Journal of Advances in Developmental Research (IJAIDR)

E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

FedX-PneuNet: A Privacy-Preserving and Clinically Explainable Federated Multi-Task Learning Framework for Pneumonia Detection and Localization in Chest X-Ray Images

Onkar Vijayrao Chandure[1], Dr. Dinesh N. Chaudhari[2]

[1]Research Scholar, Department of Computer Science & Engineering, Jawaharlal Darda Institute of Engineering & Technology, Yavatmal, Maharashtra, India affiliated with Sant Gadge Baba Amravati University, Amravati, Maharashtra, India
onkar.86it@gmail.com

[2]Professor, Department of Computer Science & Engineering, Jawaharlal Darda Institute of Engineering & Technology, Yavatmal, Maharashtra, India affiliated with Sant Gadge Baba Amravati University, Amravati, Maharashtra, India
dnchaudhari2007@rediffmail.com

Abstract:
Pneumonia remains one of the leading causes of illness and death worldwide, and catching it early makes a real difference to how well patients recover. Chest X-rays are the go-to imaging tool because they are cheap and widely available, but reading them accurately still takes a trained radiologist and takes time. Deep learning can speed this up, yet most existing models need patient data pulled into one central location to train, which raises real privacy and security concerns. On top of that, these models rarely explain how they arrive at a decision, leaving doctors with little basis to trust them.

This paper proposes FedX-PneuNet, a framework built to detect pneumonia from chest X-rays without compromising patient privacy or behaving as a black box. It is trained on the RSNA Pneumonia Detection Challenge dataset and handles two tasks at once: classifying whether pneumonia is present and locating the affected region of the lung. Training takes place through Federated Learning using the FedAvg approach, so raw images never leave the hospital where they were captured. Differential Privacy and Secure Aggregation add a further layer of protection to whatever is shared during training, and Grad-CAM together with LIME let the model's predictions actually be inspected rather than taken on faith. The framework is evaluated using accuracy, precision, recall, F1-score, AUC, and IoU, benchmarked against the centralized and federated baselines discussed in detail in Section 2.

Keywords: Pneumonia Detection, Chest X-ray Image, Federated Learning, Differential Privacy, Secure Aggregation, Explainable Artificial Intelligence (XAI), Grad-CAM, Deep Learning, Medical Imaging

1. Introduction

Pneumonia is a lung infection in which the air sacs fill with fluid or pus, causing cough, fever, chest pain, and difficulty breathing. It can affect anyone, but the risk is highest for children, the elderly, and people with weakened immunity. The World Health Organization still counts pneumonia among the leading infectious causes of death in children worldwide [35], which is part of why catching it early and correctly matters so much — a missed or delayed diagnosis has a direct, measurable cost to patient outcomes.

Chest X-ray remains the default imaging choice for pneumonia because it is inexpensive, fast, and available even in modestly equipped clinics; on the image, pneumonia typically shows up as cloudy or opaque patches in the lung fields [35]. Reading that image well, however, is harder than it sounds. Machine calibration, how the patient is positioned, and exposure settings all change how the image looks, and pneumonia's visual pattern can be mistaken for tuberculosis, COVID-19, or bronchitis. Diagnosis therefore still leans heavily on the reading radiologist's experience, and even two experienced radiologists do not always agree on the same film. In many rural or under-resourced hospitals, a trained radiologist may not be available at all, which delays diagnosis further.

There is also a second, less technical barrier: trust. Clinicians have good reason to be cautious about plugging an automated tool into their diagnostic workflow if they cannot ask it why it reached a conclusion, and hospital administrators face their own separate hurdles around patient-data governance long before any model is deployed. Put together, a pneumonia-detection system meant for real clinical use has to satisfy three different audiences at once: it must be accurate enough to be worth using, private enough to clear legal and ethical review across institutions, and transparent enough that a radiologist is willing to act on what it says.

1.1 Deep Learning for Pneumonia Detection: Progress and Limitations

Artificial intelligence, particularly deep learning, has been explored extensively as a way to support this diagnostic gap. Convolutional Neural Networks (CNNs) have shown that they can learn diagnostic features directly from chest X-ray images without manual feature engineering. Kermany et al. demonstrated one of the earlier large-scale applications of this idea, using deep learning to classify pediatric chest X-rays for pneumonia [36]. Later works such as ChestX-ray8 and CheXNet extended this direction, showing that deep networks trained on large-scale radiograph datasets could detect multiple thoracic diseases, including pneumonia, with accuracy comparable to practicing radiologists [37, 38].

Despite this progress, nearly all of these models share a common limitation: they are trained using centralized data. This means chest X-ray images from different hospitals or clinics must first be collected and pooled into a single location before training can begin. In a healthcare setting, this requirement creates a serious obstacle. Patient images are sensitive data, and hospitals are bound by privacy regulations and ethical review that restrict how such data can be shared or transferred outside their own systems. Consequently, most publicly available pneumonia detection models are trained on a narrow set of hospitals or a single public dataset, which limits how well they generalize to imaging equipment, patient populations, or disease presentations they were not trained on.

1.2 Federated Learning as a Privacy-Aware Alternative

Federated Learning has emerged as a practical response to this data-centralization problem. Rather than moving patient data to a central server, Federated Learning moves the model to the data. Each participating hospital trains a local copy of the model using its own chest X-ray images, and only the resulting model updates — not the raw images — are sent to a central server for aggregation [1]. McMahan et al. formalized this idea through the Federated Averaging (FedAvg) algorithm, which enabled communication-efficient training across decentralized clients [39]. Sheller et al. later confirmed that this approach is practical in real clinical settings, showing that multiple medical institutions could jointly train a model without ever exchanging patient data [40].

For pneumonia detection specifically, this collaborative structure is valuable because different hospitals typically see different patient demographics, imaging hardware, and disease severity distributions. A model that learns from several institutions at once, without seeing any institution's raw data, tends to generalize better than one trained on a single, narrow dataset. Several recent studies have applied Federated Learning to pediatric pneumonia classification and broader chest X-ray-based lung disease diagnosis, reinforcing that this direction is both feasible and increasingly active [9, 10].

1.3 Remaining Gaps: Privacy Leakage and the Black-Box Problem

Federated Learning reduces, but does not eliminate, privacy risk. Model updates exchanged during training can still leak information about the underlying data — prior work has shown that gradients or parameter updates can, in some cases, be used to infer or partially reconstruct details about the training samples they came from. To close this gap, two techniques are commonly layered on top of Federated Learning. Differential Privacy introduces controlled statistical noise into model updates, making it difficult to trace any single patient's contribution to the trained model [41]. Secure Aggregation, on the other hand, ensures that the central server only ever sees a combined update from all participating clients, never an individual hospital's update in isolation [42]. Together, these mechanisms address the residual privacy concerns that Federated Learning alone does not fully resolve.

A second, equally important gap concerns trust. Deep learning models, federated or otherwise, are frequently criticized as "black boxes" — they output a prediction without indicating which parts of the image influenced that decision. In a clinical setting, this is not a minor inconvenience; a radiologist or physician is unlikely to act on a prediction they cannot verify. Explainable AI (XAI) techniques attempt to address this. Grad-CAM produces a heatmap over the input image, highlighting the regions that most influenced the model's output [43], while LIME approximates the model locally to identify which image regions supported a given prediction [44]. When applied to chest X-rays, both techniques allow a clinician to visually confirm whether the model is attending to clinically relevant lung regions rather than incidental artifacts in the image.

1.4 Scope and Organization of This Paper

Federated Learning, Differential Privacy, Secure Aggregation, and Explainable AI have each received considerable individual attention in medical imaging research, but surprisingly few studies look at how they work together — specifically for detecting and localizing pneumonia in chest X-rays. This review

brings together the current state of research across all four areas, compares the methods and datasets each one uses, points out where the open problems still sit — particularly the joint use of privacy protection and explainability inside one deployable system — and lays out where this line of work could reasonably go next.

The key contributions of this paper can be summarized as follows:

1) It brings together, in one comparative framework, four strands of prior work — deep learning for pneumonia detection, Federated Learning, privacy preservation, and Explainable AI — that have mostly been reviewed on their own until now.

2) It points to a specific, so-far-unaddressed gap: no existing pneumonia-focused study combines Federated Learning, formal privacy guarantees (Differential Privacy and Secure Aggregation), explainability (Grad-CAM and LIME), and lesion localization inside one deployable framework.

3) It proposes FedX-PneuNet, a multi-task architecture and training pipeline built to handle classification and bounding-box localization together on the RSNA Pneumonia Detection Challenge dataset, trained under a federated, privacy-preserving, and explainable regime.

4) It lays out a concrete evaluation protocol covering classification (Accuracy, Precision, Recall, F1, AUC), localization (IoU), and a qualitative assessment of explainability — giving future federated pneumonia-detection systems a benchmark to be measured against.

The remainder of this paper is organized as follows. Section 2 reviews the literature across deep learning, Federated Learning, privacy-preserving techniques, and Explainable AI for pneumonia detection, and identifies the resulting research gap. Section 3 details this research gap and the motivation for the proposed work. Section 4 presents the proposed FedX-PneuNet framework, including its architecture, dataset, workflow, and anticipated implementation challenges. Section 5 concludes the paper with directions for future work.

2. Literature Review

Chest X-rays are one of the fastest and most affordable ways to detect pneumonia, but reading them accurately still depends on experienced radiologists, who are not always available, particularly in rural areas. Deep learning can support this task, though most existing approaches require pooling patient data centrally, which raises privacy concerns. Federated Learning addresses this by letting hospitals train a shared model without exchanging raw data, while techniques such as Differential Privacy and Secure Aggregation protect the model updates themselves, and explainability methods such as Grad-CAM and LIME help clinicians understand and trust the resulting predictions. Tables 1 to 3 summarize the major contributions, methods, and limitations of the studies most relevant to this research.

2.1 Deep Learning and Federated Learning for Pneumonia Detection

Early work in this space relied entirely on centralized training. Kermany et al. [36] showed that a CNN trained on pediatric chest X-rays could reliably tell pneumonia apart from normal cases, and larger benchmark efforts like ChestX-ray8 [37] and CheXNet [38] pushed accuracy close to radiologist level. All three, however, depended on a single pooled dataset — not how hospitals actually share data.

Federated Learning offered a way around this. McMahan et al. [39] introduced the Federated Averaging (FedAvg) algorithm, and Sheller et al. [40] confirmed it works across real medical institutions. Since then, several studies have applied FL directly to pneumonia detection: Liu et al. [47] and Khan & Alam [49] confirmed FL's feasibility for chest X-ray classification using basic setups without privacy or explainability layers, Kareem et al. [50] extended this to distributed detection with the same gap, and Feki et al. [48] applied a similar FL setup to COVID-19 screening from chest X-rays, further evidencing FL's applicability across respiratory-disease imaging more broadly. More recent work has targeted FL's practical weaknesses — Keerthika et al. [19] addressed non-IID data across hospitals, Rana & Marwaha [23] explored unsupervised, autoencoder-based detection, and Pan et al. [10] focused on heterogeneous pediatric hospital data. Sharma et al.'s FedPneu [22] and Mabrouk et al. [20] pushed toward multi-hospital and ensemble-based diagnosis, Vasilevschi et al. [24] proposed a smart-healthcare FL framework, and Trivedi et al. [21] introduced a lightweight FL model for resource-constrained deployment across COVID-19, pneumonia, and TB. A recent survey by Kumari et al. [4] broadens this picture further, reviewing computational intelligence for lung disease diagnosis more generally and confirming that explainability and reliability remain open concerns across the field.

Taken together, this body of work shows a clear progression — from centralized CNNs, to FL setups solving the data-sharing problem, to more recent FL variants handling non-IID data and hospital heterogeneity. What's largely missing is any study combining privacy protection and explainability together.

Table 1: Literature Review on Deep Learning and Federated Learning for Pneumonia Detection

Authors	Year	Method / Technique	Dataset	Major Contribution	Limitation
Kumari et al. [4]	2026	Survey on Computational Intelligence for Lung Disease Diagnosis	Chest CT Scan	Comprehensive review of AI, deep learning, and hybrid computational intelligence techniques for lung disease diagnosis; highlights importance of explainable and reliable AI in medical imaging.	Focuses on CT images and lung cancer rather than chest X-ray pneumonia detection or Federated Learning.
Vasilevschi et al. [24]	2025	Federated Deep Learning	Chest X-ray	Smart healthcare framework for pneumonia detection	Limited privacy mechanisms

Authors	Year	Method / Technique	Dataset	Major Contribution	Limitation
Keerthika et al. [19]	2025	Proximal Guided FL	Chest X-ray	Handles non-IID data efficiently	No explainability analysis
Rana & Marwaha [23]	2025	Unsupervised FL	Chest X-ray	Decentralized pneumonia detection using autoencoders	Limited validation
FedPneu (Sharma et al.) [22]	2024	Cross-silo Federated Learning	Chest X-ray	Multi-hospital collaborative pneumonia detection	No Differential Privacy
Pan et al. [10]	2024	Efficient FL	Pediatric Chest X-ray	Handles heterogeneous hospital data	Explainability not included
Trivedi et al. [21]	2024	Lightweight FL	COVID-19, Pneumonia, TB	Resource-efficient FL model	Limited privacy protection
Mabrouk et al. [20]	2024	Ensemble Federated Learning	Pneumonia Images	Improved collaborative diagnosis	High communication overhead
Kareem et al. [50]	2023	Federated Learning	Chest X-ray	Distributed pneumonia detection	Basic FL only
Khan & Alam [49]	2023	Federated Learning	Chest X-ray	FL for pneumonia classification	No XAI or privacy layer
Liu et al. [47]	2021	Federated Learning	COVID-19 X-ray	Demonstrated FL feasibility in medical imaging	COVID-19 specific
Sheller et al. [40]	2020	Federated Learning	Medical Imaging	Multi-institution collaborative learning	No privacy enhancement
McMahan et al. [39]	2017	FedAvg	General	Introduced Federated Averaging	No security/privacy mechanism
Rajpurkar et al. [38]	2017	CheXNet (DenseNet)	Chest X-ray	High pneumonia detection accuracy	Centralized learning
Wang et al. [37]	2017	ChestX-ray8	ChestX-ray8	Large benchmark dataset	Centralized approach
Kermany et al. [36]	2018	CNN	Pediatric Chest X-ray	Deep learning for pneumonia diagnosis	Centralized dataset

Table 1 traces the evolution from centralized deep learning models to privacy-preserving federated approaches for pneumonia detection from chest X-rays. While recent work has improved collaborative

diagnosis and handled challenges such as non-IID data, most studies still lack strong privacy mechanisms or explainable AI — pointing to the need for a more secure, interpretable federated learning framework.

2.2 Privacy-Preserving Techniques in Federated Learning

FL by itself does not fully protect patient data, since model updates can still leak information about the data used to produce them. Dwork et al. [41] laid the mathematical groundwork for Differential Privacy, and Bonawitz et al. [42] introduced Secure Aggregation, ensuring the central server only ever sees a combined update rather than any single hospital's contribution. Building on these foundations, Ziegler et al. [45] showed Differential Privacy can defend against reconstruction attacks in FL, though at a measurable accuracy cost, and Ahmed et al. [46] applied an adaptive DP approach to chest X-ray classification, though scoped specifically to COVID-19.

A large part of the privacy literature is survey-based rather than implemented. Koutsoubis et al. [26], Pati et al. [25], Zhu et al. [28], and Guan et al. [27] each review privacy threats and defense in medical-imaging FL, and more recent surveys by Ghosh et al. [5] and Karmode & Shahade [7] extend this further — the former reviewing FL privacy and non-IID challenges, the latter reviewing Blockchain-integrated FL as an emerging direction for secure healthcare AI. None of these implement a privacy-preserving FL system specifically for pneumonia detection, leaving an open question of how these techniques hold up once applied to a disease-specific classification and localization task. Complementary privacy architectures have also been explored outside this DP/Secure-Aggregation pairing: Markkandan et al. [11] proposed a privacy-preserving expert system for collaborative diagnosis across institutions, and Bhardwaj and Sumangali [12] combined Federated Learning with blockchain and explainable AI to secure healthcare data end-to-end. Both point toward the same conclusion from a different angle rather than replacing it.

Table 2: Literature Review on Privacy-Preserving Federated Learning

Authors	Year	Privacy Technique	Contribution	Limitation
Ghosh et al. [5]	2026	Systematic Literature Review on Federated Learning	Reviews recent FL techniques, privacy preservation, communication efficiency, and non-IID data challenges in medical imaging; recommends combining FL with stronger privacy and explainability.	Survey paper; no implementation for pneumonia detection.
Karmode & Shahade [7]	2026	Review of Blockchain-integrated Deep Federated Learning	Highlights integration of Blockchain, Federated Learning, Differential Privacy, and Secure Aggregation for secure healthcare AI and discusses future research directions.	General healthcare review; not specific to chest X-ray pneumonia detection.

Authors	Year	Privacy Technique	Contribution	Limitation
Ziegler et al. [45]	2024	Differential Privacy	Defends against reconstruction attacks in FL	Accuracy–privacy trade-off
Ahmed et al. [46]	2024	Adaptive Differential Privacy	Privacy-aware FL for chest X-ray classification	COVID-19 focused
Koutsoubis et al. [26]	2024	Privacy-preserving FL Survey	Medical imaging privacy challenges	Survey only
Pati et al. [25]	2023	Privacy Survey	Reviews FL privacy threats	No implementation
Zhu et al. [28]	2023	Privacy-preserving Medical AI	Reviews DP, encryption, FL	General review
Guan et al. [27]	2023	FL Survey	Comprehensive medical image FL review	Disease independent
Bonawitz et al. [42]	2017	Secure Aggregation	Protects client model updates	Does not include DP
Dwork et al. [41]	2006	Differential Privacy	Mathematical foundation of DP	General framework

Table 2 summarizes privacy-preserving techniques applied to federated learning in medical imaging, including Differential Privacy, Secure Aggregation, and blockchain-integrated approaches. Much of this work is either survey-based or narrowly scoped to a specific application, leaving room for a more comprehensive privacy-preserving federated learning framework built specifically for pneumonia detection.

2.3 Explainable AI for Pneumonia Detection

The third strand of work addresses model transparency. Grad-CAM, from Selvaraju et al. [43], and LIME, from Ribeiro et al. [44], remain the two most widely used explainability techniques, both originating from general computer vision rather than medical imaging. At a broader level, Yeole and Dharmik [2] proposed an explainable Federated Learning framework for privacy-preserving lung-disease detection generally, and Houssein et al. [6] surveyed explainable AI techniques for medical imaging systems at large; both confirm that explainability and federated privacy are increasingly studied together, even where pneumonia is not the specific focus. Their use in pneumonia detection has grown steadily — Shahi & Bagale [33] applied Grad-CAM for weakly supervised localization, Khadidos et al. [29] combined Grad-CAM and LIME for pediatric pneumonia detection, and Colin & Surantha [31] and Mahamud et al. [30] applied interpretable and explainable transfer-learning approaches, though both stayed within a centralized setup. Hendryani et al. [32] combined a hybrid CNN-ViT architecture with Grad-CAM, again centralized. Two

more recent works move closer to the intersection this paper is concerned with. FLPneXAINet by Biswas et al. [8] combines FL, GAN-based augmentation, and Grad-CAM for pneumonia prediction but omits Differential Privacy and Secure Aggregation. FLEM-XAI by Durga et al. [9] similarly combines FL with Grad-CAM and SHAP, but targets multi-disease lung diagnosis rather than pneumonia specifically, and likewise stops short of formal privacy mechanisms.

Table 3: Literature Review on Explainable AI for Pneumonia Detection

Authors	Year	Explainability Method	Contribution	Limitation
FLPneXAINet (Biswas et al.) [8]	2025	Grad-CAM	FL + GAN + XAI for pneumonia prediction	No Differential Privacy or Secure Aggregation
FLEM-XAI (Durga et al.) [9]	2025	Grad-CAM, SHAP	Explainable FL for lung disease diagnosis	Multi-disease focus
Hendryani et al. [32]	2025	Hybrid CNN-ViT + Grad-CAM	Explainable pneumonia detection	Centralized learning
Khadidos et al. [29]	2024	Grad-CAM, LIME	Explainable pediatric pneumonia detection	No Federated Learning
Mahamud et al. [30]	2024	Explainable Transfer Learning	Multi-lung disease diagnosis	Centralized framework
Colin & Surantha [31]	2024	Interpretable Deep Learning	Clinically interpretable pneumonia detection	No privacy protection
Shahi & Bagale [33]	2023	Grad-CAM	Weakly supervised pneumonia localization	No FL integration
Selvaraju et al. [43]	2017	Grad-CAM	CNN visual explanations	General computer vision
Ribeiro et al. [44]	2016	LIME	Local interpretable explanations	General explainability

Table 3 reviews explainability methods — chiefly Grad-CAM, LIME, and SHAP — used in pneumonia and lung-disease detection. These methods have improved clinicians' trust in AI-based diagnosis, but most of the underlying models are centralized or lack robust privacy protection, reinforcing the case for combining explainable AI with a secure federated learning framework.

2.4 Summary of the Gap

Across all three areas, a consistent pattern emerges: studies tend to solve one or two of three problems — accuracy, privacy, or explainability — but rarely all three together. Almost none combine Federated Learning, Differential Privacy or Secure Aggregation, and Grad-CAM or LIME within a single

pneumonia-specific framework that also performs localization alongside classification. This gap motivates the proposed FedX-PneuNet framework.

3. Research Gap and Motivation

3.1 Research Gap

Deep learning has shown strong results for pneumonia detection from chest X-rays, but most existing models rely on centralized training, which means sharing patient data — exactly what Federated Learning (FL) avoids by keeping raw images at each hospital.

The closest existing work, FLPneXAINet, combines FL, GAN-based augmentation, and Explainable AI for pneumonia prediction [8], but only handles classification, not localization, and skips Differential Privacy (DP) and Secure Aggregation (SA). A related study improves FL under heterogeneous hospital data [41] but also leaves out XAI and formal privacy protection. Ziegler et al. showed FL updates are vulnerable to reconstruction attacks and used Rényi Differential Privacy to reduce this risk [45], but their focus is general chest X-ray classification, not pneumonia, and excludes explainability. Other work pairing Grad-CAM and LIME with DenseNet121 and EfficientNet-B0 improved transparency [28, 29], but stayed centralized with no FL or privacy layer.

Most FL-based pneumonia studies also stop at classification. The RSNA Pneumonia Detection Challenge dataset includes bounding-box annotations suited for both classification and localization [34], yet this remains largely unused in FL-based work.

No existing study brings FL, DP, SA, explainability, and localization together — this is the gap the proposed work addresses.

3.2 Motivation

Pneumonia calls for early, accurate diagnosis, and chest X-rays remain the fastest and cheapest way to check for it — but reading them still depends on radiologists who are not always available, especially in under-resourced areas. Deep learning can help, but training a strong model needs data from multiple hospitals, and hospitals cannot freely share raw images because of privacy and legal restrictions. FL solves the sharing problem but not the leakage problem — updates can still expose private information, which is why DP and SA matter alongside it. Doctors also need to know why a model predicted pneumonia before they will trust it, which is where Grad-CAM and LIME come in. What drives this work, then, is the need for a single system that is accurate, private, secure, and explainable at once — one that uses RSNA's bounding-box annotations to handle detection and localization together rather than as two separate problems.

This is not only a technical preference but increasingly a legal one: regulations such as HIPAA in the United States and the GDPR in the European Union restrict cross-border and cross-institution transfer of patient-identifiable imaging data, and non-compliance carries direct legal and financial risk for hospitals. A framework that keeps raw chest X-rays local by design, rather than relying on data-use agreements and

post-hoc anonymization, is therefore easier for hospitals to adopt and easier for institutional review boards to approve.

3.3 Research Problem Selection

Based on this gap, the selected research problem is: design and development of a privacy-preserving Federated Learning framework with Explainable AI for pneumonia detection and localization using chest X-ray images. This combines distributed training to avoid raw data sharing, privacy techniques to protect model updates, and explainability to support clinical trust — with the proposed model, FedX-PneuNet, built as a multi-task model for classification and localization together.

3.4 Problem Statement

Deep learning shows strong results for pneumonia detection, but centralized training raises privacy concerns. FL avoids raw data sharing but can still leak information through updates, and existing privacy-preserving FL work rarely covers pneumonia-specific localization or explainability, while existing explainable pneumonia models mostly stay centralized. A Federated Learning framework is needed that detects and localizes pneumonia while protecting patient data and explaining its predictions through Grad-CAM and LIME.

3.5 Scope Boundaries

To keep the proposed problem well-defined, this work is scoped specifically to pneumonia detection and localization from frontal-view chest X-rays under a cross-silo Federated Learning setting, where a small number of hospital clients each hold a sizeable local dataset and participate reliably in every training round. It does not address cross-device FL (many small, intermittently available clients such as mobile devices), Byzantine or adversarial client behavior, or multi-disease diagnosis beyond pneumonia; these are recognized as related but separate problems that build on, rather than substitute for, the privacy-explainability combination targeted here. Work on lung-disease imaging beyond pneumonia specifically — including COVID-19 CT segmentation [3], multiclass lung-disease transfer learning [13, 18], explainable multi-disease transformer models [15, 16], and dual-path COVID-19 classification with explainable AI [14, 17] — confirms that deep learning and explainability are active research areas across respiratory imaging broadly, even though these studies fall outside this review's chest-X-ray, pneumonia-specific scope.

4. Proposed Framework

Based on the gaps identified in Section 3, a unified framework would need to bring Federated Learning (FL), privacy protection, and Explainable AI (XAI) together in one system, rather than treating them separately as most existing work does. In such a setup, each hospital would train the model on its own chest X-ray images, so patient data never leaves the hospital. Only model updates — not raw images — would be sent to a central server for aggregation. Since data can vary a lot between hospitals, a method like FedProx would help each hospital's local model adapt to its own data while still learning from the group.

4.1 Overview of the Proposed Methodology

The proposed direction starts by reviewing existing work on pneumonia detection, privacy-preserving methods, Federated Learning, and XAI to confirm the gaps discussed earlier. The RSNA Pneumonia Detection Challenge dataset — chosen for its bounding-box annotations — would then be prepared through preprocessing and normalization, supporting both classification and localization. A multi-task model, FedX-PneuNet, would be designed to handle both tasks together.

Training would follow a Federated Learning setup, where hospitals train locally and only share model parameters, aggregated using FedAvg. Differential Privacy and Secure Aggregation would be layered on top to protect both patient data and individual hospital updates. Grad-CAM and LIME would then be used to explain predictions by highlighting the lung regions driving each decision. Finally, the framework would be evaluated against existing pneumonia detection methods using classification, localization, privacy, and explainability metrics. This overall flow is shown in Figure 1.

4.2 Model Architecture

FedX-PneuNet is designed as a shared-backbone, multi-task network rather than two separate models. A convolutional encoder (e.g., a ResNet- or EfficientNet-style backbone, chosen for its favorable accuracy-to-compute ratio on the modest hardware typical of hospital edge deployments) extracts a common feature representation from each chest X-ray. This shared representation then feeds two lightweight task-specific heads: a classification head, which outputs a pneumonia-present/absent probability through a fully connected layer and sigmoid activation, and a localization head, which regresses bounding-box coordinates for the affected lung region in the manner of a single-stage detector. Sharing the backbone keeps the per-client model compact enough for repeated federated transmission while still letting the two heads specialize; the classification and localization losses are combined during training as a joint multi-task objective. This design choice also keeps Grad-CAM and LIME straightforward to apply, since both explainability methods operate on the shared encoder's feature maps rather than requiring separate explanations for two unrelated models.

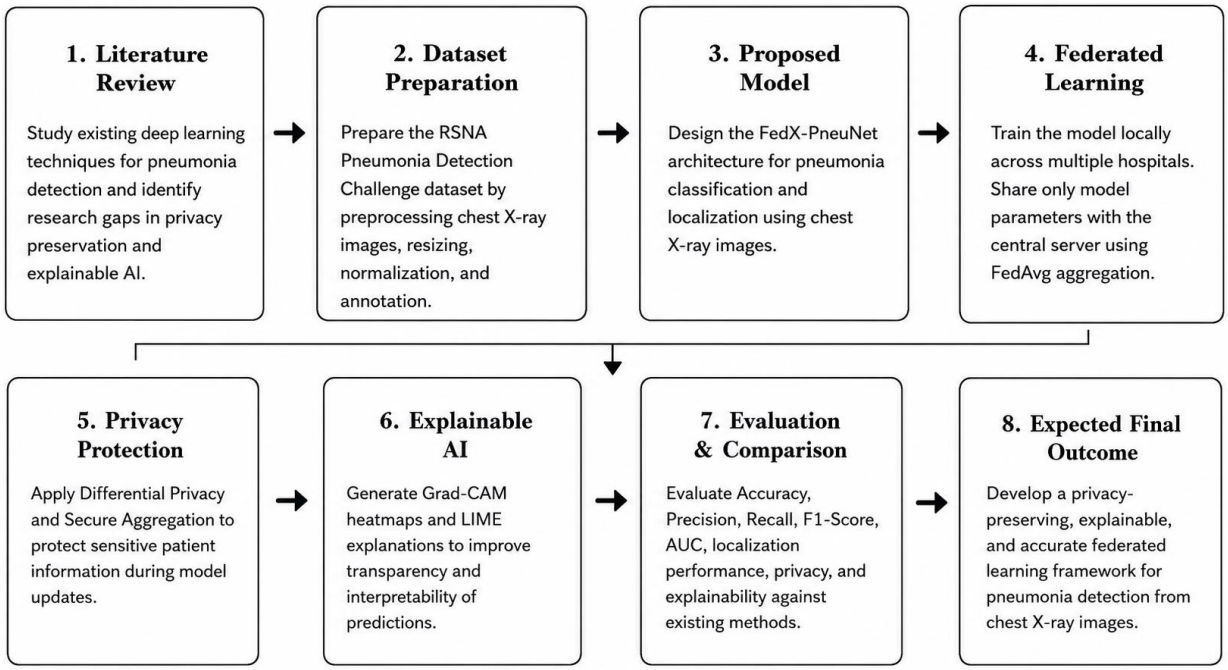
4.3 Dataset and Preprocessing

FedX-PneuNet is developed around the RSNA Pneumonia Detection Challenge dataset [34], a subset of 30,000 frontal-view chest radiographs drawn from the NIH CXR8 collection. Of these, 26,684 images carry stage-2 training annotations, roughly half showing pneumonia-consistent opacities and half split between normal and other-abnormality cases; bounding boxes for the positive class were produced by board-certified radiologists, which is what makes the dataset usable for localization and not only classification. To simulate a realistic federated setting, this pool is partitioned across a set of virtual hospital clients using a non-IID split — varying each client's positive/negative ratio and image count — since hospitals in practice see different patient mixes rather than an identical, evenly shuffled sample. Preprocessing standardizes each DICOM image to a fixed resolution, applies intensity normalization, and converts RSNA's bounding-box CSV annotations into per-image target masks usable by the localization head; class imbalance within each client is addressed through weighted sampling during local training rather than global resampling, since global resampling would require access to the full, centralized label distribution.

4.4 Summary of the Proposed Workflow

The proposed methodology follows eight steps, summarized in Figure 1: **(1)** review existing literature to confirm research gaps, **(2)** prepare the RSNA dataset through preprocessing and annotation, **(3)** design the FedX-PneuNet multi-task model, **(4)** implement Federated Learning across hospitals using FedAvg, **(5)** apply Differential Privacy and Secure Aggregation, **(6)** integrate Grad-CAM and LIME for explainability, **(7)** evaluate and compare against existing methods, and **(8)** deliver a privacy-preserving, explainable, and accurate framework for pneumonia detection and localization.

Figure 1: Proposed Research Methodology Workflow for FedX-PneuNet, from Literature Review through the Expected Final Outcome



4.5 Anticipated Implementation Challenges

Bringing FL, DP, Secure Aggregation, and XAI together in one pipeline introduces practical difficulties beyond what any single component faces in isolation. First, non-IID data across hospitals — differing patient demographics, imaging hardware, and pneumonia prevalence — can slow FedAvg convergence or bias the global model toward better-represented clients; this motivates the FedProx-style proximal term noted in Section 4 rather than plain FedAvg alone. Second, Differential Privacy's noise injection and Secure Aggregation's masking both tend to reduce raw predictive accuracy relative to a centralized, unprotected baseline, so the evaluation protocol proposed for this framework is intended to report this privacy-utility trade-off across multiple privacy budgets instead of hiding it behind a single headline accuracy number. Third, computing Grad-CAM and LIME explanations under a federated, multi-task architecture is heavier than in a standard centralized classifier, since both localization and classification heads must be interpreted jointly, and this adds inference-time overhead that a hospital-side deployment needs to budget for. Finally, communication cost across FL rounds grows with model size and client count, which is a practical constraint on how large the FedX-PneuNet backbone can be if hospitals are expected

to participate over standard clinical network connections rather than high-bandwidth research infrastructure.

5. Conclusion

Pneumonia is still a serious health problem. Chest X-rays remain the fastest and cheapest way to check for it, but reading them well depends on radiologists who are not always available, especially in under-resourced areas. Deep learning can help here, but most existing models are trained on data pooled into one place, which does not match how patient data is actually allowed to move between hospitals. This paper looked across the literature on deep learning, Federated Learning, privacy techniques, and explainable AI for pneumonia detection, and a clear pattern emerged: most studies fix one piece of the problem — accuracy, privacy, or explainability — but rarely all three together, and almost none also locate the affected lung region instead of only classifying the image.

That gap is what this paper builds on. A framework that trains locally at each hospital, adds Differential Privacy and Secure Aggregation to stop model updates from leaking patient data, and uses Grad-CAM and LIME so doctors can see why a prediction was made, treats accuracy, privacy, and trust as one connected problem instead of three separate fixes. Built around the RSNA dataset, the proposed FedX-PneuNet also goes further by identifying where the pneumonia appears, not just whether it's present. Developed as outlined in Section 4, this direction points toward pneumonia-detection tools that hospitals can trust and adopt without relinquishing control of their patients' data.

Future work includes implementing and empirically validating FedX-PneuNet on the RSNA dataset under realistic non-IID hospital partitions, quantifying the accuracy/privacy trade-off across a range of differential-privacy budgets, and conducting a structured clinician evaluation of the Grad-CAM and LIME explanations to assess whether they meaningfully support diagnostic trust in practice rather than only appearing plausible in isolated examples.

REFERENCES:

1. K.S. Archana, V.H. Shah, K. Joseph, “Chest X-Ray Pneumonia Detection Using Federated Learning: Privacy-Preserving Deep Learning for Medical Image Classification”, 2025 5th International Conference on Intelligent Technologies (CONIT), 2025.
2. A. Yeole, R. Dharmik, “Explainable Federated Learning Framework for Privacy-Preserving Lung Disease Detection Using Chest X-Ray Images”, ICAIQSA, 2025.
3. U.L. Marksia, C.Y. Rubavathi, “Accurate Segmentation of COVID-19 Infected Regions in Lung CT Scans with Deep Learning”, Neural Computing and Applications, 2024.
4. P. Kumari, L. Goel, “Emerging Computational Intelligence Based Techniques for Lung Cancer Diagnosis and Classification on Chest CT Scan Images: A Comprehensive Survey”, Artificial Intelligence Review, 2026.
5. D. Ghosh et al., “Advancements and Challenges of Federated Learning in Medical Imaging: A Systematic Literature Review”, Artificial Intelligence Review, 2026.
6. E.H. Houssein et al., “Explainable Artificial Intelligence for Medical Imaging Systems Using Deep Learning: A Comprehensive Review”, Cluster Computing, 2025.

7. S. Karmode, A.K. Shahade, “A Comprehensive Review and Future Directions on Blockchain-Integrated Deep Federated Learning for Privacy-Preserving Healthcare Applications”, International Journal of Data Science and Analytics, 2026.
8. S. Biswas, R. Mostafiz, M.S. Uddin, M.S. Uddin, “FLPneXAINet: Federated deep learning and explainable AI for improved pneumonia prediction utilizing GAN-augmented chest X-ray data”, PLOS ONE, vol. 20, no. 7, Art. no. e0324957, 2025. https://doi.org/10.1371/journal.pone.0324957
9. S.D.S. Durga, E.D. Esther Daniel, S.S. Seetha, V.K. Reshma, V. Sachnev, “FLEM-XAI: Federated learning based real time ensemble model with explainable AI framework for an efficient diagnosis of lung diseases”, Frontiers in Computer Science, vol. 7, Art. no. 1633916, 2025. https://doi.org/10.3389/fcomp.2025.1633916
10. Z. Pan, H. Wang, J. Wan, L. Zhang, J. Huang, Y. Shen, “Efficient federated learning for pediatric pneumonia on chest X-ray classification”, Scientific Reports, vol. 14, Art. no. 23272, 2024. https://doi.org/10.1038/s41598-024-74491-5
11. S. Markkandan, N.P.G. Bhavani, S.S. Nath, “A Privacy-Preserving Expert System for Collaborative Medical Diagnosis Across Multiple Institutions Using Federated Learning”, Scientific Reports, 2024.
12. T. Bhardwaj, K. Sumangali, “An Explainable Federated Blockchain Framework with Privacy-Preserving AI Optimization for Securing Healthcare Data”, Scientific Reports, 2025.
13. N.L. Yadav et al., “Transfer Learning with Fuzzy Decision Support for Multiclass Lung Disease Classification: Performance Analysis of Pre-trained CNN Models”, Scientific Reports, 2025.
14. M.M.A. Rahat et al., “An Efficient Dual Path Deep Learning Framework for COVID-19 Classification Using Lung CT Scans with Explainable AI”, Scientific Reports, 2026.
15. X. Fu et al., “Explainable Hybrid Transformer for Multi-Classification of Lung Disease Using Chest X-rays”, Scientific Reports, 2025.
16. V. Lokunde et al., “Class-Attention Pooling and Token Sparsity Based Vision Transformers for Chest X-ray Interpretation”, Scientific Reports, 2026.
17. R. Annan et al., “Privacy-Preserving Federated Learning with Optimized Ensemble Weighting and Knowledge Distillation for COVID-19 Detection from Non-IID Medical Imaging Data”, Scientific Reports, 2025.
18. M. Beula Kutti, T. Karthick, “An Automated Detection and Classification of Lung Diseases Using a Hybrid Deep Learning Technique on Chest X-ray Images”, International Journal of Computational Intelligence Systems, 2026.
19. K.P., S.P., N.K. AR, “Proximal Guided Hybrid Federated Learning Approach with Parameter Efficient Adaptive Intelligence for Pneumonia Diagnosis”, Scientific Reports, 2026.
20. A. Mabrouk, R.P. Díaz Redondo, M. Abd Elaziz, M. Kayed, “Ensemble Federated Learning: An approach for collaborative pneumonia diagnosis”, Applied Soft Computing, vol. 144, Art. no. 110500, 2023. https://doi.org/10.1016/j.asoc.2023.110500
21. N.K. Trivedi et al., “Lightweight Federated Learning for COVID-19, pneumonia, and TB from chest X-ray images”, in Proceedings of the International Conference on Artificial Intelligence for Innovations in Healthcare Industries, 2023.

22. S. Sharma, K. Guleria, A. Dogra, "FedPneu: Federated Learning for pneumonia detection across multiclient cross-silo healthcare datasets", Current Medical Imaging, vol. 21, 2025. https://doi.org/10.2174/0115734056333970241212132150
23. N. Rana, H. Marwaha, "Pneumonia detection from X-ray images using Federated Learning: An unsupervised learning approach", Measurement: Sensors, vol. 37, Art. no. 101410, 2025. https://doi.org/10.1016/j.measen.2024.101410
24. A.-M. Vasilevschi, C.-A. Coman, M. Ianculescu, O.A. Coman, "Artificial Intelligence for pneumonia detection: A Federated Deep Learning approach in smart healthcare", Future Internet, vol. 17, no. 12, Art. no. 562, 2025. https://doi.org/10.3390/fi17120562
25. S. Pati et al., "Privacy preservation for Federated Learning in health care", Patterns, vol. 5, no. 7, Art. no. 100974, 2024. https://doi.org/10.1016/j.patter.2024.100974
26. N. Koutsoubis, Y. Yilmaz, R.P. Ramachandran, M. Schabath, G. Rasool, "Privacy preserving Federated Learning in medical imaging with uncertainty estimation", arXiv preprint arXiv:2406.12815, 2024.
27. H. Guan, P.-T. Yap, A. Bozoki, M. Liu, "Federated Learning for medical image analysis: A survey", Pattern Recognition, vol. 151, Art. no. 110424, 2024. https://doi.org/10.1016/j.patcog.2024.110424
28. Y. Zhu, X. Yin, A.W.-C. Liew, H. Tian, "Privacy-preserving in medical image analysis: A review of methods and applications", arXiv preprint arXiv:2412.03924, 2024.
29. A.O. Khadidos, A. Nanyonga, A.O. Khadidos, O.M. Mirza, M.T. Yilmaz, "Explainable deep learning for pediatric pneumonia detection in chest X-ray images", arXiv preprint arXiv:2601.09814, 2026.
30. E. Mahamud et al., "An explainable artificial intelligence model for multiple lung diseases classification from chest X-ray images using fine-tuned transfer learning", Decision Analytics Journal, vol. 12, Art. no. 100499, 2024. https://doi.org/10.1016/j.dajour.2024.100499
31. J. Colin, N. Surantha, "Interpretable deep learning for pneumonia detection using chest X-ray images", Information, vol. 16, no. 1, Art. no. 53, 2025. https://doi.org/10.3390/info16010053
32. A. Hendryani, V. Nurdinawati, A. Komarudin, "Explainable Artificial Intelligence model for pneumonia detection: A hybrid CNN-ViT and Grad-CAM", Jurnal Nasional Teknik Elektro dan Technology Informasi, vol. 14, no. 4, 2025. https://doi.org/10.22146/jnteti.v14i4.19822
33. K. Shahi, A. Bagale, "Weakly supervised pneumonia localization from chest X-rays using deep neural network and Grad-CAM explanations", arXiv preprint arXiv:2511.00456, 2025.
34. Radiological Society of North America, "RSNA Pneumonia Detection Challenge 2018", RSNA, 2018.
35. World Health Organization, "Pneumonia in children", WHO, Nov. 2022.
36. D.S. Kermany et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning", Cell, vol. 172, no. 5, pp. 1122–1131.e9, 2018. https://doi.org/10.1016/j.cell.2018.02.010
37. X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, "ChestX-Ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3462–3471. https://doi.org/10.1109/CVPR.2017.369

38. P. Rajpurkar et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning", arXiv preprint arXiv:1711.05225, 2017.
39. H.B. McMahan, E. Moore, D. Ramage, S. Hampson, B.A. y Arcas, "Communication-efficient learning of deep networks from decentralized data", in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, PMLR, vol. 54, pp. 1273–1282, 2017.
40. M.J. Sheller et al., "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data", Scientific Reports, vol. 10, Art. no. 12598, 2020. https://doi.org/10.1038/s41598-020-69250-1
41. C. Dwork, F. McSherry, K. Nissim, A. Smith, "Calibrating noise to sensitivity in private data analysis", in Theory of Cryptography Conference, Lecture Notes in Computer Science, vol. 3876, pp. 265–284, Springer, 2006. https://doi.org/10.1007/11681878_14
42. K. Bonawitz et al., "Practical secure aggregation for privacy-preserving machine learning", in Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 1175–1191. https://doi.org/10.1145/3133956.3133982
43. R.R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization", in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 618–626. https://doi.org/10.1109/ICCV.2017.74
44. M.T. Ribeiro, S. Singh, C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier", in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144. https://doi.org/10.1145/2939672.2939778
45. J. Ziegler, B. Pfitzner, H. Schulz, A. Saalbach, B. Arnrich, "Defending against reconstruction attacks through differentially private federated learning for classification of heterogeneous chest X-ray data", Sensors, vol. 22, no. 14, Art. no. 5195, 2022. https://doi.org/10.3390/s22145195
46. R. Ahmed, P.K.R. Maddikunta, T.R. Gadekallu, N.K. Alshammari, F.A. Hendaoui, "Efficient differential privacy enabled federated learning model for detecting COVID-19 disease using chest X-ray images", Frontiers in Medicine, vol. 11, Art. no. 1409314, 2024. https://doi.org/10.3389/fmed.2024.1409314
47. B. Liu, B. Yan, Y. Zhou, Y. Yang, Y. Zhang, "Experiments of Federated Learning for COVID-19 chest X-ray images", arXiv preprint arXiv:2007.05592, 2020.
48. I. Feki, S. Ammar, Y. Kessentini, K. Muhammad, "Federated learning for COVID-19 screening from chest X-ray images", Applied Soft Computing, vol. 106, Art. no. 107330, 2021. https://doi.org/10.1016/j.asoc.2021.107330
49. S.H. Khan, M.G.R. Alam, "A Federated Learning approach to pneumonia detection", in Proceedings of the 7th International Conference on Engineering and Emerging Technologies, 2021. https://doi.org/10.1109/ICEET53442.2021.9659591
50. A. Kareem, H. Liu, V. Velisavljevic, "A Federated Learning framework for pneumonia image detection using distributed data", Healthcare Analytics, vol. 4, Art. no. 100204, 2023. https://doi.org/10.1016/j.health.2023.100204