

Abstract

Cancer diagnosis through medical imaging is critical for early detection and treatment planning, yet current AI approaches face significant challenges, including data silos due to privacy regulations (HIPAA, GDPR), black box models lacking interpretability, limited datasets at individual institutions, and paramount privacy concerns in healthcare AI. This research addresses the critical gap where centralized AI requires data sharing, leading to privacy violations; local models have limited generalization, resulting in poor performances, and existing federated learning lacks explainability, leading to low clinical adoption.

We propose a novel federated learning framework that enables collaborative AI model training across multiple hospitals while preserving patient privacy and providing explainable diagnostic insights. Our methodology integrates federated learning with explainable AI (XAI) techniques for lung and colon cancer diagnosis using histopathology medical imaging data sourced from Kaggle repositories. The dataset's included 25,000 histopathological images with 5 classes (Lung benign tissue, Lung adenocarcinoma, Lung squamous cell carcinoma, Colon adenocarcinoma and Colon benign tissue) . All the images are 768 x 768 pixels in size and were in jpeg file format. The images were generated from an original sample of HIPAA compliant and validated sources, consisting of 750 total images of lung tissue (250 benign lung tissue, 250 lung adenocarcinomas, and 250 lung squamous cell carcinomas) and 500 total images of colon tissue (250 benign colon tissue and 250 colon adenocarcinomas) and augmented to 25,000 using the Augmentor package. The framework used incorporates differential privacy mechanisms to ensure robust data protection while maintaining model performance.

Our experimental results demonstrate exceptional performance with 99% accuracy for lung and colon cancer diagnosis while maintaining strong privacy preservation through differential privacy ($\epsilon = 1.0$). Comparative analysis shows superior performance over traditional approaches, with CNN achieving 92.6% accuracy and MobileNetV2 achieving 91.1% accuracy. The system successfully provides explainable AI insights trusted by clinical practitioners and enables multi-institutional collaboration without compromising patient data confidentiality. This research demonstrates significant clinical impact potential, as early cancer detection with AI assistance can improve survival rates by 20-30% while maintaining stringent privacy standards essential for healthcare applications.