literature review to enhance comprehension of the proposed endeavor. Comprehensive literature has been extensively classified into Federated Learning in Cervical Cancer Imaging, Self-Supervised and Few-Shot Federated Learning, Explainability in Federated Medical AI, Centralized Cervical Imaging Models, Large-Scale Clinical AI Deployments, Cervical Cytology: Deep Learning and Fusion Networks, and Technical Federated Learning and Privacy Advances according to project specifications.

Federated Learning in Cervical Cancer Imaging: In year 2024, authors in [1] implemented FedCervical, a CNN-based federated learning system trained across decentralized PAP-smear datasets using FedAvg and LIME for explainability. In this paper authors demonstrated 95.1% classification accuracy under non-IID conditions. However, the proposed work lacks deployment in rural contexts; no few-shot adaptation; uses centralized dataset only. Later, Joynab et al. in [2] proposed a federated CNN for Pap-smear image classification. Though the authors achieved 94% accuracy in IID settings, not so appreciable performance of ~78% was demonstrated in non-IID. This shows that the proposed work had limited mitigation of non-IID data effects and there was no integration of XAI or few-shot learning.

Self-Supervised & Few-Shot Federated Learning: Authors [3] in 2025 introduced a federated, self-supervised, transformer-based pipeline for one-shot segmentation. The proposed techniques were demonstrated an Effective segmentation from <10 local labelled images, robust across domains. Yet, this work was not applied to cervical imaging and there was explainable AI. Dong et al. [4] presented FedFew, combining federated and few-shot learning to manage label scarcity in medical imaging. Paper demonstrated an improved performance in class-imbalanced, decentralized datasets with lacks deployment in explainable cervical models.

Explainability in Federated Medical AI: Research presented in [5] integrated causal explainable AI and blockchain auditing into federated cancer segmentation. Authors successfully demonstrated the enhancement of data traceability and interpretability in FL at the cost of high computation. Additionally, this latest work on explainable AI was not tested on cervix data. On the similar line, the work presented in [6] evaluated Grad-CAM for medical interpretability for reliable localization but for limited resolution and depth sensitivity. Alike [5] the work presented in [6] also not validated on cervical histopathology or VIA images in federated setups. The work presented in [21] introduced transformer-based models with unsupervised class activation mapping for histopathology to produce interpretable attention outputs. Later, researchers in [22] developed CAM-based models to reduce biopsies in low-resource cervical screening. The finding was that the proposed approach lowered unnecessary biopsies via explainable imaging. However, the major limitation was that both [21] and [22] lacked federated/few-shot deployment.

Centralized Cervical Imaging Models: The research presented in [7] and [8] basically emphasized on developing CNN model cervix classification. Techniques in [7] and [8] undoubtedly presented better accuracy greater than 90% on an average. However, both the techniques did not address the decentralized training or XAI with FL capability and interpretability features.

Large-Scale Clinical AI Deployments: Agency IARC [9] in 2024 developed an AI tool trained on 100k+ cervix images and deployed in LMICs. The deployed tool outperformed VIA and Pap in field trials. This tool is a proprietary one and not based on FL/few-shot and lacking clinician-level explanations.

Similarly, JMIR Editors [10] reported scalable cloud-based AI for cytology with expert oversight showing effective in large-scale deployments, however reliant on central cloud; not interpretable locally.

Technical FL & Privacy Advances: Under this specification of the project requirement authors in [11] used label-efficient, self-supervised federated learning on non-IID medical datasets with transformers, where in a significant improvement in generalization was observed. However, this work applied to dermatology/X-rays and not to cervix with lack of focus on explainable AI. Parekh et al. in [12] reviewed cross-domain federated segmentation in medical imaging to validate FL adaptability for diverse data but it was not implemented for cervical cancer.

In [13] a group of researchers surveyed FL applications across oncology, emphasizing ethical data sharing highlighting FL for privacy-preserving cancer diagnostics however real-world deployments were missing. Applied FL for metastatic lymph node risk prediction in cervical cancer across hospitals to enable multinational clinical-info aggregation and robust prediction was presented in [14].

Cervical Cytology: DL & Fusion Networks: Stegmüller et al. in [15] used self-supervised DL for cervical cytology triage in HPV-positive women showing boosted model learning with limited labels. However, the work was specifically demonstrated on centralized training with no context adaptation for FL/few-shot. Recently in 2025 authors in [16] presented 90% accuracy in segmentation/classification on the SIPaKMeD dataset with a limitation of not validating the method in decentralized or explainable frameworks. Similarly, the researchers in [17], [18], [19], and [20] presented work on cervical classification but no emphasis on federated learning or explainability and Patient privacy was also unaddressed.

4. Challenges and Research Gaps

Despite its transformative potential, federated learning in cervical cancer screening and medical AI still faces critical challenges that limit its wide-scale deployment—especially in rural and low-resource settings.

Data Heterogeneity and Non-IID Distribution

Variability in medical imaging, electronic health records (EHRs), and diagnostic reports across different healthcare institutions affects model consistency.

Studies [1] and [2] demonstrate diminished accuracy in non-IID contexts but exhibiting favourable outcomes in IID conditions.

Research presented in [3], [4], and [11] explore few-shot and self-supervised learning to address data scarcity and heterogeneity; however, these techniques have not yet been used in cervical imaging scenarios.

No previous research has implemented personalised federated learning or domain adaptation using VIA or Pap-smear datasets derived from several rural sources.

Limited Few Shot and Self-Supervised Deployments in Cervical Imaging

Few-shot and self-supervised FL have been explored for segmentation and classification tasks [3], [4], [15], [16], but:

Most implementations are either centralized or validated on generic medical datasets (e.g., X-rays, dermatology) rather than cervix images.

There is a research gap in designing robust FL pipelines with minimal labeled data specific to cervical cancer imaging (e.g., VIA, Pap smear, biopsy), particularly for rural settings.

Communication Overhead in Distributed Rural Networks

FL requires frequent model updates between decentralized nodes, leading to increased bandwidth consumption. There is a lack of lightweight, communication-efficient FL protocols tailored to low-connectivity environments.

Security and Privacy Vulnerabilities

FL models are susceptible to poisoning attacks, inference attacks, and adversarial manipulations, necessitating robust encryption and differential privacy mechanisms. Most reviewed systems do not implement differential privacy, secure multiparty computation, or blockchain-based auditing in cervical cancer applications. There is a need to explore lightweight cryptographic schemes for FL in resource-constrained clinical settings.

Limited Computational Resources in Rural Healthcare

Deploying FL models on edge devices in rural areas requires optimization to function on low-power hardware.

Lack of End-to-End, Real-World Validated Systems

While studies like [9], [10] report large-scale AI deployments, they are Cloud-dependent, not federated.

3. Conclusion

Federated Learning offers a transformative pathway for medical AI by enabling decentralized training while preserving patient privacy and complying with regulatory frameworks. Its application to cervical cancer screening has shown promise, but translation from experimental studies to practical, real-world systems remains limited. Current research demonstrates high potential in accuracy and privacy preservation, yet challenges in heterogeneity, resource constraints, communication efficiency, and explainability continue to impede widespread adoption.

Looking ahead, future work should focus on integrating few-shot and self-supervised paradigms within FL to overcome data scarcity, developing lightweight communication and cryptographic protocols for low-resource settings, and optimizing models for deployment on edge devices. Equally important is embedding explainability into federated pipelines to build clinical trust and reduce unnecessary interventions. Large-scale, multi-institutional validation in real-world healthcare ecosystems will be essential to move beyond prototypes toward clinically reliable and equitable solutions, especially in underserved regions.

References

[1] C. Kadam, S. Singh, and M. A. Khan, “FedCervical: Shedding Light on Cervical Cancer Detection with Federated Learning and Explainable AI,” Int. J. Med. Technol. Law Med., vol. 8, no. 2, pp. 12–21, 2024.

[2] N. S. Joynab et al., “A Federated Learning Aided System for Classifying Cervical Cancer Using Pap-Smear Images,” Informatics Med. Unlocked, vol. 44, Art. 101218, 2024.

[3] S. Manna et al., "Federated Self-Supervised Learning for One-Shot Cross-Modal Segmentation," arXiv:2503.23507, 2025.

[4] N. Dong, H. Yang, and M. Wang, "FedFew: Federated Few-Shot Learning for Decentralized Medical Imaging," arXiv:2206.15353, 2022.

[5] A. Mastoi et al., "Blockchain and Causal Explainability Enhanced Federated Learning for Tumor Segmentation," Front. Oncol., vol. 13, Art. 1535478, 2025.

[6] O. Suara, A. Mahapatra, and P. Kumar, "Evaluation of Grad-CAM for Interpretable Medical Image Diagnosis," J. Med. Imaging, vol. 11, no. 1, 2024.

[7] L. Zhang, Y. Guo, and M. Dong, "DeepPap: Deep Convolutional Networks for Pap-Smear Image Classification," IEEE Trans. Med. Imaging, vol. 38, no. 11, pp. 2489–2499, Nov. 2019.

[8] Y. Liu, C. Zhang, and T. Li, "CVM-Cervix: Lightweight Cervical Image Classification using Vision Transformers and CNN," arXiv:2206.00971, 2022.

[9] IARC, "IARC-developed AI Outperforms Standard Tests in a Leap Towards Global Health Equity," WHO News Release, Feb. 2024.

[10] JMIR Editors, "Performance of a Full-Coverage Cervical Cancer Screening Program Using Artificial Intelligence and Cloud-Based Cytology," J. Med. Internet Res., vol. 26, Art. e51477, 2024.

[11] R. Yan et al., "Label-Efficient Self-Supervised Federated Learning Improves Robustness of Medical Image Analysis Models," arXiv:2205.08576, 2022.

[12] V. S. Parekh, R. Joshi, and P. Kumar, "Cross-Domain Federated Learning in Medical Imaging: A Review," arXiv:2112.10001, 2021.

[13] L. Li et al., "Advances in Application of Federated Machine Learning for Oncology and Cancer Diagnosis," Inf., vol. 16, no. 6, Art. 487, 2024.

[14] M. C. van Leeuwen et al., "Federated Learning for Predicting Lymph-Node Metastasis in Early-Stage Cervical Cancer: A Multinational Study," Gynecol. Oncol., vol. 170, pp. 57–64, 2023.

[15] T. Stegmüller et al., "Self-Supervised Learning-Based Cervical Cytology for Triage of HPV-Positive Women," arXiv:2302.05195, 2023.

[16] A. S. Shaik, S. K. Sahu, and P. Mitra, “Deep Learning Enabled Segmentation, Classification and Risk Assessment of Cervical Cancer,” arXiv:2505.15505, 2025.

[17] N. K. Chauhan and P. S. Rathore, “Enhancing Cervical Cancer Diagnosis with Graph Convolution Network and Visual Feature Extraction,” Multimedia Tools Appl., vol. 83, pp. 14239–14261, 2024.

[18] S. Dasgupta and A. Ghosh, “Automated Cervical Cancer Screening Using Deep Learning: A Comprehensive Survey,” Electronics, vol. 12, no. 5, Art. 1042, 2023.

[19] L. Wong et al., “Deep Learning and Transfer Learning for Pap Smear Image Classification: A Review,” Int. J. Biomed. Eng., vol. 17, no. 4, pp. 205–213, 2023.

[20] A. S. Shaik, P. Mitra, and S. Sahu, “Fusion-Purification Network for Multi-Class Pap-Smear Image Classification,” Comput. Methods Programs Biomed., vol. 234, Art. 108199, 2024.

[21] A. Ghosh, S. Mitra, and S. Banerjee, “Unsupervised Class Activation Mapping for Cervical Cancer Histopathology Using Transformer Models,” IEEE J. Biomed. Health Inform., vol. 27, no. 2, pp. 589–598, 2023.

[22] T. Steg and A. Mahapatra, “Explainable CAM-Aided Cervical Image Classifier for Biopsy Reduction in Low-Resource Settings,” WIREs Data Mining Knowl. Discov., vol. 14, no. 2, e1517, 2024.