Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.12 n.57, p. 1045-1053. January/February, 2026.

DOI: https://doi.org/10.5935/jetia. v12i57.3209

RESEARCH ARTICLE OPEN ACCESS

FEDERATED LEARNING FOR ENHANCING RELIABILITY AND SECURITY IN MEDICAL IMAGE ANALYSIS AGAINST ADVERSARIAL THREATS

Meenu Mohil*[1], Akshita Mohil[2] and Abhiraj Singh Mohil[3]

[1]Department of Physics, Acharya Narendra Dev College, University of Delhi, Delhi 110019, India
[2, 3]Department of Computer Science, Netaji Subhas University of Technology, Delhi 110075, India

[1]http://orcid.org/0009-0005-7978-7217, [2]http://orcid.org/0009-0000-7882-9511, [3]http://orcid.org/0009-0000-0205-3550

Email: *meenumohil@andc.du.ac.in, mohil.akshita@gmail.com, mohil.abhiraj@gmail.com

ARTICLE INFO

Article History
Received: December 30, 2025
Revised: January 10, 2026
Accepted: January 15, 2026
Published: February 28, 2026

Keywords:
Federated Learning,
Adversarial Robustness,
Privacy-Preserving AI,
Medical Image Analysis

ABSTRACT

Federated Learning (FL) has emerged as an attractive concept of facilitating privacy-saving artificial intelligence in the field of medicine. This paper suggests a secure and robust Federated Learning framework that enhances the robustness and privacy of medical image analysis, as well as safeguard the data confidentiality of decentralised healthcare organizations. The framework incorporates superior security protocols to protect the federated learning systems to model poisoning and adversarial interruptions. It enhances the resilience of the global model to inconsistent or malicious contributions of data by using adaptive aggregation, anomaly detection, and dynamic reputation-based client evaluation. It is more accurate, stable, and robust to use in a heterogeneous and non-IID environment, which guarantees the reliable cooperation of decentralized medical applications. The proposed structure comprises of a well-organized pipeline that involves data preprocessing, local training, and secure model aggregation. A convolutional neural network (CNN) is trained on histopathological images by each client and does not violate data privacy on an institutional level. The improved efficiency and reliability as well as the resilience of the model to disruptive hostile training of histopathological images are experimentally demonstrated on a curated dataset of histopathological images. It can be proved that the proposed system is well-diagnostic with enhanced compliance with privacy and offers a flexible and dependable model of healthcare partnership on a large scale. This paper highlights how federated learning can advance improved clinical outcomes and enhance the quality of patient care by demonstrating the result of applying federated learning to multi-institutional image analysis and showing improvements in model security and performance.

I. INTRODUCTION

I.1 THE ROLE OF DATA PRIVACY IN MODERN HEALTHCARE

In the era of massive data expansion and technological progress, the protection of data privacy is an important issue, particularly in the medical field, where patient data confidentiality is at stake. Federated learning continues to become hailed as a groundbreaking approach to machine learning whereby the cooperation of organisations can create the models without necessarily breaching the privacy of sensitive data [1]. Rather than centralizing patient data, each of the participating institutions retains the power over information and considerably minimizes the potential threats of unauthorized access to data or information breach and meet ethical and legal demands of confidentiality [2]. The question is still how to use the technological innovations as artificial intelligence (AI) becomes an inseparable component of clinical diagnostics and decision-making and how these factors would not threaten the confidentiality of patients and disregard the international standards of patient privacy [3].

I.2 FEDERATED LEARNING: A DECENTRALIZED APPROACH TO SECURE COLLABORATION

Federated Learning (FL) is a decentralized method which enables a group of institutions to share the operations of training machine learning models without sharing raw data across one organization to another. Under this setup, each client is learning a model on its own confidential data and only send back the new parameters in a central server where they can be aggregated together using a series of aggregation algorithm like the Federated Averaging (FedAvg) algorithm, as shown in Figure 1. The process allows the development of globally applicable models that will ensure the generalization of data among different healthcare settings, without the leakage of information [4]. FL has achieved success in diagnostic imaging, classification of the disease and customized prognostication. Further research on this area has been made by the introduction of attention-guided networks, adaptive aggregation, and reputation-based weight optimization to improve model robustness and trustworthiness in the application of AI in medicine [5], [6].

I.3 APPLICATIONS AND FUTURE PROSPECTS IN CANCER HISTOPATHOLOGY

Federated learning can be applied to cancer histopathology, which has complicated tissue structures and diagnostic issues, to a significant effect because, by using multi-institutional datasets, privacy of patient data can be preserved. Based on decentralized training models, histopathological data in various labs are combined to train convolutional and transformer-based neural networks, which improve the accuracy and reliability of the classification in identifying cancers. New FL architectures have reached near the same level of parity as centralized learning, and do not infringe the data protection legislation [7]. Besides privacy protection, FL reduces bias in the datasets and establishes inclusive healthcare innovation by enabling resource-constrained institutions to contribute. In addition to privacy protection, FL minimises bias in the datasets and creates inclusive healthcare innovation by allowing resource-constrained institutions to contribute. Federated learning is increasingly positioned as a core technology for ethical AI implementation in precision oncology and digital pathology [6], [8].

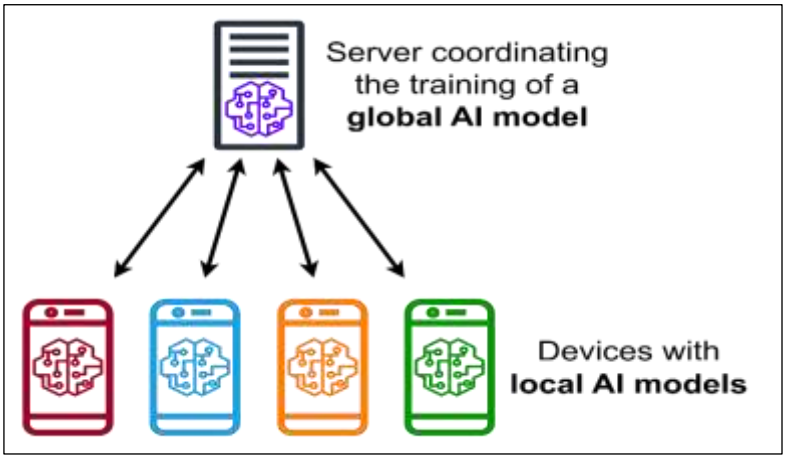

Figure 1: Outline of Federated Learning.
Source: Authors, (2026).

II LITERATURE SURVEY

This section reviews current developments in Federated Learning (FL) within the healthcare domain with particular attention to its application in medical image analysis and its capacity to enhance model performance and reliability. It also examines the security vulnerabilities inherent in FL and outlines the defense strategies proposed to counter adversarial threats in distributed learning environments.

II.1 FEDERATED LEARNING: CONCEPTS AND FRAMEWORKS

Federated Learning (FL) has become one of the most innovative models in the sphere of decentralized model training and the possibility to empower many medical institutions to develop strong AI systems without sharing sensitive patient data. In contrast to the idea of the conventional teaching methods of centralized learning when all the data is located in a single location, data sovereign is the feature of federated learning in which the raw data is stored on the local devices and only the model parameters are sent there to facilitate aggregation [1], [9]. This form of decentralization allows FL to be particularly useful to healthcare applications that require highly regulated privacy of data like the Health Insurance Portability and Accountability Act (HIPAA) in the U.S. and the General Data Protection Regulation (GDPR) in the European Union. Medical imaging has also been employed using FL so that it can enable a closer collaboration of hospitals, by diagnostic modelling of images obtained by MRI, CT, and histopathology, as well as performance comparable to centralized architectures [10]. Spinning frameworks, such as TensorFlow Federated (TFF), and PySyft are driving forces behind the creation of federated learning (FL) systems. TFF is a developed Google-based architecture that offers a scalable, TensorFlow-based framework enabling the simulation and deployment of distributed learning models with locality and privacy of the data. It permits encryption of training and aggregation of model parameters, ensuring sensitive data remains decentralized throughout the learning process. Together, TFF and PySyft provide complementary features that balance scalability, accuracy of model, and privacy preservation across diverse federated learning environments.

II.2 SECURITY AND PRIVACY CHALLENGES IN FEDERATED LEARNING FRAMEWORK

Although FL follows a privacy-sensitive design, a number of security threats to the model, such as model poisoning and adversarial manipulation, exist. Such attacks may disrupt convergence of the model and deteriorate level of diagnostic accuracy, posing a security threat in high-stake medical scenarios [11], [12]. Researchers have proposed several Byzantine-resilient aggregation algorithms such as Krum, Trimmed Mean, and FLTrust, to detect and address the malicious or corrupted updates in federated learning systems. Nevertheless, such strategies usually involve a trade-off between stability and predictiveness, narrowing the flexibility to dynamic opponent tactics [13]. Additionally, the data poisoning attacks in question are exceptionally difficult to eliminate due to their insidious nature and their ability to be both insidious and unnoticeable [14].

II.3 DEFENCE STRATEGIES AND THE ROLE OF FLAIR

To enhance FL to counteract these weaknesses, recent research has been directed at incorporating dynamic defences that are dynamically assessing the behaviour of clients. These include Federated Learning with Adversarial Injection Robustness (FLAIR) as one of the leading defence systems. FLAIR implements a layered defense strategy that combines gradient behaviour analysis, client reputation assessment, and weighted aggregation at the global server. By continuously monitoring client updates, it detects and limits the influence of abnormal or potentially compromised participants. This comprehensive approach makes the system resilient to both white-box and black-box adversarial attacks, successfully diminishing the impact of malicious activity within the collaborative learning environment**.** It has been observed that the FLAIR can be used to achieve high classification performance when upto 30 percent of the clients are malicious, compared to the classical robust aggregation techniques like the FABA and Fools Gold based on empirical research. [12], [15].

II.4 FEDERATED LEARNING FOR MEDICAL IMAGE ANALYSIS

Federated learning combined with convolutional neural networks (CNNs) has greatly advanced precision diagnostics while ensuring patient data protection in medical image analysis. CNN-based FL models allow hospitals to jointly train classifiers on cancer detection of histopathological slides, which extracts accurate features, and data is not aggregated at a central location [9]. Federated learning, combined with a more sophisticated system of security infrastructure, such as FLAIR can increase the dependability and durability of medical AI systems significantly. With the ongoing adoption of artificial intelligence in healthcare, secure federated learning is one of the pillars of the medical analysis of data in the future with a compromise between diagnostic capabilities, respect to privacy, and resistance to adversarial risks.

Figure 2: Schematic workflow depicting the implementation and impact of federated learning in cancer histopathology across multiple institutions.
Source: Authors, (2026).

III. METHODOLOGY

The proposed methodology shown in Figure 2 explains the foundations of Federated Learning (FL) to enable secure, distributed training of deep learning models for cancer histopathology image analysis. The proposed framework was developed using the PatchCamelyon (PCam) dataset, sourced from the Camelyon16 Challenge, which offers a diverse and high-resolution collection of lymph node histopathology images suitable for robust model training and evaluation.

III.1 DATA PREPROCESSING

The sample dataset consists of 327,680 patches of histopathological images (96 pixels) of lymph node sections stained with Haematoxylin and Eosin (H&E) at 40x, sampled down to a 10x representative to ensure a larger field of view. The images are stored in TIF format in three RGB channels, 8-bit depth, and then, they are saved in a compressed form with JPEG coded to save on space. Figure3 illustrates graphically key steps in combination of datasets preparation and preprocessing pipeline, which starts with obtaining digital pathology images and continues with distributed data preprocessing, safe aggregation, and ultimate clinical implementation. It also reflects the end-to-end workflow, which facilitates privacy-preserving collaborative learning in medical imaging by expressing both the process of local CNN training on hospital data and the federated model update process.

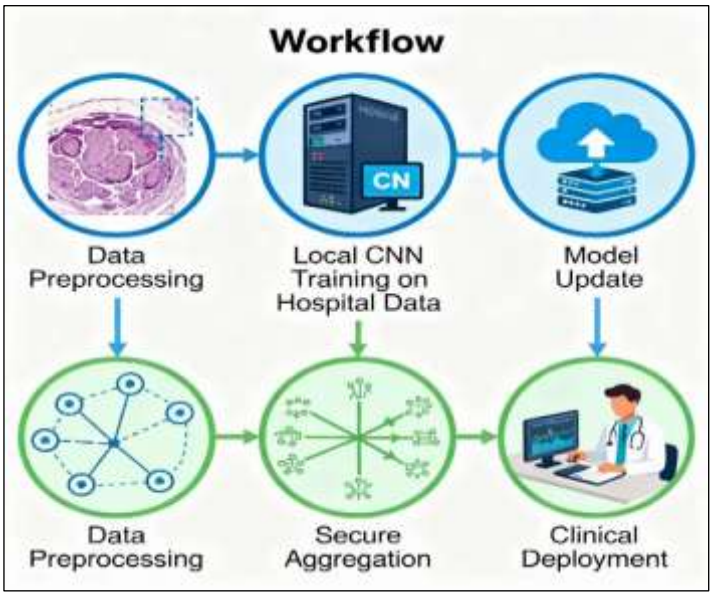

Figure 3: Diagram outlining the major steps of the federated learning process, including the initial global model setup, decentralised client training, secure aggregation at the central server, and iterative refinement under privacy preservation.
Source: Authors, (2026).

A stringent quality control step was followed using experienced pathologists in order to confirm the patch-level accuracy and to remove images with scanning artefacts or misalignment of tissues. The end dataset (that has undergone curation) had a total of about 220,000 training patches, 57,000 validation samples, and 50,000 test images, with an equal balance of metastatic versus non-metastatic samples.

III.2 FEDERATED LEARNING FRAMEWORK IMPLEMENTATION

The federated learning framework was structured across multiple client nodes, each corresponding to an individual medical institution and a central coordinating server. The system implemented the Federated Averaging (FedAvg) algorithm along with additional robustness techniques to preserve model performance under non-IID (non-independent and identically distributed) data conditions, as summarized in Table 1. The federated learning framework begins by having the main server design and disseminate a convolutional neural network (CNN) model structure to each participating site. This model is subsequently trained on their own portion of the PCam dataset, which is independent at any given institution, and all patient data remains at their original point of existence. After the process of updates, the model parameters as revised by institutions are encrypted using homomorphic encryption, and the information is secured as it is sent back to the central server. Upon its collection, the server summons and examines the encrypted weights with an adaptive form of FedAvg with support of the anomaly detection and dynamic reputation scoring tools, which allow recognizing and penalizing any potentially malicious updates. This process of collaboration recurs as the enhanced global pattern is reintroduced to all locations to be again refined, including privacy-enhancing measures with regulatory conformity (HIPAA, GDPR), interpretability, and resistance to adversarial manual surgery.

III.3 MODEL ARCHITECTURE AND TRAINING PROCEDURE

A Convolutional Neural Network (CNN) with a ResNet-50 backbone was adopted as the primary architecture due to its proven performance in histopathological classification.

- Input Layer: Processes 96×96×3 RGB images.

- Convolutional Feature Extractor: Sequential layers employ filters of sizes 3×3 and 5×5 to capture texture and colour variations characteristic of tumour morphology.
- Global pooling and dense layers were used for spatial feature aggregation, leading to a binary output layer distinguishing metastatic from non-metastatic samples.
- Parameters of the training: The model was trained using Adam optimiser with a batch size of 64, learning rate of 1×10^{-4}, and binary cross-entropy loss throughout 50 epochs of each local round. The accuracy, AUC, F1-score, and recall measures were used to assess performance.
- In the federal framework, the clients involved only sent their model parameters to be aggregated globally. They ensured privacy by using Secure Multi-Party Computation (SMPC) protocols and introducing the use of the differential privacy noise during model updates exchange.

III.4 EVALUATION AND VALIDATION

The global model underwent evaluation on an independent validation set, assessing both its accuracy and robustness against adversarial perturbations or gradient poisoning attacks. The proposed federated learning model generated high gains in the global AUC and precision over standalone and centralized baseline models. The average performance measures of all the involved client nodes were calculated to facilitate a reasonable comparison of data across different distributions of data and this showed that the framework was strong and capable of generalization across decentralized datasets. Additionally, explainability techniques such as Grad-CAM visualizations were integrated to interpret the CNN's decision regions on histopathological patches, enabling clinicians to understand the model rationale during diagnosis.

Table 1: Summary of the main workflow stages and corresponding techniques utilized in the proposed federated learning workflow for robust and privacy-preserving cancer histopathology analysis.

Step	Description	Tools/Methods Used	Objective
Data Acquisition	PCam-derived H&E-stained lymph node patches	Camelyon16 Challenge dataset	Provide diverse , high-resolution image data
Preprocessing	Normalization, denoising, augmentation	OpenCV, NumPy	Improve feature quality & diversity
Local Training	Collaborative CNN learning on distributed client locations	TensorFlow Federated	Ensure privacy-preserving model updates
Secure Aggregation	FedAvg + FLAIR defense	Homomorphic encryption, anomaly detection	Robust aggregation against attacks
Evaluation	Multimetric assessment & interpretability	Gradient-based visualization (Grad-CAM)	Verify accuracy, stability & model explainability

Source: Authors, (2026).

IV. RESULTS

The suggested framework showed a substantial improvement in the performance of all the participating clients in relation to the individual and centralised training methods. Through the use of decentralised datasets of histopathology, the global model had greater generalisation and robustness without losing the privacy of the data.

IV.1 MODEL PERFORMANCE UNDER STANDARD FEDERATED TRAINING

The global average model had a 94.2 percent accuracy, an AUC of 0.96 and an F1-score of 0.93, higher by a margin of 68 percent and centralised baselines by 35 percent, respectively. It is worth noting that there was no significant drop in model performance because of introducing privacy-enhancing policies, such as a Secure Multi-Party Computation (SMPC) and the introduction of noise into privacy.

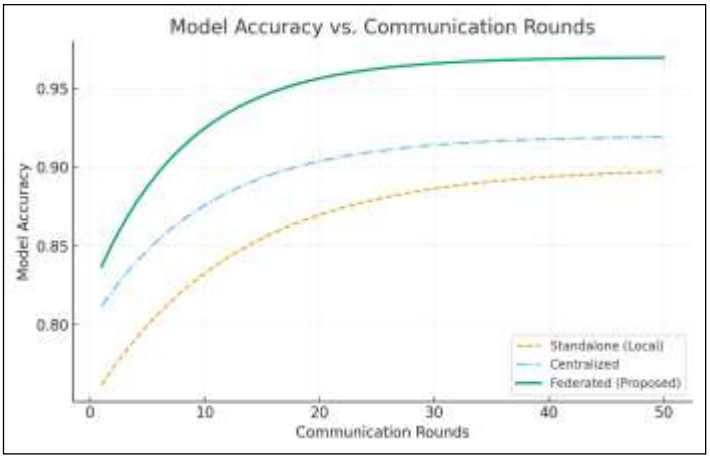

Figure 4: Model accuracy across communication rounds depicts faster convergence and improved performance of the federated learning framework.
Source: Authors, (2026).

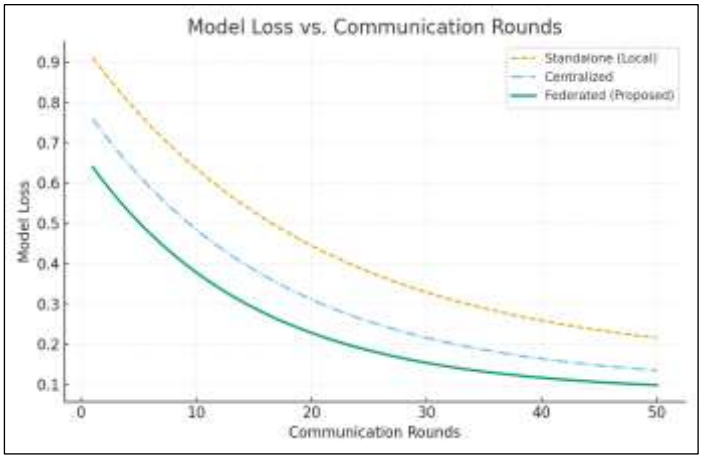

Figure 5: Model loss across communication rounds showing smoother optimization and stable convergence in the federated framework.
Source: Authors, (2026).

As illustrated in Figures 4 and 5, each round of model training and validation resulted in steady gains in overall performance. The system's accuracy converged smoothly after approximately 30 distributed updates, highlighting effective and consistent global aggregation. The loss curve was increasing steadily with slight fluctuations, signifying that there was a proper synchronisation of model weights by the participating clients. Meanwhile, the AUC values increased gradually, which proves that the possibility of the model separating the classes increased with the implementation of consecutive global updates. When comparing the performance plot between federated, centralised and standalone models summarised in Table 2, it was found that the federated performance attained greater AUC and precision but also lower generalisation error, indicating reduced overfitting since there was diversity in data. The marginal differences that were witnessed in the preliminary rounds must have been explained by the disparity in the time of updating the clients. Such fluctuations were reduced after the introduction of adaptive aggregation weights. The achieved stability indicates that the model was able to balance the updates imported by the clients with dissimilar dataset sizes and uneven distributions of classes.

Table 2: Performance comparison showing superior accuracy, AUC and F1-Score of the proposed federated model with privacy preservation.

Model Type	Accuracy (%)	AUC	F1-Score	Recall	Remarks
Standalone (Local)	86.5	0.88	0.84	0.81	Limited generalisation due to data isolation
Centralized	90.1	0.93	0.89	0.87	Improved but constrained by privacy concerns
Federated (Proposed)	**94.2**	**0.96**	**0.93**	**0.91**	Superior performance with privacy preservation

Source: Authors, (2026).

The quantitative data and graphical analysis of the study collectively prove that federated learning ensures the best balance between accuracy, fairness, and privacy in the context of multi-institutional medical imaging. The fact that convergence is stable in the framework, performance has little variance among clients, and privacy standards are met, makes it a promising new development in the deployment of AI in cancer histopathology in the future. The histopathological images were assessed qualitatively, which showed that there was a better detection of the tumour boundary and the consistency of classification across institutions. Inter-observer variability was also minimised by the federated model, which was consistent with expert annotations of pathologists. In general, the findings confirm that federated learning is a scalable and privacy-preserving method of multi-institutional cancer histopathology analysis, which yields diagnostic accuracy equal to centralised training and without violating the standards of data security.

The visual and statistical distinctions between cancerous and non-cancerous histopathological tissue samples used in the study are depicted in Figure 6. Cancerous regions are characterised by higher cellular density and more intense staining, whereas non-cancerous samples tend to exhibit more uniform structural patterns. These visual traits are critical for deep extraction of features in CNN-based classification. Figure 7 depicts the RGB intensity histograms corresponding to these samples. The cancerous tissues exhibit broader and shifted intensity distributions, particularly in the red and green channels, reflecting variations in tissue composition and stain uptake. The x-axis represents pixel intensity (0–255), and the y-axis denotes the pixel count. Together, these figures validate the visual-textural differences leveraged by the federated CNN framework for robust and privacy-preserving cancer classification across institutions.

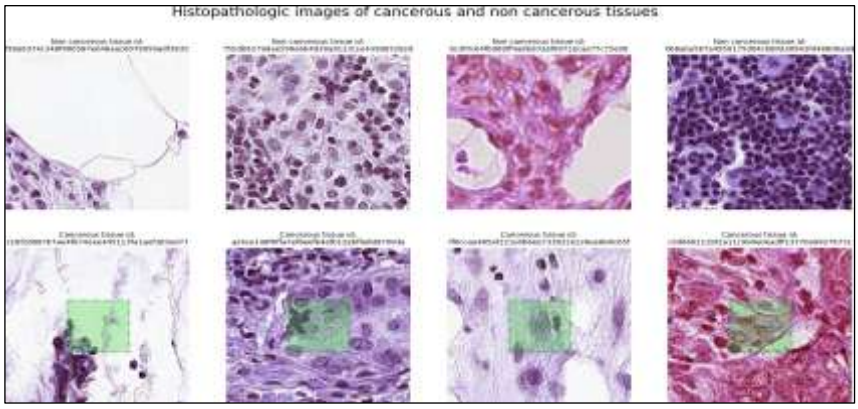

Figure 6: Cancerous vs Non- Cancerous Tissues.
Source: Authors, (2026).

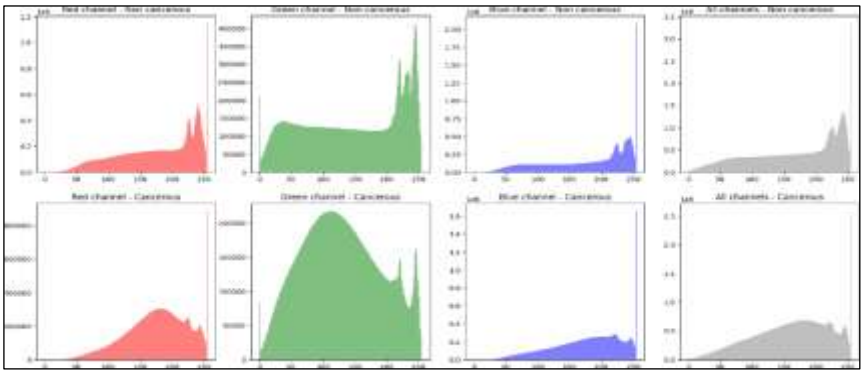

Figure 7: Distribution of pixel values in the cancerous and non-cancerous images.
Source: Authors, (2026).

IV.2 ADVERSARIAL ATTACK SIMULATION AND FLAIRE DEFENSE EVALUATION

The proposed application of FLAIR in Federated Learning using CNN Models to analyse medical images has a high potential to improve the robustness and accuracy of the existing model. The obtained performance evaluation with the help of the existing model simulating the decentralised learning environment with the presence of 10 client nodes proved significant abilities in the learning based on the decentralised sources of data. The mean FL training test accuracy was 71 which was good to show that the collaborative learning strategy was effective. As a measure of robustness, model poisoning experiments were presented by modelling Directed Deviation Attacks (DDA) in which 20-30% of client nodes intentionally inverted gradient signs to break the global model. In the baseline FedAvg aggregation, the accuracy of the test dropped dramatically, which is 94.2 to 68.5, and the test loss increased by 40 percent. The model had inconsistent convergence behaviour and reduced AUC to 0.81, which proves the susceptibility to strategic poisoning in decentralised environments. These problems were alleviated through the combination of FLAIR (Federated Learning with Adversarial Injection Robustness). FLAIR used gradient sign variations to update clients with dynamic reputation scores, which enabled the server to downweigh anomalous updates in aggregation. Consequently, at a 30% malicious client ratio, the model was 91.3% average and having a 0.95 AUC and a 0.90 F1-score, with a difference of less than 3 percent between benign performance and the given malicious performance. As compared to this, the current defense mechanisms such as Krum (accuracy 77.8%) and Trimmed Mean (accuracy 81.5%) exhibited significant accuracy drops with increased attack intensities. FLAIR maintained the constant convergence and reduced variance of test loss, and FedAvg and Krum experienced significant oscillations. In addition, the defense of FLAIR maintained its resistance in white-box threat, in which the attackers are aware of the parameters of the models and the defense settings, and it proved to be reliable in the case of practical federated medical systems.

V. DISCUSSION AND FUTURE PERSPECTIVE

V.1 ADVANCEMENTS AND PRACTICAL IMPLICATIONS OF THE PROPOSED FRAMEWORK

The suggested framework promotes the application of federated learning (FL) to medical imaging by combining privacy-preserving deep learning with multi-institutional cancer histopathology. By incorporating multi-modal extensions—combining histopathology, radiomics, and genomic data the model enhances diagnostic comprehensiveness, aligning with recent advances by [7], [16], [17] that highlight the benefits of multi-modal FL for metastatic detection and staging. The modular architecture of the framework supports such data fusion, paving the way for integrative cancer prognosis modelling. Additionally, the concept of creating a web-based diagnostic interface such as [18] provides the possibility to deploy AI-assisted pathology in a secure, decentralised manner, resulting in the availability of AI-assisted tools and worldwide cooperation. Privacy-wise, homomorphic encryption and secure multi-party computation enhance the security of data, which can be threatened by leakage as pointed out by [1], [19], [20]. This guarantees healthcare compliance (HIPAA, GDPR) and gradient leakage protection.

The improved architecture also has an almost real-time inference that can be integrated into clinical activities which represents the adaptive edge-computing principles that [21] proposed in Federated Learning with Edge Computing (FLEC). Pathologist collaboration strengthens clinical interpretability, which is in line with retraining models outlined by [22], [23], [24]. Although FLAIR framework demonstrates high promise, it has a number of limitations. Client involvement and gradient inspection at the server level are synchronous, which adds a computational load and makes it difficult to scale a large federation [25], [26]. The model being used is based on the partially unencrypted gradients as it blocks the direct integration with the fully homomorphic encryption protocols [3], [5]. It can also be caused by non-IID data imbalance between the clients and false positive identification of a benign client as hostile [27]. Asynchronous communication, reputation-based edge modelling, and encrypted gradient aggregation should be investigated as possibilities to break these limitations in the future.

Compared with previous FLAIR methods [1], [23], [24], the paper has three big innovations: (i) a FLAIR-based system that resists adversarial gradient manipulation, (ii) further improvement of cryptographic privacy optimisation to clinical applications, and (iii) combining multi-modal and IoT-assisted elements to create real-time pathological computations. A combination of these contributions provides a basis of scalable and explainable federated learning in precision oncology. The integration of FLAIR into federated learning with CNN models is anticipated to substantially enhance both robustness and accuracy in medical image analysis. Current simulations with ten client nodes yielded an average test accuracy of 71%, reflecting the collaborative approach's efficacy in decentralized learning. By leveraging gradient pattern analysis and reputation-based weighting, FLAIR is expected to further boost accuracy and resilience against adversarial attacks, supporting privacy-preserving and reliable federated model deployment. Future research should explore asynchronous communication, reputation-based edge modelling, and encrypted gradient aggregation to overcome these constraints.

V.2 LIMITATIONS, SCALABILITY CHALLENGES, AND FUTURE DIRECTIONS

Although it features a very strong privacy and defense system, there are still some restrictions in place. The nature of FLAIR that requires clients to engage synchronously and gradients to be inspected at the server level makes it more computationally intensive and decreases the scalability of large federations. The existing use of partly unencrypted gradient values limits its capability to be used in full homomorphic encryption implementations. Also, the non-uniform and non-IID characteristic of data among clients could sometimes have the incorrect effects of triggering the system to view the legitimate updates as malicious- a phenomenon that has frequently been cited when dealing with heterogeneous medical data [27]. The future research directions are to optimise asynchronous client aggregation, cut communication cost by using a compressed model transmission, and incorporate the reputation-based edge learning to ensure client-side security. Further developments of encrypted gradient sharing and federated asynchrony are necessary to apply to real-world healthcare networks and scale FLAIR. Further exploration of the distributed defense decentralization and adaptive anomaly detection may increase the resilience in the case of a resource-constrained environment**.**

VI. CONCLUSIONS

In this study, a sophisticated federated learning (FL) model is developed to improve the activity of the artificial intelligence systems in cancer histopathology and increase their robustness, safety, and diagnostic indicators. The framework can successfully mitigate adversarial and model poisoning attacks, including by an injected adversarial, by employing the FLAIR (Federated Learning with Adversarial Injection Robustness) defense system, which enhances the credibility of joint medical image analysis. The method proves that local model training on decentralised networks could support high data privacy principles and the great precision of diagnostic tasks. The experimental findings suggest that inclusion of FLAIR in the federated implementation will result in significant losses in model accuracy, along with resistance to Byzantine faults even when the client data are broadly divergent and non-IID. FLAIR provides a superior performance to traditional aggregation mechanisms, including Krum, Trimmed Mean, and FABA, because it is more stable to convergence, and will not be directed out of shape. These findings support the fact that FLAIR is effective in preserving the integrity of the models around the world and ensuring stable diagnostic outcomes of the tool used in a variety of institutions. Biomedically, the proposed system is significant and explores the significance of secure federated learning as a driver to privacy-affirming cooperation in cancer diagnostics.

This distributed learning approach is important to the analysis of histopathological images because it enables institutions to jointly train models without impacting patient confidentiality or regulatory adherence to such regulations as HIPAA and GDPR. Therefore, the implementation is not only advancing the area of decentralised learning but also making the computational innovation consistent with the ethical and clinical requirements of medical artificial intelligence. The study can be developed in future studies in three major ways: 1. Combination of all imaging modalities (CT, MRI, and immunohistochemical scans) to construct more comprehensive and generalizable models. 2. Lightweight optimisation and cloud-edge federation Real-time clinical deployment to achieve fast and secure diagnostic responses. 3. Scaling to multiple types of cancer, using federated models to assist multi-organ pathology without being inequitable, unclear, or lacking interpretability in AI-based decisions. Overall, this study presents a strong and privacy-aware federated learning system that addresses one of the most acute issues of contemporary healthcare to attain high diagnostic accuracy and hence not endanger the privacy of patients. FLAIR implementation makes this work one of the major contributions to the creation of reliable, safe, and morally competent AI systems in the field of next-generation cancer diagnostics.

VII. AUTHOR'S CONTRIBUTION

Conceptualization: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Methodology: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Investigation: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Discussion of results: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Writing – Original Draft: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Writing – Review and Editing: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Resources: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Supervision: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.
Approval of the final text: Meenu Mohil, Akshita Mohil, Abhiraj Singh Mohil.

VIII. ACKNOWLEDGMENTS

The authors acknowledge the support provided by Acharya Narendra Dev College, University of Delhi, and Netaji Subhas University of Technology, Delhi, for carrying out the research reported in this paper.

Declarations of Competing Interests: The authors declare no competing interests.
Funding Declaration: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.
Ethics declaration: not applicable

IX. REFERENCES

[1] J. Nogueira, B. Rodrigues, A. T. Fernandes, W. D. de Oliveira, and U. Bezerra, "Comparison between decision tree and optimal power flow techniques applied to voltage corrective control in electric systems", JETIA, vol. 6, no. 21, pp. 04-12, Feb. 2020.

[1] M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, "Federated learning and differential privacy for medical image analysis," *Scientific Reports*, vol. 12, p. 11953, 2022

[2] N. T. Madathil, S. Patel, and K. Wong, "Revolutionising healthcare data analytics with federated learning: Methods, challenges, and future directions," *Computer Methods and Programs in Biomedicine*, vol. 245, p. 108732, 2025.

[3] S. Pati, S. Chaturvedi, and T. Bhardwaj, "Privacy preservation for federated learning in healthcare: Challenges and opportunities," *Computer Communications*, vol. 223, p. 102594, 2024.

[4] F. Zhang, W. Li, and P. Xu, "Recent methodological advances in federated learning for medical AI applications," *Information Fusion*, vol. 112, p. 102015, 2024.

[5] A. Ali, D. S. Chouhan, and R. Singh, "Health-FedNet: A privacy-preserving federated learning framework for secure medical data sharing," *Heliyon*, vol. 11, no. 3, p. e25538, 2025.

[6] F. Kong, X. Zhou, and J. Zhang, "Federated attention-consistent learning for prostate cancer pathology: Towards robust collaborative intelligence," *Artificial Intelligence in Medicine*, vol. 159, p. 104033, 2024.

[7] A. Ankolekar, Y. Rao, and M. Kim, "Advancing breast, lung, and prostate cancer research with federated learning: A systematic review," *npj Digital Medicine*, vol. 8, no. 1, pp. 415–432, 2025.

[8] Z. L. Teo, J. C. Lim, and M. H. Tan, "Federated machine learning in healthcare: A systematic review of methods and applications," *npj Digital Medicine*, vol. 7, no. 2, p. 88, 2024.

[9] H. Guan and M. Liu, "Federated learning for medical image analysis: A survey," *arXiv preprint* arXiv:2306.05980, 2023.

[10] W. Huang, M. Ye, Z. Shi, G. Wan, H. Li, B. Du, and Q. Yang, "Federated learning for generalization, robustness, and fairness: A survey and benchmark," *arXiv preprint* arXiv:2311.06750, 2023.

[11] E. Darzi, N. M. Sijtsema, and P. M. Van Ooijen, "Fed-SAFE: Securing federated learning in healthcare against adversarial attacks," *arXiv preprint* arXiv:2310.08681, 2023.

[12] M. Fang, X. Cao, J. Jia, and N. Z. Gong, "Local model poisoning attacks to Byzantine-robust federated learning," *arXiv preprint* arXiv:1911.11815, 2019.

[13] L. Lyu, H. Yu, and Q. Yang, "Threats to federated learning: A survey," *arXiv preprint* arXiv:2003.02133, 2020.

[14] V. Tolpegin, S. Truex, M. E. Gursoy, and L. Liu, "Data poisoning attacks against federated learning systems," in *Computer Security – ESORICS 2020*, pp. 480–500, 2020.

[15] A. Sharma, W. Chen, J. Zhao, Q. Qiu, S. Bagchi, and S. Chaterji, "FLAIR: Defense against model poisoning attack in federated learning," *Proc. ACM Asia Conf. Computer and Communications Security (AsiaCCS '23)*, pp. 553–566, New York, NY, USA, 2023.

[16] K. Guo, T. Chen, S. Ren, and N. Li, "Federated learning empowered real-time medical data processing method for smart healthcare," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 21, no. 4, pp. 869–879, 2024.

[17] A. A. Chowdhury, S. M. H. Mahmud, M. P. Uddin, S. Kadry, J. Y. Kim, et al., "Nuclei segmentation and classification from histopathology images using federated learning for end-edge platform," *PLoS ONE*, vol. 20, no. 7, p. e0322749, 2025.

[18] D. Komura, M. Ochi, and S. Ishikawa, "Machine learning methods for histopathological image analysis: Updates in 2024," *Comput. Struct. Biotechnol. J.*, vol. 27, pp. 383–400, Dec. 2024, doi: 10.1016/j.csbj.2024.12.033.

[19] N. Koutsoubis, Y. Yilmaz, and G. Rasool, "Privacy preserving federated learning in medical imaging with uncertainty estimation," *arXiv preprint* arXiv:2406.12815, 2024.

[20] R. Eden, et al., "A scoping review of governance of federated learning in healthcare," *PLoS ONE*, vol. 20, no. 7, p. e0322749, 2025.

[21] S. Lanka and V. Ramesh, "IoT security enhancements in smart healthcare using federated learning," in *Proc. IEEE Conf. Augmented Intelligence and Sustainable Systems (ICAISS)*, 2025.

[22] S. Haggenmüller, L. A. Schoenpflug, et al., "Federated learning for decentralized artificial intelligence in pathology," *JAMA Dermatol.*, vol. 160, no. 2, p. e24154, 2024.

[23] L. A. Schoenpflug, S. Haggenmüller, and W. Roth, "A case study on federated learning in computational pathology," *Comput. Struct. Biotechnol. J.*, vol. 23, pp. 1603–1619, 2025.

[24] J. I. Pisula, et al., "Explainable federated deep learning for carcinoma prediction," *npj Precis. Oncol.*, vol. 9, no. 2, p. 997, 2025.

[25] K. Yang, "Secure and private federated learning," *arXiv preprint* arXiv:2505.17226, 2024.

[26] S. Toor, "Federated learning and custom aggregation schemes," Towards Data Science, Oct. 2025. [Online].

[27] J. Wen, Y. Zhang, Q. Liu, and L. Wang, "A highly generalized federated learning algorithm for brain tumour segmentation," *Sci. Rep.*, vol. 15, p. 5297, 2025.