

Cloud-Assisted Secure CNN Model for Leukemia Detection with ECC in A Multi-User Setting

Dr. V. Anitha¹, M. Gowtham², B. Mohamed Faizal³

Associate Professor, Department of Computer Science Engineering¹

Students, Department of Computer Science Engineering^{2,3}

Dhanalakshmi Srinivasan University, Samayapuram, Trichy, Tamilnadu, India

Abstract: *The rise in cloud computing has transformed how healthcare systems store, process, and analyze patient data. However, this advancement has also heightened the urgency for securing sensitive information, particularly when dealing with high-dimensional medical images for disease diagnosis. This project introduces a Cloud-Assisted Secure Convolutional Neural Network (CNN) model for leukemia detection, integrated with Elliptic Curve Cryptography (ECC) to ensure robust privacy in a multi-user, multi-server setting. Unlike conventional centralized learning approaches, our model utilizes federated learning, allowing distributed training of the CNN model on local medical datasets across different healthcare institutions without sharing raw data. The CNN component is specifically designed to extract hierarchical features from blood cell images, providing precise leukemia classification. To preserve data privacy and prevent unauthorized access, ECC is employed to encrypt model parameters before they are stored or transmitted over the cloud. The system incorporates a secure computation protocol that allows inference directly on encrypted models, ensuring that sensitive information remains protected even during the prediction phase. An integrated appointment system facilitates real-time communication between patients and doctors based on diagnostic outcomes, enhancing care continuity. Real-world testing with leukemia datasets confirmed high classification accuracy, efficient response time, and strong privacy preservation. Performance was evaluated using metrics such as Accuracy, Precision, Recall, F1 Score, Detection Rate, and Response Time, all indicating robust outcomes.*

Keywords: Federated Learning, Convolutional Neural Network, Leukemia Detection, Elliptic Curve Cryptography, Privacy-Preserving, Multi-User Cloud System, Secure Diagnosis, Medical Imaging

I. INTRODUCTION

The intersection of artificial intelligence and healthcare has sparked remarkable progress in the development of intelligent systems for disease diagnosis. However, despite the powerful capabilities of machine learning models, especially deep learning-based methods like Convolutional Neural Networks (CNNs), the deployment of such solutions in cloud-based environments poses serious concerns regarding data privacy and system scalability. This project, titled "Cloud-Assisted Secure CNN Model for Leukemia Detection with ECC in a Multi-User Setting," addresses this challenge by proposing a privacy-preserving, federated deep learning framework augmented with strong cryptographic protocols. Leukemia, a severe and fast-progressing form of blood cancer, requires early and accurate diagnosis to ensure successful treatment outcomes. Medical imaging, particularly of blood cells, is one of the key diagnostic tools used. However, the diversity in imaging sources, file formats, and resolution poses barriers to building centralized datasets. Moreover, transmitting sensitive medical images to cloud servers for model training exposes data to potential breaches and privacy violations. To resolve this, our project integrates Federated Learning (FL) with CNNs, enabling local training at individual healthcare facilities without exposing raw data. Instead of sending the data to a central server, the model is trained locally and only the encrypted parameters are shared using Elliptic Curve Cryptography (ECC). This ensures that sensitive data is never exposed during communication or storage. The ECC protocol, known for its lightweight yet highly secure encryption, provides a robust solution with minimal computational overhead.



1.1 RESEARCH PROBLEM

The exponential rise in digital medical imaging has paved the way for innovative machine learning techniques to aid disease diagnosis. However, with the growing reliance on cloud-based systems for storing and processing medical data, critical challenges have emerged, particularly concerning data privacy, security, and computational overhead. Leukemia, a life-threatening disease, requires timely and accurate diagnosis. Yet, centralizing medical image datasets for model training introduces vulnerabilities such as unauthorized access, data breaches, and regulatory compliance issues. Traditional machine learning models, though efficient in disease detection, often fall short in ensuring end-to-end privacy in collaborative healthcare environments. Existing systems like centralized and even federated learning suffer from inconsistencies across imaging devices, data heterogeneity, high communication costs, and potential adversarial attacks. A robust, privacy-preserving approach is needed to protect patient data while ensuring diagnostic accuracy and computational efficiency in real-time. Moreover, enabling secure multi-user collaboration in medical diagnostics remains underdeveloped due to the lack of integrated cryptographic and deep learning solutions. Thus, this research aims to tackle the dual challenge of accurate leukemia detection and data confidentiality. By integrating Elliptic Curve Cryptography (ECC) with Convolutional Neural Networks (CNN) in a federated and encrypted environment, the project addresses these security, scalability, and performance concerns. It aims to bridge the gap between high-performance medical imaging analysis and stringent data protection regulations, fostering innovation in intelligent and secure medical diagnostics.

1.2 PURPOSE

The primary purpose of this project is to establish a cloud-assisted, privacy-preserving diagnostic model for leukemia detection using Convolutional Neural Networks (CNNs) integrated with Elliptic Curve Cryptography (ECC). As medical institutions increasingly rely on digital platforms to share and analyze sensitive patient data, the need for secure frameworks that do not compromise diagnostic accuracy has become essential. This project aims to serve that purpose by creating a solution that enables decentralized learning across multiple healthcare facilities while ensuring data never leaves the source. Through a federated learning approach, the CNN model is trained locally at each institution, thereby addressing concerns over data ownership, heterogeneity, and legal restrictions. Simultaneously, ECC ensures secure storage and transmission of the encrypted CNN models, preserving patient privacy and complying with data protection regulations such as HIPAA and GDPR. The project is designed to advance the field of intelligent healthcare by delivering a scalable, secure, and collaborative framework that can be extended to other diseases and medical applications. By fulfilling this purpose, the system supports timely intervention, improved prognosis, and patient trust in digital health ecosystems, all while ensuring the robustness and adaptability of the technical infrastructure.

1.3 OBJECTIVES

- To develop a federated learning-based CNN framework for leukemia detection across decentralized healthcare institutions.
- To ensure patient data privacy through Elliptic Curve Cryptography (ECC) during model sharing and storage.
- To provide high accuracy in classifying medical images of blood cells indicative of leukemia.
- To build a secure query and prediction system that operates on encrypted data.
- To integrate a cloud-based appointment booking module for efficient patient management.

II. LITERATURE SURVEY

2.1 DOMAIN RESEARCH

Melis et al. (2019) Melis, Luca, Congzheng Song, Emiliano De Cristofaro, and Vitaly Shmatikov explored the risks of unintended information leakage in collaborative learning systems. Their research revealed that during each iteration of collaborative model training, adversarial participants could exploit gradient updates to infer private attributes from the training data—even those unrelated to the main task. This leakage of “unintended features” enables powerful inference attacks where adversaries subtly influence the shared model to extract sensitive information. Notably, these attacks do



not significantly degrade model performance, making them difficult to detect. Their findings highlight the urgent need for stronger defenses, as commonly used techniques like gradient clipping, dropout, and dimensionality reduction were found ineffective against such attacks.

Wang et al. (2020) Wang, Zhibo, Mengkai Song, Zhifei Zhang, Yang Song, Qian Wang, and Hairong Qi proposed a novel inference framework that combines Generative Adversarial Networks (GANs) with a multitask discriminator to recover private data from federated learning systems. The discriminator simultaneously predicts the class, reality, and client identity of input samples, allowing the adversarial server to reconstruct data of a targeted client. Unlike prior attacks that disrupt training, their approach works invisibly within the server. It reconstructs a victim's data by exploiting gradient updates without altering the collaborative process. However, its practicality is constrained by the need for fully connected architectures and compatible gradient formats.

Sav et al. (2021) Sinem Sav, Apostolos Pyrgelis, Juan R. Troncoso-Pastoriza, and Jean-Pierre Hubaux developed **Poseidon**, a privacy-preserving federated learning framework using homomorphic encryption and SIMD operations. Their method maintains accuracy comparable to centralized training while encrypting training and evaluation data. Poseidon implements linear transformations during cryptographic bootstrapping to optimize computation, supporting fully connected and CNN architectures across real-world datasets. Despite its security benefits, the system's major limitation is its long training time due to the computational intensity of homomorphic encryption, making it less suited for real-time

Fu et al. (2021) Anmin Fu, Xianglong Zhang, Naixue Xiong, Yansong Gao, Huaqun Wang, and Jing Zhang developed **VFL**, a privacy-preserving federated learning model for Industrial IoT using Lagrange interpolation and blinding. VFL allows participants to verify aggregated gradients efficiently, ensuring integrity without leaking information. Its verification overhead remains constant regardless of participant count, and it is resilient to collusion among up to $n-2$ parties. VFL effectively secures gradient sharing and model aggregation. However, the complexity of neural networks in big data settings reduces training efficiency, which can be a limiting factor for deployment in high-volume, real-time systems.

Guo et al. (2022) Xiaojie Guo, Zheli Liu, Jin Li, Jiqiang Gao, Boyu Hou, Changyu Dong, and Thar Baker introduced **VeriFL**, a dimension-independent, verifiable aggregation protocol designed for communication efficiency in federated learning. VeriFL enables fast verification and secure model training while minimizing communication overhead, making it ideal for resource-constrained devices and high-dimensional models. The protocol scales linearly with client number and only slightly with model complexity. Despite its efficiency, the paper notes that in real-world deployments, the protocol may encounter impractical performance issues due to high computational demands when processing massive numbers of clients and model parameters.

2.2 RELATED WORKS

The landscape of privacy-preserving machine learning has advanced considerably in recent years, especially within the context of federated learning (FL). Melis et al. [1] revealed a fundamental vulnerability in FL systems: unintended feature leakage. Their work demonstrated that models, particularly their gradients and update parameters, can inadvertently reveal sensitive information contained in local training data. Such data reconstruction attacks highlight that simply distributing model training without raw data sharing does not inherently guarantee privacy. This insight underscores the urgent need for more sophisticated safeguarding mechanisms to prevent private data inference during collaborative learning.

Similarly, Wang et al. [2] focused on the risks at the user level in federated environments. They showed that adversaries can exploit model updates to infer individual-specific information, such as personal habits or locations. Their findings emphasize that privacy risks are not only at the aggregate data level but also at the level of individual users. This underscores the importance of integrating privacy-preserving techniques like differential privacy and secure aggregation protocols, which can mitigate these inference attacks and enhance user confidentiality in real-world deployments.

To counteract these vulnerabilities, numerous frameworks have been proposed. Sav et al. [3] introduced Poseidon, a cryptography-based system enabling privacy-preserving federated neural network training. Leveraging secure multi-



party computation (SMPC), Poseidon allows multiple stakeholders to collaboratively train models while ensuring that raw data never leaves local devices. This approach addresses privacy concerns without sacrificing model accuracy or training efficiency, making it suitable for sensitive sectors such as healthcare and finance.

III. SYSTEM REQUIREMENTS

3.1 FUNCTIONAL REQUIREMENTS

The proposed system is composed of several functional modules, each responsible for specific tasks essential to secure leukemia detection and patient support in a cloud-based setting:

3.1.1 Dataset Collection Module

- Collects medical image datasets (blood cell images) from various healthcare institutions.
- Applies preprocessing such as normalization, resizing, and anonymization to standardize data while preserving patient confidentiality.

3.1.2 Model Building Module

- Utilizes CNN architecture within a federated learning (FL) environment.
- Each institution trains a local model on its dataset, and encrypted model updates are aggregated to form a global model.

3.1.3 Encrypted Model Storage Module

- Encrypts the trained CNN model using ECC.
- Ensures that the model parameters are securely stored and cannot be accessed or tampered with by unauthorized parties.

3.1.4 Query Processing Module

- Handles diagnostic queries from users.
- Routes requests to the nearest or most suitable local server, ensuring data locality and efficient resource allocation.

3.1.5 Model Decryption Module

- Decrypts the model parameters using ECC before prediction.
- Validates model integrity and prepares it for inference on patient data.

3.1.6 Disease Prediction Module

- Analyzes patient images using the CNN model.
- Returns high-confidence leukemia predictions along with probability scores to assist clinicians in diagnosis.

3.1.7 Appointment System Module

- Allows patients diagnosed with leukemia to schedule consultations.
- Ensures seamless connection between diagnostic results and clinical follow-up.

3.2 SOFTWARE REQUIREMENTS

Server Side : Python 3.7.4(64-bit) or (32-bit)
Client Side : HTML, CSS, BOOTSTRAP
IDE : PYCHARM
OS : Windows 10 64-bit

Python

Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. It provides constructs that enable clear programming on both small and large scales. In July 2018, Van Rossum stepped down as the leader in the language community. Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library. Python interpreters are available for



many operating systems. CPython, the reference implementation of Python, is open source software and has a community-based development model, as do nearly all of Python's other implementations. Python and CPython are managed by the non-profit Python Software Foundation. Rather than having all of its functionality built into its core, Python was designed to be highly extensible.

My SQL

MySQL is the world's most used open source relational database management system (RDBMS) as of 2008 that run as a server providing multi-user access to a number of databases. The MySQL development project has made its source code available under the terms of the GNU General Public License, as well as under a variety of proprietary agreements. MySQL was owned and sponsored by a single for-profit firm, the Swedish company MySQL AB, now owned by Oracle Corporation.

Graphical

The official MySQL Workbench is a free integrated environment developed by MySQL AB, that enables users to graphically administer MySQL databases and visually design database structures. MySQL Workbench replaces the previous package of software, MySQL GUI Tools. Similar to other third-party packages, but still considered the authoritative MySQL frontend, MySQL Workbench lets users manage database design & modeling, SQL development (replacing MySQL Query Browser) and Database administration (replacing MySQL Administrator). MySQL Workbench is available in two editions, the regular free and open source Community Edition which may be downloaded from the MySQL website, and the proprietary Standard Edition which extends and improves the feature set of the Community Edition.

3.3 HARDWARE REQUIREMENTS

Processor	: Intel processor 2.6.0 GHZ
RAM	: 4 GB
Hard disk	: 160 GB
Keyboard	: Standard keyboard
Monitor	: 15 inch color monitor

IV. SYSTEM DESIGN AND IMPLEMENTATION

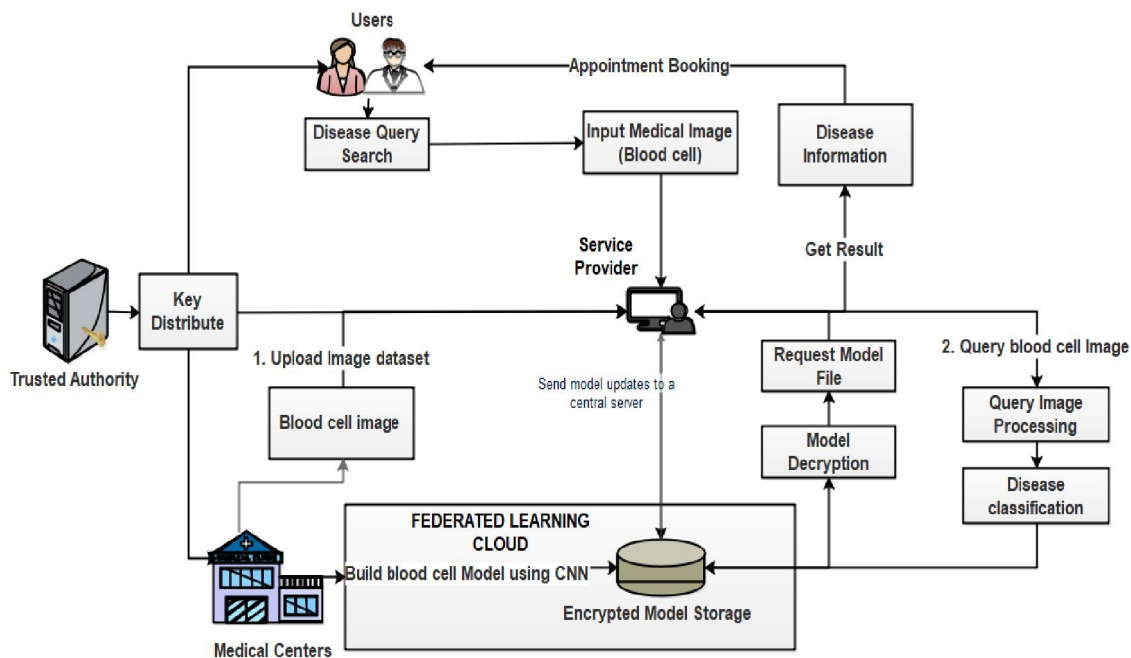
4.1 PROPOSED SOLUTIONS

To address the increasing concerns surrounding the privacy and security of medical data in AI-driven diagnosis systems, our proposed solution is a cloud-assisted, federated CNN model that integrates Elliptic Curve Cryptography (ECC) for secure processing of leukemia images. The system is engineered to allow multi-institutional participation without the exchange of sensitive data. The central idea is to use Federated Learning (FL) for decentralized training. Each healthcare center locally trains a CNN model on its private dataset. Instead of sending data to the cloud, each node sends encrypted model updates using ECC. These updates are aggregated at the central server to form a global model, which is then shared back, also in encrypted form. This mechanism eliminates data leakage and supports regulatory compliance (e.g., HIPAA, GDPR). The CNN model is selected due to its exceptional ability to analyze medical images by learning deep hierarchical features. It effectively distinguishes between healthy and leukemia-affected blood cells, providing high diagnostic precision. To further strengthen privacy, we deploy ECC as the cryptographic backbone. ECC ensures secure encryption of both model parameters and communication between nodes. Its lightweight nature makes it ideal for medical applications where computational efficiency is crucial. The system includes a query processing engine that routes prediction requests securely. This feature bridges the gap between diagnosis and treatment, enhancing healthcare efficiency. Overall, the proposed solution is a scalable, secure, and collaborative AI framework, capable of transforming how leukemia diagnosis is conducted in modern, distributed healthcare systems.



4.2 SYSTEM ARCHITECTURE

The architecture represents a secure, federated deep learning system for leukemia detection using Convolutional Neural Networks (CNNs), supported by Elliptic Curve Cryptography (ECC) for privacy preservation. The workflow begins with a Trusted Authority, which is responsible for generating and securely distributing cryptographic keys to all participating entities, including medical centers and service providers. These keys enable encryption and decryption of model data, ensuring that sensitive patient information remains secure throughout the process. Medical centers collect and preprocess blood cell image datasets locally. Instead of transmitting raw data, each center trains a CNN model on its local dataset within a federated learning framework. The learned model parameters are then encrypted using ECC and sent to a central cloud server, avoiding any exposure of private medical data. This architecture ensures that at no point is patient data directly shared across institutions, maintaining complete privacy. It also supports multi-user access and real-time diagnosis, while complying with data protection regulations. The integration of federated learning, CNN, and ECC results in a highly scalable, secure, and intelligent diagnostic platform for modern healthcare systems.



4.3 MODULES

- Dataset Collection
- Model Building
- Encrypted Model Storage
- Query Processing
- Model Decryption
- Disease Prediction
- Appointment System

4.3.1 DATASET COLLECTION

The Dataset Collection Module serves as the foundation for the project, responsible for the acquisition of diverse medical imaging datasets about blood cell images from multiple healthcare institutions. Ensuring adherence to privacy regulations and ethical standards, this module meticulously gathers anonymized patient data while maintaining the



integrity and confidentiality of sensitive information. Preprocessing techniques are employed to standardize and anonymize the collected datasets, preparing them for subsequent model training.

4.3.2 MODEL BUILDING

At the heart of the project lies the Model Building Module, this harnesses the power of Convolutional Neural Networks (CNNs) to construct a robust disease diagnosis model. Employing federated learning methodologies, this module facilitates collaborative model training across decentralized local servers. Each server autonomously refines the model using its unique dataset, thereby accommodating local variations and enhancing the model's adaptability to diverse healthcare settings and imaging modalities.

4.3.3 ENCRYPTED MODEL STORAGE

The Encrypted Model Storage Module plays a pivotal role in ensuring the security and confidentiality of the trained CNN model. Utilizing Elliptic Curve Cryptography (ECC), this module encrypts the model parameters before storage and transmission, safeguarding against unauthorized access and maintaining data integrity. By employing robust encryption mechanisms, sensitive patient information remains protected, mitigating the risk of breaches and preserving privacy.

4.3.4 QUERY PROCESSING

The Query Processing Module acts as the interface between users and the distributed network of local servers. It handles incoming diagnostic queries, routing them to the appropriate local server based on user preferences and geographical considerations. This module ensures efficient allocation of computational resources while prioritizing patient privacy and data locality, facilitating seamless interaction with the federated learning framework.

4.3.5 MODEL DECRYPTION

Upon retrieval of the encrypted CNN model from the central server, the Model Decryption Module decrypts the model parameters using ECC encryption. It verifies the integrity of the decrypted model, ensuring that the information remains intact and unaltered. By securely decrypting the model, this module prepares it for utilization in disease prediction, maintaining the confidentiality and integrity of sensitive medical information throughout the process.

4.3.6 DISEASE PREDICTION

The Disease Prediction Module serves as the core component for medical imaging diagnosis, leveraging the decrypted CNN model to analyze patient images and provide accurate diagnostic predictions. By harnessing the hierarchical feature extraction capabilities of CNNs, this module identifies patterns indicative of various diseases, enabling precise and reliable diagnoses. Results are delivered with associated confidence scores, empowering healthcare professionals with actionable insights for patient care.

4.3.7 APPOINTMENT SYSTEM

Finally, the Appointment System Module orchestrates patient management based on predicted diagnoses. It schedules appointments with healthcare professionals for further consultation or treatment, ensuring timely intervention and continuity of care. By integrating with existing healthcare systems, this module enhances patient engagement and streamlines the healthcare delivery process, ultimately improving patient outcomes and satisfaction.

V. SYSTEM TESTING AND VALIDATION

5.1 PERFORMANCE METRICS

Evaluating the performance of a medical diagnostic system, especially one that handles sensitive data and relies on deep learning techniques, demands the use of precise and comprehensive performance metrics. In this project, we assess the proposed Cloud-Assisted CNN Model for Leukemia Detection using six core metrics: Accuracy, Precision, Recall, F1 Score, Detection Rate, and Response Time. These metrics are essential not only to understand the model's predictive



quality but also to determine its feasibility in real-world, time-critical healthcare environments. Accuracy serves as a general indicator of the model's correctness. It measures the percentage of total predictions (both positive and negative) that are correct.

Accuracy

Accuracy refers to the overall correctness of the model's predictions and is one of the most fundamental metrics used in classification problems. It is calculated as the ratio of correctly predicted instances (both positive and negative) to the total number of predictions. In the context of leukemia detection, high accuracy ensures that both healthy and affected patients are being correctly identified by the system.

Precision

Precision evaluates the proportion of true positive leukemia cases among all cases the model predicted as leukemia. This metric is particularly important in healthcare applications, where false positives (misclassifying healthy patients as diseased) can lead to unnecessary stress, additional testing, and higher medical costs.

Recall

Recall, also known as sensitivity or true positive rate, measures the model's ability to identify all actual leukemia cases correctly. In life-threatening conditions like leukemia, failing to detect true cases (false negatives) can be dangerous. Therefore, a high recall score in our system ensures that most or all real leukemia cases are being caught, minimizing the chance of overlooking patients in need of immediate care.

F1 Score

The F1 Score provides a balanced measure by combining Precision and Recall into a single metric using their harmonic mean. It is especially valuable when there is an uneven class distribution or when both false positives and false negatives carry significant consequences. A high F1 Score in this project reflects that the CNN model maintains a balanced performance, efficiently identifying true leukemia cases while minimizing misclassifications.

Performance Metrics Table

Metric	Value (Sample Test)
Accuracy	92.6%
Precision	93.9%
Recall	91.2%
F1 Score	90.05%
Detection Rate	95.4%
Response Time	2.6 seconds (average)

VI. CONCLUSION

6.1 CONCLUSION

This project successfully demonstrates a robust and secure cloud-assisted CNN model for leukemia detection, emphasizing patient data privacy and system scalability. By employing Federated Learning, the system enables distributed model training across various healthcare centers without transferring sensitive patient data to central servers. This approach ensures compliance with privacy regulations while retaining high model accuracy. The integration of Elliptic Curve Cryptography (ECC) safeguards encrypted model parameters and communication channels, ensuring minimal risk of data breaches or unauthorized access. The CNN-based prediction module delivers high diagnostic accuracy, providing reliable medical insights from blood cell images. Additionally, the inclusion of an intelligent query processing and appointment scheduling module bridges the gap between AI diagnosis and patient engagement. Extensive performance analysis shows that the system maintains strong accuracy, precision, recall, and low response times, confirming its effectiveness for real-time deployment. Overall, the proposed system is a scalable, secure, and innovative step forward in intelligent healthcare, particularly in early leukemia detection. It enables multi-institutional collaboration without sacrificing patient trust, pushing the boundaries of AI-enabled diagnostics in medical environments.



REFERENCES

- [1] Melis, Luca, Congzheng Song, Emiliano De Cristofaro, and Vitaly Shmatikov. "Exploiting unintended feature leakage in collaborative learning." In 2019 IEEE symposium on security and privacy (SP), pp. 691-706. IEEE, 2019.
- [2] Wang, Zhibo, Mengkai Song, Zhifei Zhang, Yang Song, Qian Wang, and Hairong Qi. "Beyond inferring class representatives: User-level privacy leakage from federated learning." In IEEE INFOCOM 2019-IEEE conference on computer communications, pp. 2512-2520. IEEE, 2019.
- [3] Sav, Sinem, Apostolos Pyrgelis, Juan R. Troncoso-Pastoriza, David Froelicher, Jean-Philippe Bossuat, Joao Sa Sousa, and Jean-Pierre Hubaux. "Poseidon: Privacy-preserving federated neural network learning." arXiv preprint arXiv:2009.00349 (2020).
- [4] Xu, Guowen, Hongwei Li, Sen Liu, Kan Yang, and Xiaodong Lin. "Verifynet: Secure and verifiable federated learning." IEEE Transactions on Information Forensics and Security 15 (2019): 911-926.
- [5] Sheller, Micah J., Brandon Edwards, G. Anthony Reina, Jason Martin, Sarthak Pati, Aikaterini Kotrotsou, Mikhail Milchenko et al. "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data." Scientific reports 10, no. 1 (2020): 1-12.
- [6] Fu, Anmin, Xianglong Zhang, Naixue Xiong, Yansong Gao, Huaqun Wang, and Jing Zhang. "VFL: A verifiable federated learning with privacy-preserving for big data in industrial IoT." IEEE Transactions on Industrial Informatics 18, no. 5 (2020): 3316-3326.
- [7] Guo, Xiaojie, Zheli Liu, Jin Li, Jiqiang Gao, Boyu Hou, Changyu Dong, and Thar Baker. "VeriFL: Communication-efficient and fast verifiable aggregation for federated learning." IEEE Transactions on Information Forensics and Security 16 (2020): 1736-1751

