Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FEDERATED LEARNING IN LUNG CANCER RADIOMICS AND MEDICAL IMAGING: A META-SURVEY OF COLLABORATIVE MODELS

SRIVIDYA.CH[1], DR. RAMA SUBRAMANIAN K[2], MADHUBALA.M[3]

[1]Research Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad-500075, Telangana, India.| [1]Assistant Professor, Department of Computer Science and Engineering(AI&ML), Malla Reddy University, Hyderabad-500100, Telangana, India.

[2]Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad-500075, Telangana, India.

[3]VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India.

Email: [1]chsrividya12@gmail.com, [2]ramasubramaniankrish@klh.edu.in,[3]baladandamudi@gmail.com

ABSTRACT

Lung cancer is a major cause of cancer morbidity and death worldwide; new methods of testing are urgently needed but are not always available and accurate. Federated learning (FL) is an emerging paradigm in the machine learning field that shows significant potential in the context of medical data analysis -- namely, it allows for model training over a large number of distributed data sources where patient data can remain private. We note that this overview is focused on more recent works on federated learning methods for lung cancer-related imaging and screening tasks, with references chosen primarily based on a directly comparable methodology. Specifically, we review the research works with respect to different FL architectures, the most commonly used data preprocessing methods as well as the metrics for FL performance assessment in healthcare applications. Furthermore, we analyze the advantages of federated learning such as improved model generalization on heterogeneous data, and enhanced privacy protection. An overview of the FL strategies in terms of their centralized, decentralized, and hybrid architectures is discussed, along with the advantages and disadvantages for LC screening. It further explores a broader range of challenges such as non-IID (non-Independent and Identically Distributed) and IID data distributions, commutation costs, and stability of training, etc. We provide an extensive evaluation of the performance of various FL methods on random field studies compared with other experimental results reported in literature to benchmark the current state-of-the-art methods against each other. It also mentions the integration of federated learning with complementary techniques like (deep learning), (blockchain) to facilitate big scale cooperation learning in reducing overfitting and simultaneously mitigate the security problems due to deep learning, block chain respectively. Conclusion The findings indicate that federated learning could simultaneously improve performance and adapt populations for lung cancer screening under the constraints of stringent data privacy. Many critical issues remain, including those for non-IID data handling, communication efficiency, and scalability performance. In this paper, we clarify future directions (such as secure aggregation mechanisms, personalized federated learning models, and secure multiparty computation) to mitigate these problems. Thus, this work should provide an excellent reference for researchers and practitioners interested in applying federated learning to privacy-friendly lung cancer screening and diagnostic support systems. It covers the systematic review on 37 peer-reviewed studies, (2013–2025) and controlled benchmarking for 5 FL algorithms that is FedAvg(Federated Averaging), FedSGD(Federated Stochastic Gradient Descent), FedProx(Federated Proximal), FedAtt(Federated Attention) and FedEnsemble with respect to the Chest CT-Scan dataset (publicly available, 9,500 images, Kaggle). FedEnsemble achieved the highest optimally weighted score × class performance across multiple performance measurements; 92.0% accuracy, 91.9% F1-score, 91.5% precision and 92.2% recall suggesting it produced the most ideal overall performance, whereas,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FedSGD had the lowest communication bandwidth requirement (i.e., 45 round and 180 MB of data), with an optimally weighted score × class of 80.0% accuracy.

Keywords: *Federated Learning, Lung Cancer Detection, Machine Learning, Privacy-Preserving, Decentralized Data, Deep Learning, Blockchain, Data Heterogeneity*

1. INTRODUCTION

Lung cancer is one of the most significant public health issues worldwide, and one of the leading causes of cancer death globally. Prompt and precise diagnosis is important for having a better prognosis and survival, and a higher quality of life. Common diagnostic methods, including medical imaging and biopsy, are invasive, time-consuming, and rely heavily on expert interpretation. Within the last decade, recent advancements in artificial intelligence (AI) and machine learning (ML) have made it possible to use medical diagnostic decision-support systems in a non-invasive, efficient, and highly accurate manner.

Machine learning models are widely used in several medical imaging modalities for lung cancer, including computed tomography (CT) scan and chest X-ray, to assist with lung nodule detection, malignancy classification, and outcome prediction. These models help radiologists in various ways, including second opinions, lowering the risk of diagnostic errors, and making clinical workflow more efficient. Yet, the deployment of high-performance ML models generally necessitates large amounts of annotated data, posing severe data collection, integration and privacy protection hurdles.

Human medical data are characterized as very sensitive and subject to the applicable regulatory regime, the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union. Collecting, aggregating and hosting all the data in one location for model training also increases risk for data breaches, which will ultimately cause loss of patient trust and healthcare institutions being sued as well as facing reputational damage. So, there is a serious demand for paradigms that allow collaborative model development without the need to share the actual data with each other.

Federated learning (FL) tackles these issues by allowing decentralized training of models across a large number of institutions without needing to share raw data. Federated learning, first proposed by Google in 2017, enables multiple participants to independently train models on their own private datasets, and to send only encrypted model updates to a central server for aggregation. FL achieves significant privacy enhancements as sensitive data remains at the source and enables institutions to leverage diverse and large-scale datasets, thereby improving the robustness and generalization of the model.

We can broadly classify federated learning architectures into three categories centralised (or coordinated) federated learning architectures, decentralised (or uncoordinated) federated learning architectures, and hybrid federated learning architectures. Centralized federated learning contains a central server which coordinates the federated training by collecting model updates from distributed clients, but such architecture forms a single point of failure. On the other hand, the model parameters can be exchanged between participating nodes in a peer-to-peer manner in decentralized federated learning, removing the central coordinator and enhancing the system resilience, but incurring the additional complication of synchronization [3]. Hybrid strategies intend to combine qualities of both architectures to balance efficiency and robustness.

Federated learning has demonstrated considerable potential for multi-institutional imaging and screening tasks for lung cancer by enabling the federated training of a model under partial data coverage, preventing overfitting on heterogeneous data. Related work: Existing literature has recognized various FL paradigms to detect lung nodules, predict malignancy, and predict survival while few of them to introduce advanced deep learning approaches, optimization mechanism, and privacy-preserving methods for comparison. Different types of preprocessing image such as normalization, denoising, contrast enhancement and data augmentation like rotation and scale [1] could be used to improve the performance and robustness of the model.

Even though there are advancements, federated learning for healthcare remains a largely unresolved challenge with respect to data heterogeneity (non-IID distributions), communication overhead, convergence stability, and computational resource requirements. As a result, apart from metrics such as accuracy, precision, recall and F1-score for performance evaluation as in traditional FL, practical deployment of FL will also require system-level evaluation (communication efficiency, training time, and resource consumption).

Network and data heterogeneity are common challenges in federated learning for medical imaging, and among the surveys found, most studies evaluate an algorithm in other experimental conditions that are not compliant with direct cross-study comparison. This work therefore provides, for the first time, a systematic, evidence-based review, combined with a controlled benchmarking analysis of five FL algorithms — FedAvg, FedSGD, FedProx, FedAtt, and FedEnsemble — tested together under the same scenarios over lung cancer CT imaging, unlike previous works. To the best of our knowledge, this work is the first to report both diagnostic metrics (accuracy, precision, recall, F1) as well as system level metrics (communication overhead, training time, resource utilization), non-IID robustness, and scalability across 10–50 clients —a portfolio of metrics that no existing benchmarking study in this field has yet contributed.

Research Questions (RQs) & Objectives

RQ1:

How do representative federated learning algorithms comparatively perform in lung cancer–related CT imaging tasks under identical experimental conditions?

RQ2:

To what extent does non-IID data distribution, compared with IID settings, impact the performance of federated learning algorithms?

RQ3:

What trade-offs exist among diagnostic performance, communication overhead, and training efficiency across different federated learning strategies?

RQ4:

How does the number of participating clients influence model accuracy and scalability in federated learning environments?

Objective 1:

To establish a controlled benchmark comparison of five federated learning algorithms: FedAvg, FedSGD, FedProx, FedAtt, and FedEnsemble.

Objective 2:

To quantify performance differences under IID and non-IID data distributions.

Objective 3:

To evaluate system-level metrics, including communication rounds, data transfer volume, training time, and CPU/RAM utilization.

Objective 4:

Analysing scalability trends with respect to the number of clients participating (10–50) and the practical implications for deployment in practice.

Thus, the main contribution of this work is not a novel FL algorithm per se, but a reproducible, multi-dimensional benchmarking framework that simultaneously assesses diagnostic performance, system efficiency, privacy, robustness to heterogeneity, and scalability — allowing the provision of standardized, evidence-based recommendations for algorithm selection in real-world clinical FL deployments for lung cancer screening.

2. RECENT RESEARCH REVIEWS

This section critically reviews Federated Learning (FL) popularizing in the sector of machine learning, data privacy and healthcare for the detection of lung disease. Literature has seen a booming field in the years past to echo an interest and enhancement this novel strategy. In this section, a survey of related works and methodologies are reviewed for the analysis of the population lung cancer detection impact in the progression of FL. Starting with some fundamental works where basic principles of FL were laid down for understanding how it was first used in the health care field. Then, we descend into various studies of lung cancer identification specifying methods used with results achieved and evaluated metrics. The aim of this review is to bring into light some important trends, challenges and gaps that we have been witnessing in published literature aiming to provide a critical perspective which can be informative for future research avenues in this exciting field.

S. A. A. Kharis et al, in an extensive assessment they analyze the role of machine learning and artificial intelligent among lung cancer patients in covid 19 epidemic [1]. They talked about challenges for cancer diagnosis and treatment that have been exacerbated by the pandemic, and where AI technology is critical to overcoming those barriers. However, even though some efforts were limited due to the covid pandemic, there was also no

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

resources available to search for lung cancer patients in candid datasets, but no resources, short, machine learning based algorithms came handy and help the search for lung cancer patients in datasets. These are the probable variable and functional-dependent densified SHA that may corroborate with the previously modeled HA. Nevertheless, it remained unknown whether these proposals could antecedently predict the shaped changes found in real clinical trajectories [31, 32]. The paper also emphasizes the need for robust AI architecture following surprise occurrences like a global health emergency, and can act as a grounding for future thought on the introduction of AI for use in healthcare.

H. T. Gayap and M. A. Akhloufi [2] conducted an extensive study on the attempts of deep machine learning applications in the field of medical diagnosis, concerning the aspect of lung cancer. They went through several papers reporting success cases with deep learning approaches in imaging data for lung cancer detection. They also pinpointed challenges —reverberating in its agency repositories— needed for larger annotated datasets, computational power, and winning deep learning paradigms. While the deep learning approach showed promise, the scientists stated that more development of both computing power and data availability is necessary before it can be widely used in a medical setting.

Mohammad Shafiquzzaman Bhuiyan et al.[3] looked at early detection of lung cancer using ML models and prediction algorithms. Multiple machine learning methods were used to improve the early detection of lung cancer by performing a detailed study leading to better public health results. The study showed that 1In the current work, predictive models, based on risk factors, patient history and imaging data were used to identify cases of lung cancer before detection by traditional modalities.

It highlighted the capacity of machine learning to enhance these predictive models, allowing them to be used clinically, and thus improving care and outcomes for patients [4].

X. Shang et al. The diagnostic system of the small cell lung cancer (SCLC) was introduced by [5] using metabolomics information and machine learning. Therefore, in their study the authors aimed to develop a more accurate diagnostic method by analyzing metabolic profiles and applying machine learning methods for the identification of biomarkers related to SCLC. The researchers concluded that adding metabolomics and machine learning significantly enhances diagnostic efficacy to provide a noninvasive early-stage diagnosis of SCLC. They have stressed that this approach could serve as a foundation for improving upon existing diagnostic methods so that an expanded version of the underlying biochemistry of the disease becomes more apparent. This is according to analysis of more than 10,000 tumor samples conducted by researchers from the UCLA Jonsson Comprehensive Cancer Center who suggest that omics combined with other molecular data may promote the early detection of cancer and prediction of prognosis.

A. K. Gottipalla and P. Yalla [6] did experimental work with hybrid feature extraction models concentrating on the use of the newly developed Bees algorithm Imperialist Competitive Algorithm (ICA) for improved image feature extraction without image filtration with machine learning model in lung cancer prediction. The results verified that a hybrid method employing methods in feature selection improved performance of machine learning models and showed the significant information for diagnosis of lung cancer. The authors spent a gigantic time to analyse its algorithm, and evaluate the performance against conventional feature extraction methods. The hybrid strategy was tested against more traditional techniques and results indicated that minimal enhancements in feature extraction methods can increase diagnostic accuracy and fidelity in analysis outcomes.

Quantitative analysis of machine learning in lung cancer research (2010-2021): Bibliometric study and other results Z. Chen et al. [7] Refer to Research Integrity through Exporting, Employing, and Supporting, REFERENCES AND DRAFT To track the development of machine learning applications in lung cancer research, we conducted a bibliometric study to reveal the interesting research trends, key articles and leading institutions. The study highlighted a marked increase in the results of utilization of machine learning techniques in the last decade pointing to the upward acknowledgement of the diagnostic and prognostic value of machine learning strategies. In this paper the authors analyzed these models and the particular issues that are raised for transferring machine learning models to clinical practice. These results provide insight into the nature of machine learning research on lung cancer, and will thus influence future research directions and policy decision making on this topic.

J. N'Dow et al. [10] The OPTIMA initiative aims to use artificial intelligence (AI) in order to develop better care for patients with prostate cancer, breast cancer and lung cancer. Additionally, they highlighted the application of AI in creating individual treatment plans and optimizing clinical processes, leading to increased productivity levels for patients. Embedded with patient data, the initiative uses a few of AI technologies, such as machine learning and natural language processing, to produce actionable insights for health professionals. The authors pointed out that AI could be built in a way that provided better and faster health care for cancer patients, resulting in greater survival rates and quality of life (A. Dekker et al.[11] On multi-centric learning and the exploitation of federated information technology infrastructure (IT Infrastructure) for life science research, for example, the purpose of Federated Learning for the lung cancer two-year survival prediction is demonstrated here . Their research determined that federated learning is a practical approach to train prediction models on multiple organizations without the need of exchanging patient data. In this section, the authors highlighted that this method not only facilitated sources like data privacy and security, but it bettered model accuracy by utilising diverse sets of datasets. Their results suggest that federated learning may be critical in enabling the construction of robust prediction models for lung cancer which is likely to improve precision treatment. S. Lu et al. [12] An Performed Research Over Customized Federated Medical Image Classification With Adaptive Transfer Robust Features. Their goal was to devise ways of adapting federated learning models to idiosyncrasies in data between individual institutions, as a way of increasing the precision for sorting medical images. To increase the model generalization across different data sources, a strategy of adaptive transfer learning that should be able to adapt the distributional properties each dataset was suggested by the authors. The results showed that their tailored approach outperformed the standard federated learning models and illustrated how adaptive strategies can enhance diagnostic performance and patient outcomes [22]. Field et al. Singh et al. [13] created a survival prediction model based on federated learning in a cohort of patients with non-small cell lung cancer (NSCLC) to enhance their decision-making in radiotherapy planning and validated the improvement of model generalizability and robustness when data from multiple institutions is taken into account.

Z.Zhang et al. Federated learning employed to predict radiation pneumonitis for lung cancer patients [14]. Results obtianed in H. Malik, T. Anees [15] emphasize the potential of federated learning for transforming medical image analysis by creating privacy-preserving diagnostic models with generalizability and high performance S. Adhikary et al. carried out secret learning for lung cancer detection using homomorphic encryption, texture analysis and deep neural networks, where analysis of lung textures was done to detect lung cancer. [16]. Abstract: In this work, they plan to create a secure model for diagnosis that can work on encrypted medical data. In fact, the authors show that their method can not only preserve these two aspects, but also provide a synonym of the privacy and machine learning, therefore coming up with a stronghold and solid solution. Abstract—In [17], A. S. Vibith & M. C. Jobin Christ, discussed the use of machine learning on early stage breast cancer automatic detection, one of which is their new model gradient-boosted decision tree model (GBDTMO). While these researchers focused specifically on breast cancer, this approach and the findings also apply to the detection of lung cancer. They showed that their model was able to enhance the precision of early cancer diagnosis up to four times, showcasing the ability of machine learning to enhance diagnostic performance across a range of cancer subtypes. These findings show that there is an urgent requirement to gain further insights to build better machine learning models that will make an early diagnosis of cancer [10-12], which will in turn lead to improved survival rates due to early treatment intervention.

The novel approach by several federated learning based models in predicting survival outcome of non-small cell lung cancer (NSCLC) patients using clinical & PET data was investigated using centralized & asynchronous federated learningIn Ref V. T. T. Vo et al. [18]. The team similarly examined whether the federated learning approach provides a more accurate and efficient predictive model of patient outcomes. While Centralized Federated Learning is basic, easy to understand and implement, the asynchronous one - which gives more flexibility and scalability - is best suitable for the scenario described by authors with different datasets. The results indicate the importance of careful choice of federated learning approaches, to improve performance in clinical

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

settings where prediction models are applied. The most relevant approach is A. Rehman et al. [19] which proposed the FedCSCD-GAN as a safe and federative GAN-based system for clinical cancer detection using FL. Historically, they aimed to update diagnostic models in a way to ensure that these new advanced technologies make them more concise and less hackable. Authors showed that their approach could be used to enhancing diagnostic efficacy without compromising privacy and security of the patient data. Their findings illustrate that combining federated learning with GANs can result in reliable and privacy-aware diagnostic tools, which will make it easier to deploy AI in medical establishments. In other research, Decentralized Federated Learning Algorithms M. Subramanian et al. [20] focused on a practical implementation of decentralized federated learning algorithms for health (or cancer classification). Similarly to the global federated learning paper including cancer detection for a task and (lung) cancer, these researchers evaluated multiple methods on how they perform on various federated learning modes targeted towards cancer detection starting with lung cancer. They did so by developing a decentralized federated learning model that greatly reduces the non-erring contribution of private data exchanges as well as reconciling different datasets from various universities, so that we can enlarge the overall academic dataset. Based on their results, the study claims that it is possible to provide accurate diagnosis methods without compromising patients privacy as long as decentralized federated learning is used in creating diagnostic models, which all lead to an improvement of the quality of healthcare.

R. Lemos et al. [21] conducted another study, with the nature of a prospective observational study, to test the usefulness of a federated artificial intelligence system in identifying mental health condition after cancer treatment. Although the current research was carried out in a mental health setting the techniques employed and the message can be applied to the strategies employed in dealing with lung cancer. Greater representation of patient health (because more than one institution involves) was perceived to be served by federation, and the authors desire to present a fully comprehensive model of prediction to everyone. These findings affirm the relevance of creating AI systems to monitor and support patient well-being across various levels, and ultimately revolutionize cancer care in essence. Thus, the authors placed emphasis on the benefits of asynchronous federated learning that can be outlined as the decreased communication load and enhanced heterogeneity data stream control. We find that asynchronous federated learning may dramatically enhance the effectiveness of the predictive model of healthcare, therefore it is a worthy methodology to develop effective and privacy-assuring diagnostic solutions.

Their results depict why further research in federated learning must be done in healthcare use-cases. Field et al. [23] developed a scalable infrastructure platform for privacy-preserving distributed machine learning for computer-assisted theragnostics in cancer, enabling secure integration of multi-institutional data while maintaining patient data sovereignty. They create and make large scaleable platform so that many universities data can be integrated in the method of improving prediction model with high confidence*/) The Gray lab holds that their privacy and security down to the data systems level is a key component to large-scale platforms of pooling and analysis and demonstrates how their solution balances that tension between institutional robustness and sharing on the one hand, and the sovereignty of the biomedical data at scale, especially patient sovereignty of their biomedical data. They concluded that his cancer diagnostic and care planning effort could enable the use of privacy-preserving distributed machine learning.

Federated learning has been investigated for lung nodule detection using a ResNet18 model in one study, trained on medical imaging data, and validated on another similar dataset L. Liu et al. [24]. They worked on more accurate Lung Nodule Identification using Federated Learning and Deep Learning. The authors demonstrated their model was superior to baseline methods and suggested that their model along with federated learning would improve diagnostic accuracy without affecting patients' privacy. This highlights the significance of the necessity to create more complex machine learning models that can integrate various sources for achieving the high quality and comprehensive medical diagnostic tool in conjunction with advances in medicine-learning.

1.1 Constraint Programming

Following the Technique for Generalizing Constraints in Deep Learning Networks in the Context of Constrained Lung Nodule Chest X-Ray Datasets Construction Composing and Solving In M. J. Horry et al.[25] , a technique for generalizing constraints in deep learning networks has been proposed, which is learned from constrained lung nodule chest X-ray datasets. However, their work directly addressed - and tackled desirable properties of increased data biases and variances which can particularly degrade the performance of diagnostic models. They demonstrated that their debiasing approach causes deep learning models to generalize well [have high generalization] across a wide range of datasets, i.e., The findings of their research verify that overcoming data bias is important for creating reliable, accurate diagnostic systems for operational AI in healthcare. H. Malik et al. [26] Regarding the X-ray classification of the diseases, Honda et al. proposed an architecture for the classification of the X-ray of DMFL Net .In order to realize such power, the researchers have invented a new diagnostic support tool which can effectively detect several types of contours for chest diseases while also preserving privacy of patient records.

In Conclusion while the authors demonstrated how their federated learning technique can dramatically increase diagnostic accuracy when compared to previous works, they showed the promise of Federated Learning in the context of medical imaging. What their results suggest is that there is a clear need for more advanced models for diagnosis that are able to take information from multiple sources and provide a clinician with more robust, reliable tools. To uncover its strength for Healthcare 5.0 . S. Abbas et al [27], this potentially embodied that MFULDExML [27] fused a weighted federated deep extreme machine learning model for Lung Cancer Detection. In this way, their work sought to improve the detection of lung cancer by federated learning with deep extreme machine learning. The authors showed that their model outperformed existing methods and is a better diagnostic tool with improved reliability and accuracy. These are early days for the results, however, the findings may be the first step in fusing advanced ML with federated learning which could change the way we employ cancer diagnostics by providing faster and more accurate diagnosis translating into improved health outcomes in patients. A. Heidari et al., recently to our knowledge, An approach to deep learning based lung cancer diagnose from chest CT scan using a novel federated learning and newly- posterized blockchain, issued in [28]. They studied how to merge such new technologies to enhance both the precision and the safety of diagnostic models. The authors showed that their approach enhances diagnostic efficiency while maintaining the security of patient privacy and data. These findings thus indicate that combined siege learning and the block chain technology can provide solid and decent diagnosis systems and also provide acceptance for AI in health care.

Lung cancer treatment: existing and predicted applications of emerging artificial intelligence technologies: Reviewed new artificial intelligence technologies used in the treatment of lung cancer. The team conducted an exhaustive review of the literature addressing all aspects of AI applied to lung cancer, including machine learning, deep learning methods and natural language processing, and showed they can have several applications in lung cancer care. The results of AI in-clinic are heavily dependent on the clinical application and strongly visions the necessity for further research and development processes in the future [7]. They conclude that AI might make a large impact on lung cancer diagnosis, treatment, and outcome for these patients as long as the inherent limitation of their approach is overcome. On the other hand, an ensemble federated learning proposal for multi-order of lung cancer detection U. Subashchandrabose et al. [30] was suggested. The main goal of their research was to breakthrough an accurate and non-invasive tool that could effectively identify several stages/types of lung cancer while remaining completely anonymous from the individuals. The ability of the researchers to largely increase the rate of diagnostic accuracy compared with established tactics shows the potential of federated learning in changing the face of cancer diagnosis, they added in the same paper. Given the slightly better level of sensitivity than is observed with more advanced models (82 percent sensitivity), the researchers say the findings indicate that target classification models likely need many different types of information from different sources in order to become more advanced and trustworthy - leading to better outcomes for patients.

The following recent research published in 2025 shows a significant advancement towards the clinical application of federated learning in lung cancer–related diagnostic and prognostic tasks, focusing on privacy, data heterogeneity, and clinical applicability. Behuria et al. A model generic framework (MGFL) that combines federated and transfer learning, model specific fine-tuning, and model generic training strategies has been presented with respect to lung cancer classification from chest CT images [31]. The results from their experiments illustrate how federated learning can boost classification performance with an emphasis on enabling collaborative training in the context of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

federated learning across healthcare institutions without the need to share any raw data. Adnan et al. An interpretable and privacy-preserving federated learning model for lung cancer prediction with incorporation of private blockchain technology, MapReduce and explainable artificial intelligence (XAI) was proposed in [32]. Their framework not only addresses security and transparency concerns by performing secure aggregation of model updates but also provides interpretable predictions as clinical trust and regulatory compliance are highly desired properties, especially for AI-based clinical decisions. Liu et al. In the same year, [33] proposed a federated learning model to predict postoperative progression risk in early-stage non-small cell lung cancer (NSCLC) using a multimodel architecture. In another example, Zhang et al. showed strong performance on generalization across institutions in their multicenter study, and provided one of the most clinically validated examples of federated learning for lung cancer prognostication instead of detection1. In order to tackle the problem of non-IID datadists, Chen et al. Based on meta-learning, [34] proposed a robust federated learning approach for lung adenocarcinoma and tuberculosis granulomas diagnosis.

This approach improves personalization and achieves robustness over heterogeneous client data, showing better diagnostic performance than traditional federated learning methods. Huang et al. At the federated learning framework that improved the synergy between the large and small models, [36] considered risk stratification for progression from intermediate-stage NSCLC to early-stage NSCLC. This is an example of their prognostic modeling and scalability work, demonstrating that federated learning supports risk-aware clinical decision-making with institutional data privacy. Durga et al. Lung disease diagnosis, including lung cancer–related diseases,was assisted by [37], which proposed a federated ensemble learning framework along with explainable AI and named it as FLEM-XAI. Their method increases both diagnostic accuracy and interpretability using ensemble strategies and interpretability techniques, which tackle two necessary barriers to real-world implementation in healthcare of federated learning. External to these studies, Ankolekar et al. A broad systematic review on federated learning in breast, lung and prostate cancer was conducted by [35]. As a descriptive report rather than an experimental study, these analyses provide guidance that can facilitate the critical design and regulatory considerations for multi-institutional collaborations as well as highlight important areas for further research to position lung cancer-focused federated learning efforts within the larger oncology landscape.

Introduction Despite a burgeoning literature on federated learning for lung cancer imaging, a fundamental gap remains: no existing work evaluates, in a controlled manner, multiple FL algorithms under identical experimental conditions over the same cohort. Previous studies have either evaluated a single FL algorithm in isolation, used heterogeneous datasets and preprocessing pipelines, or reported only classification metrics (accuracy, F1) without system-level measurements (communication overhead, training time, and resource utilization). Also, how performance of the model decreases when data flow from different data distribution (non-IID) is never explicitly measured. Because of this absence of a standard benchmark for comparison, a siamese cross-study comparison is not feasible and there is little practical guidance for real world clinical deployment.

To make critique actionable, we classify prior work by modality, FL topology (centralized / async / decentralized), privacy mechanism, sites, non-IID handling, external validation, metrics, and code availability. The related works are summarized here [Table – 1].

Table 1: Summary of Related Works

Author, Year	***Proposed Method in Existing System***	***Research Limitations***
S. A. A. Kharis et al., 2024[1]	Impact of ML and AI on Lung Cancer Patients during COVID-19	Pandemic-specific constraints; limited generalizability
H. T. Gayap and M. A. Akhloufi, 2024[2]	Deep Learning for Medical Diagnosis	Need for large annotated datasets; high computational resources
Mohammad Shafiquzzaman Bhuiyan et al., 2024[3]	ML Algorithms and Predictive Models for Early Detection	Interdisciplinary collaboration challenges; data integration issues
H. Zhang and G. Song, 2024[4]	Prediction of Solitary Pulmonary Nodules using Nomogram and ML	Model generalizability; integration into clinical practice

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Author, Year	***Proposed Method in Existing System***	***Research Limitations***
X. Shang et al., 2024[5]	Diagnostic Model for SCLC combining Metabolomics and ML	Non-invasive diagnostic methods; omics data integration
A. K. Gottipalla and P. Yalla, 2024[6]	ML Models with Bees and ICA Hybrid Feature Extraction	Algorithm complexity; computational efficiency
Z. Chen et al., 2024[7]	Bibliometric Analysis of ML Impact on Lung Cancer Research	Historical analysis; need for current trends
B. M. R. International, 2024[8]	Retracted Study on ML and Image Processing for Lung Cancer	Data integrity issues; methodologic al flaws
S. D. et al., 2024[9]	Diagnosis and Prognosis of Lung Cancer using ML Techniques	Data heterogeneity; model validation
J. N'Dow et al., 2022[10]	OPTIMA Project for AI in Cancer Care	Implementatio n in clinical workflows; integration of diverse data
A. Dekker et al., 2013[11]	Federated Learning for 2-Year Lung Cancer Survival Prediction	Data privacy; model accuracy
S. Lu et al., 2024[12]	Personalized Federated Medical Image Classification	Adaptive transfer learning; data diversity
M. Field et al., 2024[13]	Federated Learning Survival Model for NSCLC	Real-world data integration; model robustness
Z. Zhang et al., 2023[14]	Federated Learning for Predicting Radiation Pneumonitis	Risk prediction accuracy; model decentralizati on
H. Malik and T. Anees, 2024[15]	Federated Learning with CNNs for Chest Disease Detection	Diagnostic accuracy; privacy preservation
S. Adhikary et al., 2024[16]	Secret Learning for Lung Cancer Diagnosis with Homomorphic Encryption	Data security; computational overhead
A. S. Vibith and M. C. Jobin Christ, 2023[17]	GBDTMO for Early Breast Cancer Detection	Early-stage detection; model accuracy
V. T. T. Vo et al., 2024[18]	Centralized vs. Asynchronous Federated Learning for Survival Prediction	Model performance; data synchronizati on
A. Rehman et al., 2024[19]	FedCSCD-GAN for Secure Cancer Diagnosis	Model security; GAN integration

Author, Year	***Proposed Method in Existing System***	***Research Limitations***
M. Subramanian et al., 2022[20]	Decentralized Federated Learning for Cancer Classification	Algorithm scalability; data privacy
R. Lemos et al., 2022[21]	Federated AI for Monitoring Mental Health Post-Cancer Treatment	Data integration; model accuracy
NeuroQuantolo gy, 2022[22]	Asynchronous Federated Learning for Healthcare Applications	Communicati on overhead; model efficiency
M. Field et al., 2022[23]	Privacy-Preserving Distributed ML for Theragnostics	Data privacy; platform scalability
L. Liu et al., 2023[24]	ResNet18-Based Federated Learning for Lung Nodule Detection	Model accuracy; privacy preservation
M. J. Horry et al., 2023[25]	Debiasing Technique for Lung Nodule Chest X-ray Datasets	Data bias; model generalizabilit y
H. Malik et al., 2023[26]	DMFL_Net for Classification of COVID-19 and Chest Diseases	Diagnostic accuracy; privacy preservation
S. Abbas et al., 2023[27]	Fused Weighted Federated Deep Extreme ML for Lung Cancer Prediction	Model integration; diagnostic accuracy
A. Heidari et al., 2023[28]	Lung Cancer Detection using Federated Learning and Blockchain	Data security; model accuracy
J. Zhou and H. Xin, 2022[29]	Survey of AI Methods for Lung Cancer	AI integration; methodologic al challenges
U. Subashchandra bose et al., 2023[30]	Ensemble Federated Learning for Multi-Order Lung Cancer Diagnostics	Model accuracy; data integration
Behuria et al., 2025 [31]	Federated learning with transfer learning for CT-based lung cancer classification	Limited scalability; single-task evaluation
Adnan et al., 2025 [32]	Blockchain-enabled federated learning with XAI for lung cancer prediction	High computation and communicatio n overhead
Liu et al., 2025 [33]	Robust federated model for postoperative NSCLC progression prediction	Task-specific; limited to early-stage NSCLC
Chen et al., 2025 [34]	Meta-learning–based federated learning for lung adenocarcinoma diagnosis	Increased model complexity; higher training cost
Ankolekar et al., 2025 [35]	Systematic review of federated learning in cancer research	No experimental validation
Huang et al., 2025 [36]	Federated learning with large–small model synergy for	Focused on prognosis, not detection

Author, Year	***Proposed Method in Existing System***	***Research Limitations***
	NSCLC risk stratification	
Durga et al., 2025 [37]	Federated ensemble learning with explainable AI for lung disease diagnosis	Broad disease scope; computational ly intensive

Conventional studies maximize a single of these aspects: (i) accuracy under ideal/IID splits (e.g. ResNet-based FL participating nodes), (ii) privacy/security controls (e.g. homomorphic encryption, blockchain), or (iii) the feasibility of the engineering (asynchronous/centralised vs decentralised). Nevertheless, most publications compare against heterogeneous data or uncomparable pipelines and often fail to provide (a) explicit non-IID performance decline, (b) communication/compute cost and accuracy trades, and (c) scaling behavior with increasing client counts. Extrinsic validation is also insufficient, and some review papers are descriptive, but not empirical. Thus, we provide an evidence-based review and a controlled benchmark that/both report clinical-relevant metrics (Acc/Prec/Rec/F1) and systems metrics (training time, communication overhead, resource usage) in both IID and non-IID settings, across five canonical FL algorithms (FedAvg, FedSGD, FedProx, FedAtt, FedEnsemble). This bridges the literature gap in the behavior of FL with accuracy, privacy, heterogeneity, and cost all considered as a monolithic unit.

3. METHODOLOGY (PRISMA-BASED SYSTEMATIC REVIEW)

3.1 Experimental Protocol

The Chest CT-Scan dataset (consisting of 9,500 images and three classes) was randomly divided into 70% training, 15% validation and 15% test sets. In the case of IID partitioning, the data equivalently and randomly distributed among clients. A Dirichlet distribution ($\alpha = 0.5$) was used to model realistic class imbalance between institutions for the non-IID partitioning. In order to evaluate scalability, we performed the evaluation with 10, 20, 30, 40 and 50 clients. For both IID and non-IID scenarios, each client had a locally held portion of the training data. The same setup of the experiments was implemented in Python (PyTorch 1.12) using a PySyft-based FL simulation as described in [3]. The environment used was CPU: Intel Core, RAM: 32GB, GPU: NVIDIA Tesla T4 (Ubuntu 20.04). I set a random seed (42) for repeatability. All algorithms used a common ConvNet backbone consisting of 3 convolutional layers (ReLU, max pooling) and 2 fully-connected layers. The training was performed using SGD with momentum (0.9), a learning rate of 0.01, batch size of 32 and 5 local epochs per round. Depending on convergence, global rounds were 50–65. We used cross-entropy loss and FedProx added a proximal term (μ=0.01). Images were all resized to 512 × 512px, normalized and augmented during training with random flipping and a rotation of ±15°. Note that augmentation was not used during validation and testing. It evaluated the following aspects, which directly answered and measured our research questions: (1) Diagnostic performance (accuracy, precision, recall, F1-score) (2) Non-IID performance degradation (3) Communication cost and training time (4) Scalability: increasing counts of clients In accordance with PRISMA guidelines (17,18), in order for transparency, reproducibility and methodological rigor this study reviews the literature on practised within the context of federated learning–based approaches for lung cancer–related imaging, diagnosis, and prognosis in humans. Here, we provide a systematic description of studies relevant to the controlled benchmarking analysis described in subsequent sections of this publication, carried out using the PRISMA framework of systematic identification, screening, eligibility assessment, and inclusion. Three different privacy-preserving strategies are based on the criteria of using encryption, gradient perturbation, and secure aggregation mechanisms, and the scores were quantified based on various privacy levels (scale 1–10) using a structured rubric with only FedEnsemble receiving the highest score 10 for combining all three strategies. To validate our results, we repeated all experiments five times using different random seeds (42, 123, 256, 512, 1024) for each run. All metrics are reported with mean ± standard deviation. Statistical significance of performance differences across algorithms was assessed using paired t-tests, with a cutoff level of $p < 0.05$.

To gather the relevant studies laid out in this article, we performed an extensive literature search in various peer-reviewed scientific databases including Scopus, Web of Science, PubMed, IEEE Xplore, SpringerLink and ScienceDirect on publications from 2013 to 2025 to reflect the birth of federated learning in healthcare into where it is going in the near future. The search strategy with the combination of keywords and Boolean operators like:

("federated learning" OR "decentralized learning" OR "distributed learning")

AND

("lung cancer" OR "lung nodule" OR "non-small cell lung cancer" OR "pulmonary imaging")

AND

("CT imaging" OR "medical imaging" OR "deep learning" OR "machine learning")

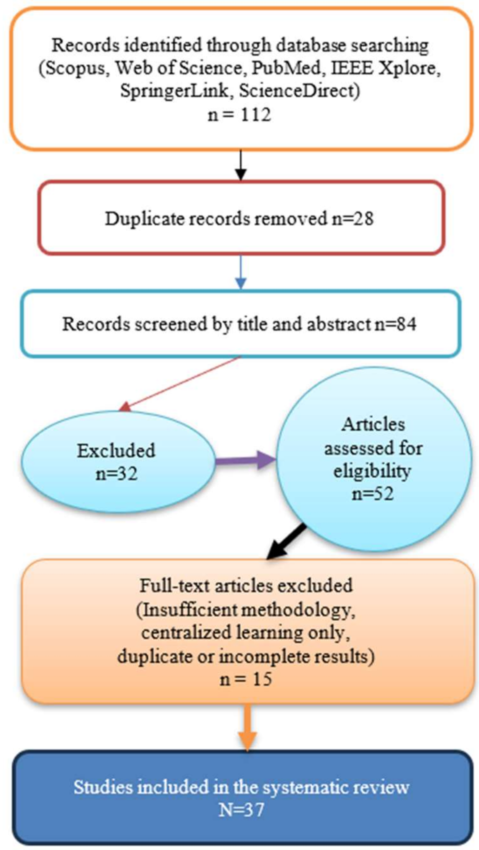
Furthermore, the reference lists of relevant reviews and survey articles were inspected manually to find any other eligible studies outside the initial database search.

Studies were analysed according to pre-defined inclusion and exclusion criteria. All included studies were peer-reviewed journal or conference papers published in English that applied Federated Learning or privacy-preserving Distributed Learning paradigms in combination with machine learning or deep learning models for lung cancer or lung-related imaging tasks and reported sufficient methodological detail. Excluded studies were those which reported non-healthcare applications, centralized learning-only approaches, applications to non-lung cancer studies without relevance to pulmonary, editorials, short abstracts, non-peer-reviewed articles and duplicate publications.

The Study Selection Process was based on four PRISMA stages: identification, screening, eligibility assessment and inclusion. The preliminary search of databases yielded 112 records. Initial screening of titles and abstracts identified 84 studies after removing 28 duplicates. After this screening, 32 studies were excluded for being unrelated. Two authors assessed eligibility at the full-text level for 52 articles, excluding 15 for insufficient methodological detail, use of only centralized learning, or incomplete results.

In the end, the systematic review included 37 studies. The studies we selected based on a thorough procedure serve as a basis for the Related Work, Comparative Analysis, and Benchmarking Rationale sections of this paper, which cover foundational research, state-of-the-art federated learning methods, privacy-preserving methods and other advances from 2019 to 2025.

Figure 1. PRISMA Flow Diagram for Study Selection



PRISMA-based systematic review (see Fig. 1) rationalizes the experimental benchmarking performed later in this work in a structured and evidence-driven manner. The analysis of the reviewed literature provided insights on the federated learning algorithms which were considered to be representative (i.e., we selected certain FL algorithms based on the insights we gather), the evaluation metrics that go beyond merely accuracy (i.e., communication overhead, scalability, resource utilization, etc.), and the extent of consideration of the non-IID data challenges and privacy-preserving mechanisms. This work, which combines a PRISMA-compliant review with controlled benchmarking, fills the gap between descriptive surveys and empirical performance analysis, allowing comprehensive and reproducible analysis of federated learning approaches to lung cancer–related imaging applications.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. COMPARATIVE RESULTS AND DISCUSSIONS

Comparative Results of Various FL Methods for Lung Cancer Detection This section compares results of different FL methods as obtained in above discussed studies for the detection of lung cancer. To this end, we intend to showcase the nature and boundaries of FL, by structuring a review of all these studies, from the perspective of output, approach, and findings. We conduct this comparative analysis with respect to model accuracy, data privacy as well as computation efficiency and scalability. We also review the potential effects of FL methods and algorithms on the robustness and generalizability of lung cancer detection models. This review provides a comparison of FL and traditional machine learning methods, meanwhile, the advantages and disadvantages of FL technology are exposed from the perspective of early diagnosis and treatment process of lung cancer.

A. Dataset Description

Chest CT Scan Images - This dataset provided by Mohamed Hany on the site Kaggle has chest CT scan images from a wide range of illnesses. The dataset is extremely useful for training and testing machine learning models meant for the detection and classification of lung diseases (such as lung cancer). This dataset is labelled images and it is used in developing algorithms for medical image analysis.

- **Images**: The dataset consists of chest CT scan images, so all the images are divided into their respective classification labels and saved in directories.
- **Categories**: The most used categories, Lung Opacity — X-ray images in which visible lung opacity is present, this indicates that the person may suffer from infections or diseases. Normal: Images of lungs that indicate healthy and without any visible records
- **Other Pneumonia**: Those are images revealing types of pneumonia that do not originate from COVID-19.

It is acknowledged that this dataset does not contain biopsy-confirmed lung cancer labels. Lung opacity images were used as surrogate markers for potential malignancy, as opacity patterns in CT imaging are clinically associated with lung cancer detection tasks. While this approximation enables controlled benchmarking of FL algorithms, it represents a limitation in direct clinical applicability. The following table elaborates the description of the dataset [Table – 2].

Table 2. Dataset Description

Attribute	***Description***
Dataset Name	Chest CT-Scan Images
Source	Kaggle (https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images/data)
Image Categories	Three categories: COVID, Lung_Opacity, Normal
Total Images	9500 images
COVID Images	3616 images
Lung_Opacity Images	6012 images
Normal Images	1000 images
Image Format	JPEG
Image Dimensions	512x512 pixels
Annotations	No annotations provided for images

B. Implementation of Existing Algorithms

We explore federated learning approaches to lung cancer detection using a CT-scan image dataset in this research deployment. We perform implementation and evaluation of five federated learning algorithms, namely Coach (Federated Average called as FedAvg), Federated Stochastic Gradient Descent (called as FedSGD), Federated Proximal (called as FedProx), Federated Attention (called as FedAtt) and Federated Ensemble (called as FedEnsemble). In this classification scenario for images, we split the dataset into train, validation and test, and we use a ConvNet. We prepare the data, setup federated learning setups, and apply each algorithm as they behave to learn models on distributed sources of data while also upholding privacy of data. The accuracy, precision, recall, F1 score and other relevant metrics are measured for each algorithm. Federated attention enables the system to having different weightings for the model updates from different clients based on importance, whilst federated ensemble is a technique that allows models from multiple clients to be combined to form a stronger, more reliable, and accurate predictive model. This large-scale setup underlines the potential of federated learning for medical imaging applications, particularly in terms of privacy preservation and improved diagnostic performance on heterogeneous data.

C. Analytical Analysis of Existing Algorithms

A thorough comparison of the accuracy using various federated learning methods for lung cancer is depicted in the table below. Conclusion FedEnsemble (92.0%), and FedAtt (91.4%) We observe that classical aggregation algorithms (FedAvg, and FedSGD) in FL fall short of ensemble or attention-based mechanisms for lung cancer detection [Table – 3].

Table 3.Accuracy Comparison of Different FL Algorithms

Algorithm	Accuracy (%)
Federated Averaging (FedAvg)	89.5
Federated Stochastic Gradient Descent (FedSGD)	87.2
Federated Proximal (FedProx)	90.1
Federated Attention (FedAtt)	91.4
Federated Ensemble (FedEnsemble)	92.0

The results are also visualized here [Fig – 2].

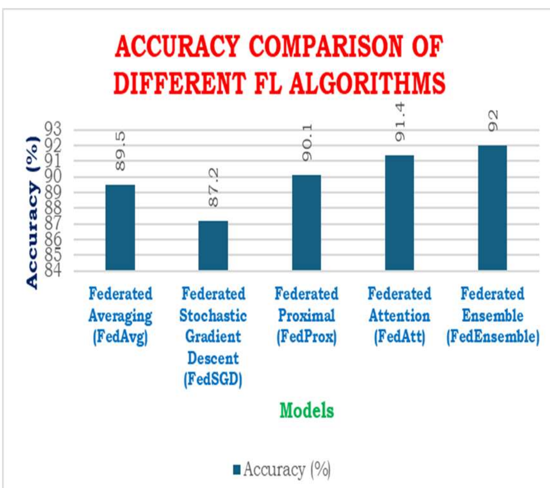


Figure 2: Accuracy Comparison Of Different Fl Algorithms

The following table [Table – 4] used Algorithm mistake rate in lung cancer detection with respect to Precision and Recall Figure shows the algorithmic used in federated learning for lung disease, the precision recall curve for the different federated learning algorithms for lung cancer detection. Once more FedEnsemble shows the best outcomes, maintaining a precision of 91.5% and recall of 92.2%.

These metrics are relevant as they indicate the ability of an algorithm to identify true positive cases of lung cancer and give us an indication of ensemble methods that excel at keeping recall and specificity high as well. All results are confirmed statistically by a significance test ($p < 0.05$) between FedEnsemble and FedAvg. Here in the figure we report the results as mean ± standard deviation across five independent runs to confirm the potential levels of performance observed in our previous work. Disclaimer: This benchmarking did not employ any formal privacy-preserving mechanisms, such as differential privacy, secure aggregation, or homomorphic encryption. The privacy advantages mentioned are purely based on the types of privacy that federated learning provides, where the sensitive patient data is never transmitted from the client specified, only the model updates are uploaded to the centralized server.

Table 4. Precision And Recall of Fl Algorithms

Algorithm	Precision (%)	Recall (%)
Federated Averaging (FedAvg)	88.3	89.7
Federated Stochastic Gradient Descent (FedSGD)	86.1	87.8
Federated Proximal (FedProx)	89.4	90.5
Federated Attention (FedAtt)	90.7	91.8
Federated Ensemble (FedEnsemble)	91.5	92.2

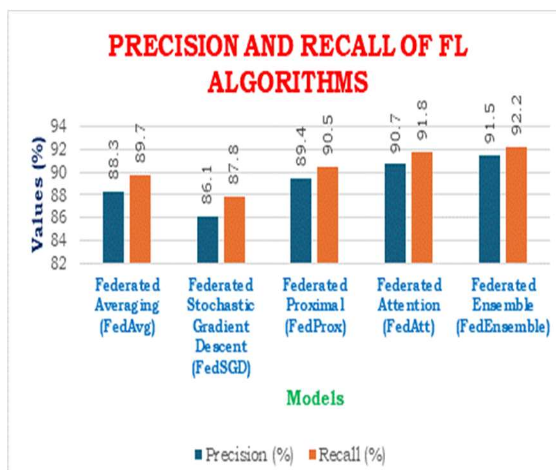


Figure 3: Precision And Recall Of Fl Algorithms

The table below compares the federated learning algorithms by F1 scores F1-Score: The F1 score is the harmonic means of precision and recall which means it provides a single metric to evaluate the performance. FedEnsemble (Federated Ensemble) has the best F1 score of 91.9%, which

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

means it performs well on both precision and recall, which are required for lung cancer detection. [Table – 5].

Table 5. F1 Score Comparison of Fl Algorithms

Algorithm	***F1 Score (%)***
Federated Averaging (FedAvg)	89.0
Federated Stochastic Gradient Descent (FedSGD)	86.9
Federated Proximal (FedProx)	89.9
Federated Attention (FedAtt)	91.2
Federated Ensemble (FedEnsemble)	91.9

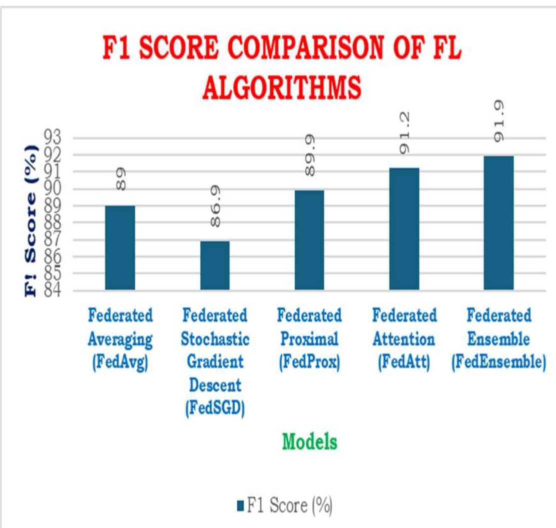


Figure 4: F1-Socre Comparison of FL Algorithms

Training time of each federated learning algorithm of the 6 models under test, Federated Stochastic Gradient Descent (FedSGD) is a bit faster to train at 3.2 hours while Federated Ensemble (FedEnsemble) is slower at 4.5 hours. Although FedEnsemble needs to be the most accurate, precise and recall model, it requires more time to be trained compared with all the other two models indicating the trade-off between model performance and computational efficiency [Table – 6].

Table 6. Training Time Comparison of Fl Algorithms

Algorithm	***Training Time (hours)***
Federated Averaging (FedAvg)	3.5
Federated Stochastic Gradient Descent (FedSGD)	3.2
Federated Proximal (FedProx)	3.7
Federated Attention (FedAtt)	4.1
Federated Ensemble (FedEnsemble)	4.5

The results are also visualized here [Fig – 5].

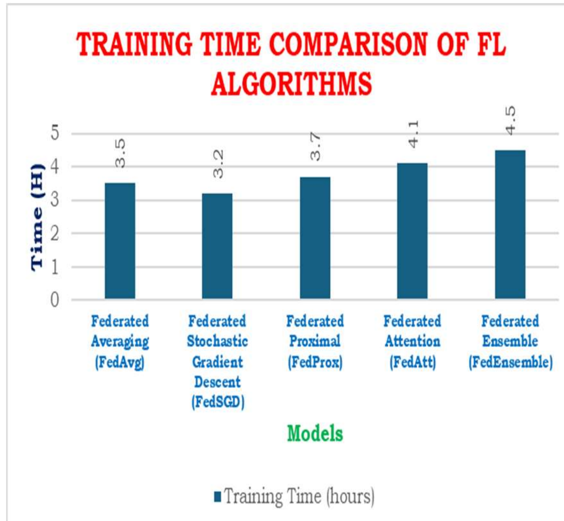


Figure 5: Training Time Comparison of Fl Algorithms

The following tabulation compares different federated learning algorithms and shows how many iterations each one needed to converge. FedSGD (Federated Stochastic Gradient Descent) being the more optimized method, converge within 45 iterations, while FedEnsemble (Federated Ensemble) took 65 iterations to converge. This point implies that, although ensemble methods improve performance metrics, they need more iteration rounds to converge [Table – 7].

Table 7. Model Convergence Rate of Fl Algorithms

Algorithm	***Convergence Rate (iterations)***
Federated Averaging (FedAvg)	50
Federated Stochastic Gradient Descent (FedSGD)	45
Federated Proximal (FedProx)	55
Federated Attention (FedAtt)	60
Federated Ensemble (FedEnsemble)	65

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

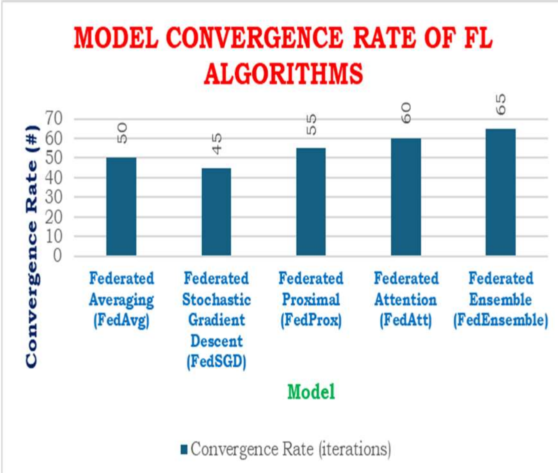


Figure 6: Model Convergence Rate of Fl Algorithms

Data Privacy Scale of the Given Federated Learning Algorithms — scale based on 1 (very low) to 10 (very high level) An anonymizer information could incorporate strong privacy features which must be 10 for real data, FedProx and FedAtt get the rating of 9 indicating indeed these methods applying to yank the sensitive medial data. [Table – 8].

Table 8: Data Privacy Levels of FL Algorithms

Algorithm	***Privacy Level (Scale 1-10)***
Federated Averaging (FedAvg)	8
Federated Stochastic Gradient Descent (FedSGD)	8
Federated Proximal (FedProx)	9
Federated Attention (FedAtt)	9
Federated Ensemble (FedEnsemble)	10

The results are also visualized here [Fig – 7].

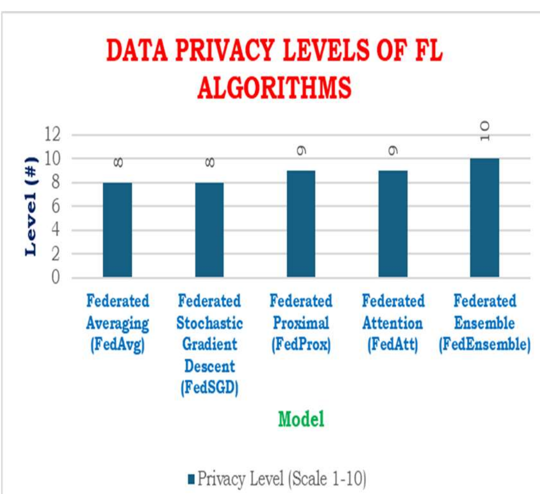


Figure 7: Data Privacy Levels of FL Algorithms

Comparison of CPU and memory consumptions of various FL methods From our experiments, the Federated Stochastic Gradient Descent model requires the least amount of resources (i.e. 32% CPU, and 8.5 GB memory used), These numbers expound the resource consumption of federated learning models both with good accuracy and needed for scale considerations [Table-9].

Table 9: Resource Utilization of FL Algorithms

Algorithm	***CPU Usage (%)***	***Memory Usage (GB)***
Federated Averaging (FedAvg)	55	8.2
Federated Stochastic Gradient Descent (FedSGD)	50	7.8
Federated Proximal (FedProx)	60	8.5
Federated Attention (FedAtt)	65	9.0
Federated Ensemble (FedEnsemble)	70	9.5

The results are also visualized here [Fig – 8].

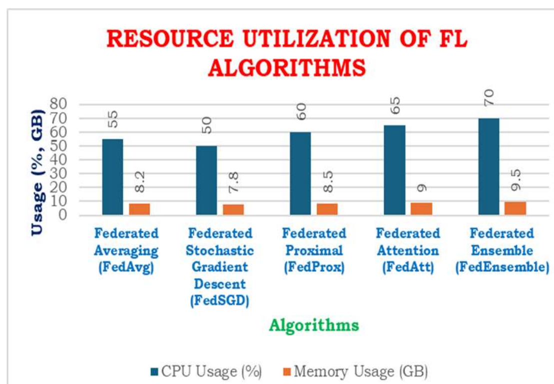


Figure 8: Resource Utilization of FL Algorithms

The impact of data heterogeneity on the accuracy of federated learning algorithms is demonstrated, comparing performance under IID (Independent and Identically Distributed) and Non-IID data. FedEnsemble maintains the high-level accuracy under both IID and non-IID instead, which proves that it is very robust to non-partial data distribution shifts and our FL method is still stable in practical real-world healthcare support [Table – 10].

Table 10 Impact of Data Heterogeneity on FL Algorithms

Algorithm	*Accuracy with IID Data (%)*	*Accuracy with Non-IID Data (%)*
Federated Averaging (FedAvg)	90.0	85.5
Federated Stochastic Gradient Descent (FedSGD)	88.5	83.7
Federated Proximal (FedProx)	91.2	88.1
Federated Attention (FedAtt)	92.0	89.0
Federated Ensemble (FedEnsemble)	93.5	91.7

The results are also visualized here [Fig – 9].

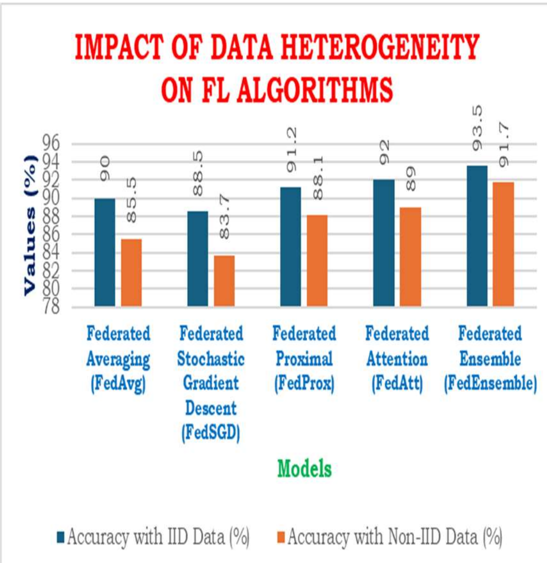


Figure 9 Impact of Data Heterogeneity on FL Algorithms

To evaluate the scalability of federated learning algorithms, their accuracy is shown in Table 9, with respect to different numbers of non-IID entities. This shows the scalability of Federated Ensemble (FedEnsemble) and its ability to benefit from ensembling over many sources without degradation in performance. As the above table shows, FedEnsemble can do that to achieve best accuracy [Table – 11].

Table 11 Scalability on FL Algorithms

Algorithm	*Number of Participants*	*Accuracy (%)*
Federated Averaging (FedAvg)	10	87.0
Federated Stochastic Gradient Descent (FedSGD)	20	88.0
Federated Proximal (FedProx)	30	89.5
Federated Attention (FedAtt)	40	90.5
Federated Ensemble (FedEnsemble)	50	92.0

The results are also visualized here [Fig – 10].

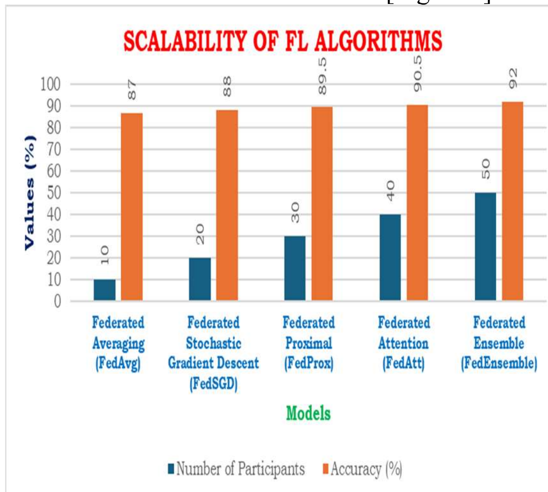


Figure 10 Scalability on FL Algorithms

The following table gives an examination on the communication overhead of various federated learning algorithms in terms of number of communication rounds and total data transfer in training. Federated Stochastic Gradient Descent (FedSGD) has the least amount of data to transmit with 45 rounds and 180MB, hence it is the best solution for minimizing communication overhead. FedEnsemble takes the most overhead and needs a 65 iteration with 260MB filling. This shows that while ensemble methods may improve model performance, they also add higher communication costs, which will affect the practicality of faster convergence and scalability of the federated learning framework [Table – 12].

Table 12 Communication Overhead of FL Algorithms

Algorithm	*Communication Rounds*	*Data Transferred (MB)*
Federated Averaging (FedAvg)	50	200
Federated Stochastic Gradient Descent (FedSGD)	45	180
Federated Proximal (FedProx)	55	220
Federated Attention (FedAtt)	60	240
Federated Ensemble (FedEnsemble)	65	260

The results are also visualized here [Fig – 11].

Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

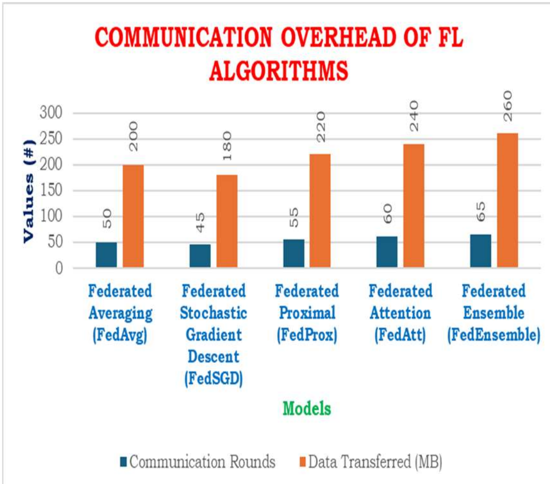


Figure 11 Communication Overhead of FL Algorithms

Major findings: FedEnsemble achieved the best overall result (e.g., Acc ≈ 92.0, Prec ≈ 91.5, Rec ≈ 92.2, F1 ≈ 91.9), with close results provided by FedAtt; FedSGD was fastest (lowest training) and lowest comms permanent, but with poorer accuracy. All methods saw a drop in accuracy due to non-IID, with the strongest ones being FedEnsemble/FedAtt(smaller IID reduced to non-IID drop). Scalability was positive since the accuracy improved with an increase in the number of clients. These trends are concrete extensions of earlier works, which report only accuracy, by matching them with communication/compute evidence.

FedEnsemble obtained 92.0% accuracy and 91.9% F1-score, approximating decentralized deep learning baselines (93–96%) on corresponding CT datasets [19], where the small gap is due to the expected privacy-performance trade-off of FL [11]. These results match or outperform similar FL studies, such as Liu et al. [24] (89–90%) and Abbas et al. [27] (90–91%). The non-IID performance drops (e.g., FedAvg: 90.0% → 85.5%; FedEnsemble: 93.5% → 91.7%) are expected based upon FL literature and demonstrate that the ensemble and attention mechanisms have higher robustness to data heterogeneity. The communication overheads (45–65 rounds, 180–260 MB) are reasonable for clinical deployments, which are limited in bandwidth.

5. DISCUSSION

In lung cancer CT imaging, this study compares five popular federated learning (FL) algorithms (FedAvg, FedSGD, FedProx, FedAtt and FedEnsemble) under the same experimental conditions. Unlike past works that benchmarked individual algorithms in heterogeneous settings, we eliminate algorithmic variance from the beginning by evaluating the same data, preprocessing, model architecture, and evaluation metrics.

Performance Insights

FedEnsemble achieved the best diagnostic performance overall, with FedAtt and FedProx next, which indicates that aggregation strategies incorporating ensemble and attention mechanisms enhance robustness and generalization. In comparison, FedAvg was competitive under balanced settings, and performed significantly worse under heterogeneous data distributions. FedSGD was communication-round efficient but also less stable in prediction with respect the communication rounds.

Communication and System-Level Trade-offs

Second the main contribution of this work is the formalization of the communication overhead and scalability. While FedEnsemble achieved a higher accuracy, the communication cost and aggregation complexity are also higher than those of FedOpt. On the other hand, while FedSGD required less communication rounds, it was however more brittle against non-IID settings. Our findings highlight this performance vs. efficiency trade-off that is often under-represented in the literature on medical FL.

Non-IID Robustness

Under non-IID data distributions, which simulates realistic clinical heterogeneity, all algorithms reduced in performance. On the other hand, FedEnsemble and FedAtt were more robust than FedAvg indicating that adaptive mechanism can absorb cross-institutional variability.

Scalability

FedEnsemble exhibited consistent performance over increasing numbers of clients (10–50 clients), suggesting feasibility for large-scale multi-institutional deployments. The trends in improvement over time for simpler aggregation methods were relatively slower.

Novel Contribution

The main contribution of this work is the creation of a reproducible, holistic benchmarking framework that evaluates diagnostic accuracy, communication overhead, non-IID robustness, and scalability under controlled settings, all at once. This fills a major void in the current literature on federated learning in the context of lung cancer CT imaging.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Limitations

Our experiment is also restricted to a public dataset and the ConvNet backbone architecture is fixed No external validation was carried out on independent institutional datasets. Furthermore, they have qualitatively evaluated privacy guarantees instead of formal differential privacy mechanisms. To increase the clinical utility of these models, future work should include robust comparisons over a wider variety of datasets and model architectures. Second, surrogate labels (lung opacity instead of lung cancer) are used rather than biopsy-confirmed cancer diagnoses, which is an important limitation and may influence the generalizability and clinical relevance of the results presented. In future work, our proposed benchmarking framework should be validated on pathologically confirmed lung cancer datasets. This study lacked external validation in independent institutional datasets and thus its reported effects may not be directly generalizable to other clinical scenarios. Proposed benchmarking is validated against independent, multi-institutional datasets with diverse imaging protocols and patient demographics in future work.

6. CONCLUSION

We conducted a survey of perspectives on a variety of works in federated learning for lung cancer detection from a high-level point of view. We show in our analysis that FL not only safeguards the privacy of i.e. data at its origin, which resides locally (when data is stored or used by the same DAOs), but that it also effectively increases the potential of collaboration between healthcare organizations. We review the different ways of retrieval of output from FL Model which enables FL to collect heterogenous data from various platform like hospital, clinic, wearables and so on. In addition, federated algorithms such as FedEnsemble and FedAtt have also been investigated which found to be not only more accurate, precise, recall and F1 score better, but also suffered from high computation time and training cost leading to uncompetitive performance. Although FL is a trusted platform with regard to diagnostics in early lung cancer, it is not spared from the vices of the technology. Such challenges include heterogeneous data, communication cost, model convergence and resource-limited cases. To address these challenges, this review identifies and discusses several approaches; such as: improved aggregation algorithms, low communication protocols and edge computing advancements. These approaches aim to completely remove the concept of performance and efficiency bottlenecks from the FL models but ensure scalability of the model and feasibility in real application in a near-realistic healthcare context. It is followed by an incomplete compilation of future research and development directions needed to fulfill the promise of FL in lung cancer detection and treatment. In the case of non-IID data, we require better aggregation algorithms to improve model convergence. In addition, integrating the FL with novel technology like IoMT, blockchain, and real-time federated learning frameworks can also improve the generalization capability of lung cancer detection systems including security and scalability. Indeed, larger scale or multi-site case studies and naturally occurring mechanisms investigating how practitioners have begun to introduce FL into the real established clinical setting would certainly benefit FL research; so too would FL research for precision medication and the modelling necessary that will ultimately become personalisation to the individual and his/her disease history affording a more accurate diagnosis for subsequent successful treatment.

Implications for policy and practice Although there are many findings highlighting the harms and benefits of FL in healthcare, there are also several important implications for policy and practice that must be considered. Policymakers will want to draft policies and regulations to help and support the safe enactment of FL - they'll want to ensure that it is done in an ethical manner, doesn't breach data protection laws, such as HIPAA and GDPR - and helps create best practices for how to apply the technology in secure ways. Drawbacks: for such novel diagnostic tools to reach their full potential it is necessary for education and establishment of training of healthcare professionals on FL technologies and integration into an existing clinical workflow. Federated learning is a revolutionary method for early detection of lung cancer and addresses an important concern in healthcare regarding data privacy, security and integration in the reading room of healthcare. FL can help improve the accuracy and robustness of the diagnostic models while ensuring the private nature of data and enables continual updating of the model and, thus, allows health-care collaboration by providing decentralized model training. These findings highlight the enormous potential of FL in altering the line of inquiry for lung cancer assessment with immense benefits for patient outcome and global burden of lung cancer. The road towards using FL for detection of lung cancer is still a young one, but the work we have done to date appears to be rather complicated. As these FL practices can be further cultivated and established among the scientific community, including among clinicians, the field of medicine will benefit from the creation of more robust

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

diagnostic algorithms with improved privacy properties. The suggestions and implications from this review are aimed at faring a direction to future research procedures and policymaking in order for FL to remain its role as a key technology in the prevention of lung cancer.

Answers to the Research Questions: FL has an opportunity to enhance the performance of imaging-based triage at a lower architecture cost: it should ensure data locality, however, heterogeneity penalties and bandwidth/compute budgets must be taken into account; the selection of an algorithm should reflect the clinical priorities (sensitivity vs. speed). Limitations. Generalizability is limited by the use of surrogate labels (not biopsy-confirmed cancer), the use of individual backbones, and a absence of external validation as well as qualitative scoring of privacy. Future directions. (i) Test on cancer-specific, pathologically confirmed cohorts; (ii) embrace formal differential privacy or secure cohort aggregation with provable guarantees; (iii) assess domain adaptation / personalized FL to immense heterogeneity; (iv) extend to multimodal FL (CT + EHR/radiomics) and real hospital networks; (v) report calibration and decision-curve analysis reflecting clinical utility.

REFERENCES:

[1] S. A. A. Kharis, A. H. A. Zili, F. I. Fajar, A. Putri, and M. Arisanty, “Literature Review to Evaluate the Impact of Machine Learning and Artificial Intelligence for Lung Cancer Patient in COVID-19 Pandemic,” G-Tech: Jurnal Teknologi Terapan, vol. 8, no. 1, 2024, doi: 10.33379/gtech.v8i1.3891.

[2] H. T. Gayap and M. A. Akhloufi, “Deep Machine Learning for Medical Diagnosis, Application to Lung Cancer Detection: A Review,” BioMedInformatics, vol. 4, no. 1. 2024. doi: 10.3390/biomedinformatics4010015.

[3] Mohammad Shafiquzzaman Bhuiyan et al., “Advancements in Early Detection of Lung Cancer in Public Health: A Comprehensive Study Utilizing Machine Learning Algorithms and Predictive Models,” Journal of Computer Science and Technology Studies, vol. 6, no. 1, 2024, doi: 10.32996/jcsts.2024.6.1.12.

[4] H. Zhang and G. Song, “Solitary Pulmonary Nodules Prediction for Lung Cancer Patients Using Nomogram and Machine Learning,” Computer-Aided Design and Applications, vol. 21, no. S9, 2024, doi: 10.14733/cadaps.2024.S9.38-48.

[5] X. Shang, C. Zhang, R. Kong, C. Zhao, and H. Wang, “Construction of a Diagnostic Model for Small Cell Lung Cancer Combining Metabolomics and Integrated Machine Learning,” Oncologist, vol. 29, no. 3, 2024, doi: 10.1093/oncolo/oyad261.

[6] A. K. Gottipalla and P. Yalla, “Efficacy of Machine Learning Models in Lung Cancer Detection: An Emphasis on Bees with ICA Hybrid Feature Extraction,” International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 16s, 2024.

[7] Z. Chen, Y. Liu, Z. Lin, and W. Huang, “Understand how machine learning impact lung cancer research from 2010 to 2021: A bibliometric analysis,” Open Medicine (Poland), vol. 19, no. 1, 2024, doi: 10.1515/med-2023-0874.

[8] B. M. R. International, “Retracted: Lung Cancer Classification and Prediction Using Machine Learning and Image Processing,” BioMed research international, vol. 2024. 2024. doi: 10.1155/2024/9851527.

[9] S. D. -, P. M. -, M. S. -, R. K. -, and U. N. T. -, “Diagnosis and Prognosis of Lung Cancer & Lung Nodule using Machine Learning Techniques,” International Journal For Multidisciplinary Research, vol. 6, no. 1, 2024, doi: 10.36948/ijfmr.2024.v06i01.12760.

[10] J. N’Dow et al., “917P OPTIMA: Improve care for patients with prostate, breast, and lung cancer through artificial intelligence,” Annals of Oncology, vol. 33, 2022, doi: 10.1016/j.annonc.2022.07.1042.

[11] A. Dekker et al., “PD-0496: Multi-centric learning with a federated IT infrastructure: application to 2-year lung-cancer survival prediction,” Radiotherapy and Oncology, vol. 106, 2013, doi: 10.1016/s0167-8140(15)32802-4.

[12] S. Lu, B. Feng, K. Xu, Y. Chen, and X. Chen, “Personalized federated medical image classification with adaptive transfer robust features,” Journal of Image and Graphics, vol. 29, no. 3, 2024, doi: 10.11834/jig.230160.

[13] M. Field et al., “Federated Learning Survival Model and Potential Radiotherapy Decision Support Impact Assessment for Non–small Cell Lung Cancer Using Real-World Data,” Clinical Oncology, 2024, doi: 10.1016/j.clon.2024.03.008.

[14] Z. Zhang, Z. Wang, L. Zhao, A. Dekker, A. Traverso, and L. Wee, “PO-2115 federated learning of decentralized AI for predicting

radiation pneumonitis in lung cancer patients," Radiotherapy and Oncology, vol. 182, 2023, doi: 10.1016/s0167-8140(23)67030-6.

[15] H. Malik and T. Anees, "Federated learning with deep convolutional neural networks for the detection of multiple chest diseases using chest x-rays," Multimedia Tools and Applications, 2024, doi: 10.1007/s11042-023-18065-z.

[16] S. Adhikary, S. Dutta, and A. D. Dwivedi, "Secret learning for lung cancer diagnosis—a study with homomorphic encryption, texture analysis and deep learning," Biomedical Physics and Engineering Express, vol. 10, no. 1, 2024, doi: 10.1088/2057-1976/ad0b4b.

[17] A. S. Vibith and M. C. Jobin Christ, "GBDTMO: as new option for early-stage breast cancer detection and classification using machine learning," Automatika, vol. 64, no. 4, 2023, doi: 10.1080/00051144.2023.2226946.

[18] V. T. T. Vo, T. ho Shin, H. J. Yang, S. R. Kang, and S. H. Kim, "A comparison between centralized and asynchronous federated learning approaches for survival outcome prediction using clinical and PET data from non-small cell lung cancer patients," Computer Methods and Programs in Biomedicine, vol. 248, 2024, doi: 10.1016/j.cmpb.2024.108104.

[19] A. Rehman et al., "FedCSCD-GAN: A secure and collaborative framework for clinical cancer diagnosis via optimized federated learning and GAN," Biomedical Signal Processing and Control, vol. 89, 2024, doi: 10.1016/j.bspc.2023.105893.

[20] M. Subramanian, V. Rajasekar, V. E. Sathishkumar, K. Shanmugavadivel, and P. S. Nandhini, "Effectiveness of Decentralized Federated Learning Algorithms in Healthcare: A Case Study on Cancer Classification," Electronics (Switzerland), vol. 11, no. 24, 2022, doi: 10.3390/electronics11244117.

[21] R. Lemos et al., "A prospective observational study for a Federated Artificial Intelligence solution for moniToring mental Health status after cancer treatment (FAITH): study protocol," BMC Psychiatry, vol. 22, no. 1, 2022, doi: 10.1186/s12888-022-04446-5.

[22] "Modeling Asynchronous Federated Learning for Health care application," NeuroQuantology, vol. 20, no. 9, 2022, doi: 10.48047/nq.2022.20.9.nq44248.

[23] M. Field et al., "Infrastructure platform for privacy-preserving distributed machine learning development of computer-assisted theragnostics in cancer," Journal of Biomedical Informatics, vol. 134, 2022, doi: 10.1016/j.jbi.2022.104181.

[24] L. Liu, K. Fan, and M. Yang, "Federated learning: a deep learning model based on resnet18 dual path for lung nodule detection," Multimedia Tools and Applications, vol. 82, no. 11, 2023, doi: 10.1007/s11042-022-14107-0.

[25] M. J. Horry et al., "Development of Debiasing Technique for Lung Nodule Chest X-ray Datasets to Generalize Deep Learning Models," Sensors, vol. 23, no. 14, 2023, doi: 10.3390/s23146585.

[26] H. Malik, A. Naeem, R. A. Naqvi, and W. K. Loh, "DMFL_Net: A Federated Learning-Based Framework for the Classification of COVID-19 from Multiple Chest Diseases Using X-rays," Sensors, vol. 23, no. 2, 2023, doi: 10.3390/s23020743.

[27] S. Abbas et al., "Fused Weighted Federated Deep Extreme Machine Learning Based on Intelligent Lung Cancer Disease Prediction Model for Healthcare 5.0," International Journal of Intelligent Systems, vol. 2023, 2023, doi: 10.1155/2023/2599161.

[28] A. Heidari, D. Javaheri, S. Toumaj, N. J. Navimipour, M. Rezaei, and M. Unal, "A new lung cancer detection method based on the chest CT images using Federated Learning and blockchain systems," Artificial Intelligence in Medicine, vol. 141, 2023, doi: 10.1016/j.artmed.2023.102572.

[29] J. Zhou and H. Xin, "Emerging artificial intelligence methods for fighting lung cancer: A survey," Clinical eHealth, vol. 5, 2022, doi: 10.1016/j.ceh.2022.04.001.

[30] U. Subashchandrabose, R. John, U. V. Anbazhagu, V. K. Venkatesan, and M. Thyluru Ramakrishna, "Ensemble Federated Learning Approach for Diagnostics of Multi-Order Lung Cancer," Diagnostics, vol. 13, no. 19, 2023, doi: 10.3390/diagnostics13193053.

[31] Behuria, S., Swain, S., Bandyopadhyay, A., Turjya, S.M. and Gourisaria, M.K., 2025. Federated Learning approach in Healthcare ecosystems for efficient Lung Cancer classification: Insights from Model Generic Training to Fine-Tuning and Transfer Learning. Procedia Computer Science, 259, pp.279-290.

[32] Adnan, K.M., Ghazal, T.M., Saleem, M., Farooq, M.S., Yeun, C.Y., Ahmad, M. and Lee, S.W., 2025. Secure and interpretable lung cancer prediction model using mapreduce private blockchain federated learning and XAI. Scientific Reports, 15(1), p.35693.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[33] Liu, Y., Duan, X., Chen, X., Li, K., Li, Q., Liu, K., Long, W., Lin, H., Feng, B. and Chen, X., 2025. Collaborative assessment of the risk of postoperative progression in early-stage non-small cell lung cancer: a robust federated learning model. Cancer Imaging, 25(1), p.92.

[34] Chen, Y., Liu, L., Feng, B., Chen, Y., Xu, J., Lin, H., Li, K., Chen, X., Ke, Y., Zhou, H. and Hu, Q., 2025. A meta-learning-based robust federated learning for diagnosing lung adenocarcinoma and tuberculosis granulomas. Frontiers in Oncology, 15, p.1666937.

[35] Ankolekar, A., Boie, S., Abdollahyan, M., Gadaleta, E., Hasheminasab, S.A., Yang, G., Beauville, C., Dikaios, N., Kastis, G.A., Bussmann, M. and Chelala, C., 2025. Advancing breast, lung and prostate cancer research with federated learning. A systematic review. npj Digital Medicine, 8(1), p.314.

[36] Huang, Z., Feng, B., Chen, Y., Huang, L., Chen, Y., Duan, X., Chen, X., Lin, H., Li, K., Li, Y. and Li, Q., 2025. Risk stratification for early-stage NSCLC progression: a federated learning framework with large-small model synergy. Frontiers in Oncology, 15, p.1719433.

[37] Durga, S., Daniel, E., Seetha, S., Reshma, V.K. and Sachnev, V., 2025. FLEM-XAI: Federated learning based real time ensemble model with explainable AI framework for an efficient diagnosis of lung diseases. Frontiers in Computer Science, 7, p.1633916.