

## Article

# A Fog-Based Privacy-Preserving Federated Learning System for Smart Healthcare Applications

Maryum Butt <sup>1</sup>, Noshina Tariq <sup>1</sup> , Muhammad Ashraf <sup>1</sup>, Hatoon S. Alsagri <sup>2,\*</sup>, Syed Atif Moqurrab <sup>3</sup>, Haya Abdullah A. Alhakbani <sup>2</sup> and Yousef A. Alduraywish <sup>2</sup>

<sup>1</sup> Department of Avionics Engineering, Air University, Islamabad 44000, Pakistan; bmaryum95@gmail.com (M.B.); noshina.tariq@mail.au.edu.pk (N.T.); muhammad.ashraf@mail.au.edu.pk (M.A.)

<sup>2</sup> College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11673, Saudi Arabia; hahakbani@imamu.edu.sa (H.A.A.A.); yalduraywish@imamu.edu.sa (Y.A.A.)

<sup>3</sup> School of Computing, Gachon University, 1342, Seongnam-daero, Sujeong-gu, Seongnam-si 13120, Republic of Korea; atif@gachon.ac.kr

\* Correspondence: hsagri@imamu.edu.sa

**Abstract:** During the COVID-19 pandemic, the urgency of effective testing strategies had never been more apparent. The fusion of Artificial Intelligence (AI) and Machine Learning (ML) models, particularly within medical imaging (e.g., chest X-rays), holds promise in smart healthcare systems. Deep Learning (DL), a subset of AI, has exhibited prowess in enhancing classification accuracy, a crucial aspect in expediting COVID-19 diagnosis. However, the journey to harness DL's potential is rife with challenges: notably, the intricate landscape of medical data privacy. Striking a balance between utilizing patient data for insights while upholding privacy is formidable. Federated Learning (FL) emerges as a solution by enabling collaborative model training across decentralized data sources, thus bypassing data centralization and preserving data privacy. This study presents a tailored, collaborative FL architecture for COVID-19 screening via chest X-ray images. Designed to facilitate cooperation among medical institutions, the framework ensures patient data remain localized, eliminating the need for direct data sharing. Addressing imbalanced and non-identically distributed data, the architecture is a robust solution. Implementation entails localized and fog-computing-based FL models. Localized models utilize Convolutional Neural Networks (CNNs) on institution-specific datasets, while the FL model, refined iteratively, takes precedence in the final classification. Intriguingly, the global FL model, fortified by fog computing, emerges as the frontrunner in classification after weight refinement, surpassing local models. Validation within the COLAB platform gauges the model's performance through metrics such as accuracy, precision, recall, and F1-score. Remarkably, the proposed model excels across these metrics, solidifying its efficacy. This research navigates the confluence of AI, FL, and medical imaging, unveiling insights that could reshape healthcare delivery. The study enriches scientific discourse by addressing data privacy in collaborative learning and carries potential implications for enhanced patient care.

**Keywords:** smart healthcare; federated learning; COVID-19; convolutional neural networks; fog computing



**Citation:** Butt, M.; Tariq, N.; Ashraf, M.; Alsagri, H.S.; Moqurrab, S.A.; Alhakbani, H.A.A.; Alduraywish, Y.A. A Fog-Based Privacy-Preserving Federated Learning System for Smart Healthcare Applications. *Electronics* **2023**, *12*, 4074. <https://doi.org/10.3390/electronics12194074>

Academic Editor: Simeone Marino

Received: 25 August 2023

Revised: 17 September 2023

Accepted: 21 September 2023

Published: 28 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The bare minimum list of essential smart healthcare services includes emergency, inpatient, hospital, medical care, outpatient, and preventative medical services [1,2]. A data leak can harm both doctors' and their patients' reputations. Cyberattack risk is decreased by enhancing the security of IT systems that store and process medical records. To preserve medical information, laws establish security standards and place protections around patient data and healthcare facilities. Regarding data storage, "on-premise" refers to the healthcare provider's data center [3]. Medical professionals use Artificial Intelligence

(AI) (e.g., Machine Learning (ML) and Deep Learning (DL)) to assist with patient care and clinical data management [4]. DL is progressively integrated into advanced systems with notable clinical implications. Some of the most promising applications involve patient-facing innovations and surprisingly mature approaches to enhancing the user experience of health IT [5]. Among these, Convolutional Neural Networks (CNNs), a category of DL, demonstrate remarkable proficiency in analyzing images, such as X-rays or findings from MRI scans [6]. CNNs were intentionally designed by computer scientists at Stanford University to excel in image processing, with the goal of enabling them to operate more efficiently and manage larger images. Consequently, certain CNNs surpass human diagnosticians in accurately identifying crucial details within diagnostic imaging assessments [7]. Federated Learning (FL) enables non-affiliated hospitals to participate without sharing their rich datasets, alleviating the need to co-locate all the data in a central location and enabling independent hospitals to gain benefits from them [8]. Significant problems, such as data security, privacy, and access to diverse data, are solved by this method [9,10].

FL involves the collective sharing of data among multiple individuals to collaboratively train a singular DL model and progressively enhance its capabilities; it resembles a collaborative group presentation or report [11]. Each participant receives the model from a cloud data center, often starting with a foundational model that has previously undergone training. It is a good option for private information belonging to patients, persons, businesses, or other sectors that must adhere to strict privacy regulations [12]. FL offers an alternative to modeling problems by allowing numerous edge devices or organizations to train a global model using various local data. FL mainly or partially supports data privacy and security challenges [13,14]. However, there is a conflict when we discuss ML and privacy. Indeed, to perform well, ML in general—and DL models in particular—require access to massive datasets. Unfortunately, due to privacy issues and liability risks, these data are frequently maintained by many companies. Privacy, data ownership, legal, and technical issues make it challenging to acquire most data, especially in the healthcare industry.

Furthermore, in the healthcare industry, data management strategies are severely impacted by international regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union [15,16]. Data collection from clients in bulk without a clear service goal is no longer an issue. The GDPR provides the legal framework to protect personal data in the European Union. The GDPR increases corporate accountability by imposing additional requirements on service providers concerning data handling, specifically emphasizing controlling its centralization. Privacy protection has assumed paramount importance at the heart of data processing now more than ever [17]. These challenges pose a serious concern for organizations and data enthusiasts dealing in the healthcare industry to build AI healthcare applications. In other words, one must disclose data to utilize these full diagnostics.

A distributed ML approach called FL, which Google first unveiled in 2017 [18], enables cross-organization collaboration for training DL models without disclosing their datasets for cases for which training data are to be maintained locally by device users (nodes) as opposed to a centralized data center; FL emerges as a compelling ML approach [19]. These nodes use their data to compute, and then they update a global model. Since these nodes can be located in various geographical regions subjected to different regulations and with different data generation and preparation techniques, the data distribution patterns across these nodes can have a high level of variance. In a heterogeneous multi-node environment, some clients may have more or different data than others. Therefore, identifying a data sample that generalizes the entire data distribution is an unlikely event. Two primary features that distinguish federated optimization problems from their distributed counterparts are highlighted in this context. First, the data used by each user for training are non-identically and independently distributed (Non-IID), eliminating the existence of a consolidated representation of the entire population's dataset. Second, data distribution is unbalanced, with varying amounts of data held by each user compared to others. Leveraging the

benefits of Federated Learning (FL), we have employed FL-based techniques to address a compassionate issue in the healthcare sector: the coronavirus disease (COVID-19). This infectious ailment, caused by the SARS-CoV-2 virus, was initially reported in Wuhan, China, in December 2019 [20]. The outbreak spread rapidly throughout China and most of the world's nations. Stopping the virus is extremely difficult given the pandemic's rapid escalation (thousands of deaths and hundreds of thousands of infections). There are currently numerous diagnostic techniques available for the identification of coronavirus.

Chest X-ray imaging and CT scans, which are widely utilized [21,22], play a pivotal role. The lung damage caused by COVID-19, which targets respiratory tract epithelial cells, is assessed through chest X-ray images [23]. Furthermore, the availability of X-ray imaging equipment globally and its heightened diagnostic accuracy confers a notable advantage in COVID-19 diagnosis compared to specialized testing kits. AI-based models for processing chest X-ray images offer an expedient and cost-effective alternative for COVID-19 screening, prompting substantial interest in the scientific community to develop ML-based strategies for outbreak prediction and diagnosis [24]. This study's objective—the cornerstone of the article's distinctive concepts—is to formulate and assess an FL-based system for identifying COVID-19 through chest X-ray images. A pioneering effort, this research employs FL techniques to detect COVID-19 in chest X-rays proficiently. The principal contributions of this research encompass:

- This research proposes a collaborative and decentralized system specifically designed to empower clinicians with valuable insights for diagnosing COVID-19 while ensuring stringent safeguards are in place for patient data privacy.
- The study introduces a robust, decentralized fog-based Federated Learning model. This model is unique in its ability to effectively handle dispersed and unevenly distributed data, yielding results that are on par with those obtained from traditional centralized ML approaches.
- To ensure clarity and facilitate future research, the mathematical equations and algorithms used in this study are provided and explained in detail. These resources serve as a valuable tool for readers who wish to further understand or build upon this work.
- A comprehensive analysis of the model's performance has been conducted, demonstrating high accuracy and precision. Further examination of the model's performance reveals equally impressive recall and F1-scores. Additionally, a detailed confusion matrix is provided, offering an in-depth understanding of the model's performance across different scenarios. This demonstrates the model's robustness and reliability, underscoring the effectiveness of the proposed methodology.

The subsequent sections of the paper are structured as follows: Section 2 provides an overview of the current state-of-the-art, delving into the latest advancements. The proposed methodology is expounded upon in Section 3, alongside exploring the evaluation parameters employed. Section 4 furnishes details about the dataset, experimental procedures, and the outcomes garnered from this study. Lastly, the paper concludes in Section 5 by summarizing key findings and delineating avenues for future research.

## 2. Literature Review

This section discusses different FL-based models for detecting COVID-19.

Sadly, the dangerous COVID-19 virus caused a pandemic that spread worldwide. Researchers, scientists, medical experts, and executives worldwide face a significant risk if it requires future treatment. In a recent study by Kandati et al. [23] a comprehensive strategy is recommended amalgamating Federated Learning (FL) with a Particle Swarm Optimization algorithm (PSO). This fusion aims to accelerate the government's responsiveness in addressing COVID-19-induced chest lesions. The Federated Particle Swarm Optimization method's efficacy is evaluated using a diverse dataset of images featuring chest lesions associated with COVID-19 infections alongside pneumonia cases from Kaggle's repository of chest X-ray images. In another innovative undertaking outlined in Bian et al.'s work [20], the focus is on enhancing the efficiency and precision of global

models during training through the incorporation of pre-trained models (PTMs) within the ambit of Federated Learning (FL). This endeavor seeks to curtail computational and communication demands, ultimately fostering improved outcomes. The user discusses introducing a secure aggregation protocol that utilizes homomorphic encryption and differential privacy. The findings in reference [21] show better performance than several other strategies already in use. This strategy is suitable and efficient for classifying COVID-19 chest X-rays. Reference [25] demonstrated that FL could aggregate COVID-19 information from many participating centers while respecting patient privacy. Recently, it has been demonstrated that other fields, such as cloud computing, can benefit from FL. The authors in [24] proposed a method for identifying CT scans of COVID-19 patients utilizing data from several hospitals to create a precise collaborative model.

The COVID-19 epidemic is one of the significant medical disasters that the entire world is undergoing. Reference [26] presents a collaborative FL model that enables COVID-19 detection screening from CXR (chest X-ray) pictures by numerous medical institutions. COVID-19 was classified as a pandemic by the World Health Organization in March 2020. To slow the pandemic's spread, practical testing is essential. In this regard, ref. [27] examines recent issues using DL and FL approaches for COVID-19 detection, focusing on that. One of the biggest challenges in detecting COVID-19 is the virus's quick spread and the need for valid testing models. For clinicians, this issue continues to be the most significant burden. The study presented in [28] examined reducing medical imaging analytic techniques in prediction, e-treatment, and data transfer to identify diseases as early as feasible. The outcomes reveal that the proposed approach did best in accuracy and precision. In [29], a proposed model focuses on COVID-19 data and creates a more accurate COVID-19 diagnosis model based on patient symptoms. The data are then reanalyzed using ML, and a computational model for determining if a person has COVID-19 based only on clinical data is created. Taking advantage of FL's quick, centrally managed experiment start and enhanced data traceability and assessing the impact of algorithmic adjustments, ref. [3] used FL to improve the prediction accuracy of the websites of testing sets.

The image classification job examines an algorithm's capacity to identify the image's constituent elements without necessarily locating them. This study addresses the establishment of a standardized dataset and the ensuing possibilities in object recognition that emerge as a consequence. In a related work by Dou et al. [14], the authors detailed the extensive data collection procedure of ILSVRC (ImageNet Large Scale Visual Recognition Challenge), presented a summary, and analyzed data and failure modes pertinent to these algorithms. The challenges of procuring comprehensive ground truth annotations are discussed. At the same time, noteworthy strides in categorical object recognition are underscored. "Object recognition" is employed inclusively to encompass both image classification and object identification. Similarly, in [10], the authors aim to create a reliable, transferable model to help with patient triage. They predicted that it would surpass regional models and more truly describe healthcare systems. On the test data for each customer, they compared the trained local models with the global FL model.

Due to substantial uncertainty and insufficient availability of requisite data, standard models have exhibited restricted accuracy in long-term prediction. Several attempts to tackle this problem have been made in the literature; however, optimizing current models' essential applicability and robustness abilities is still important. To predict the detection of the COVID-19 outbreak, the authors in [3] compare ML and computing models. To gain knowledge about the probable progression and impact of infections, having access to accurate outbreak estimation methods is essential. They examine and contrast ML and ML methods to forecast the COVID-19 outbreak.

### 3. Methodology

This section outlines the methodology for identifying COVID-19 chest X-ray images while excluding non-infected instances. Following a formal introduction, we delve into a comprehensive overview of our proposed approach: including the architecture of the ML

model integrated with FL, the model training process at the client end, and the mechanism for aggregating trained models on the server.

### 3.1. Formal Description of the Proposed Framework

The proposed framework is tailored for classifying chest X-ray images into infected and non-infected categories. The foundation of this approach is the FedAvg algorithm, which facilitates efficient distributed training while safeguarding data privacy. Let  $N$  denote the total number of participating clients in the FL process.

Let  $\theta_{\text{global}}$  be the global model parameters, and  $\theta_{\text{local}}^i$  represents the local model parameters of client  $i$  ( $1 \leq i \leq N$ ).

1. Initialization: The central server initiates the global model parameters  $\theta_{\text{global}}$  and disseminates them to a subset of clients  $C_k \subset 1, 2, \dots, N$ , where  $k$  signifies the communication round index.
2. Local Training: Each client, denoted as  $i$ , undertakes local training utilizing its own dataset. The objective of this training is to minimize the local loss,  $L_i(\theta_{\text{local}}^i)$ , which is a function of the model parameters. In this context,  $\theta_{\text{local}}^i$  represents the current model parameters for client  $i$ . The process of minimizing the local loss is represented mathematically in Equation (1) as follows:

$$\theta_{\text{local}}^{i,\text{new}} = \underset{\theta_{\text{local}}^i}{\text{argmin}} L_i(\theta_{\text{local}}^i) \quad (1)$$

In the above equation,  $\theta_{\text{local}}^{i,\text{new}}$  represents the updated model parameters for client  $i$  after the local training. These are obtained by finding the argument that minimizes the local loss function  $L_i$ .

3. Model Update: Subsequent to local training, each client  $i$  transmits its updated model parameters  $\theta_{\text{local}}^{i,\text{new}}$  back to the central server.
4. Aggregation: The central server aggregates the received local model parameters via weighted averaging, as depicted in Equation (2):

$$\theta_{\text{global}}^{k+1} = \frac{1}{|C_k|} \sum_{i \in C_k} |\theta_{\text{local}}^{i,\text{new}}| \quad (2)$$

### 3.2. Deep CNN Model

The classification model accepts an X-ray image  $X$  as input and generates the probability  $P_{\text{COVID-19}}$  of being COVID-19 infected. Let  $f_{\text{CNN}}(\cdot; \theta_{\text{global}})$  denote the CNN model's mapping function parameterized by  $\theta_{\text{global}}$ , as expressed in Equation (3):

$$P_{\text{COVID-19}} = f_{\text{CNN}}(X; \theta_{\text{global}}) \quad (3)$$

### 3.3. Overall Model Training Objective

Define the combined loss function using Equation (4):

$$L_{\text{total}} = \sum_{i=1}^N \frac{|C_i|}{\sum_{j=1}^N |C_j|} L_i(\theta_{\text{local}}^i) + \lambda L_{\text{CNN}} \quad (4)$$

Algorithm 1 starts with global model parameter initialization. A random subset of clients is selected over a set number of communication rounds. For each client, local training is executed iteratively until convergence, adjusting their local model parameters using stochastic gradient descent. The updated local models are aggregated at the central server through weighted averaging. The deep Convolutional Neural Network (CNN) model processes X-ray images, and post-processing techniques, such as lung segmentation and data augmentation, enhance robustness. The overall training objective involves a combined

loss function incorporating local and CNN classification loss. Implementation details, results, and conclusions provide further insights into the methodology's effectiveness.

---

**Algorithm 1:** Federated Learning for COVID-19 classification.

---

**Input:** Client datasets  $\{D_1, D_2, \dots, D_N\}$ , communication rounds  $T$ , learning rate  $\eta$ , regularization parameter  $\lambda$

**Output:** Global model parameters  $\theta_{\text{global}}$

- 1 Initialize global model parameters  $\theta_{\text{global}}$ ;
- 2 **for**  $t = 1$  **to**  $T$  **do**
- 3     Select random subset of clients  $C_k \subset \{1, 2, \dots, N\}$ ;
- 4     **for each client**  $i \in C_k$  **do**
- 5          $\theta_{\text{local}}^i \leftarrow \theta_{\text{global}}$ ;
- 6         **while not converged do**
- 7             Compute local loss  $L_i(\theta_{\text{local}}^i)$  using Equation (1);
- 8             Compute local gradient  $\nabla_{\theta_{\text{local}}^i} L_i(\theta_{\text{local}}^i)$ ;
- 9             Update local model parameters:  $\theta_{\text{local}}^i \leftarrow \theta_{\text{local}}^i - \eta \nabla_{\theta_{\text{local}}^i} L_i(\theta_{\text{local}}^i)$ ;
- 10         **end**
- 11         Send updated model  $\theta_{\text{local}}^i$  to central server;
- 12     **end**
- 13     Aggregate local models:  $\theta_{\text{global}}^{k+1} = \frac{1}{|C_k|} \sum_{i \in C_k} \theta_{\text{local}}^i$ ;
- 14 **end**

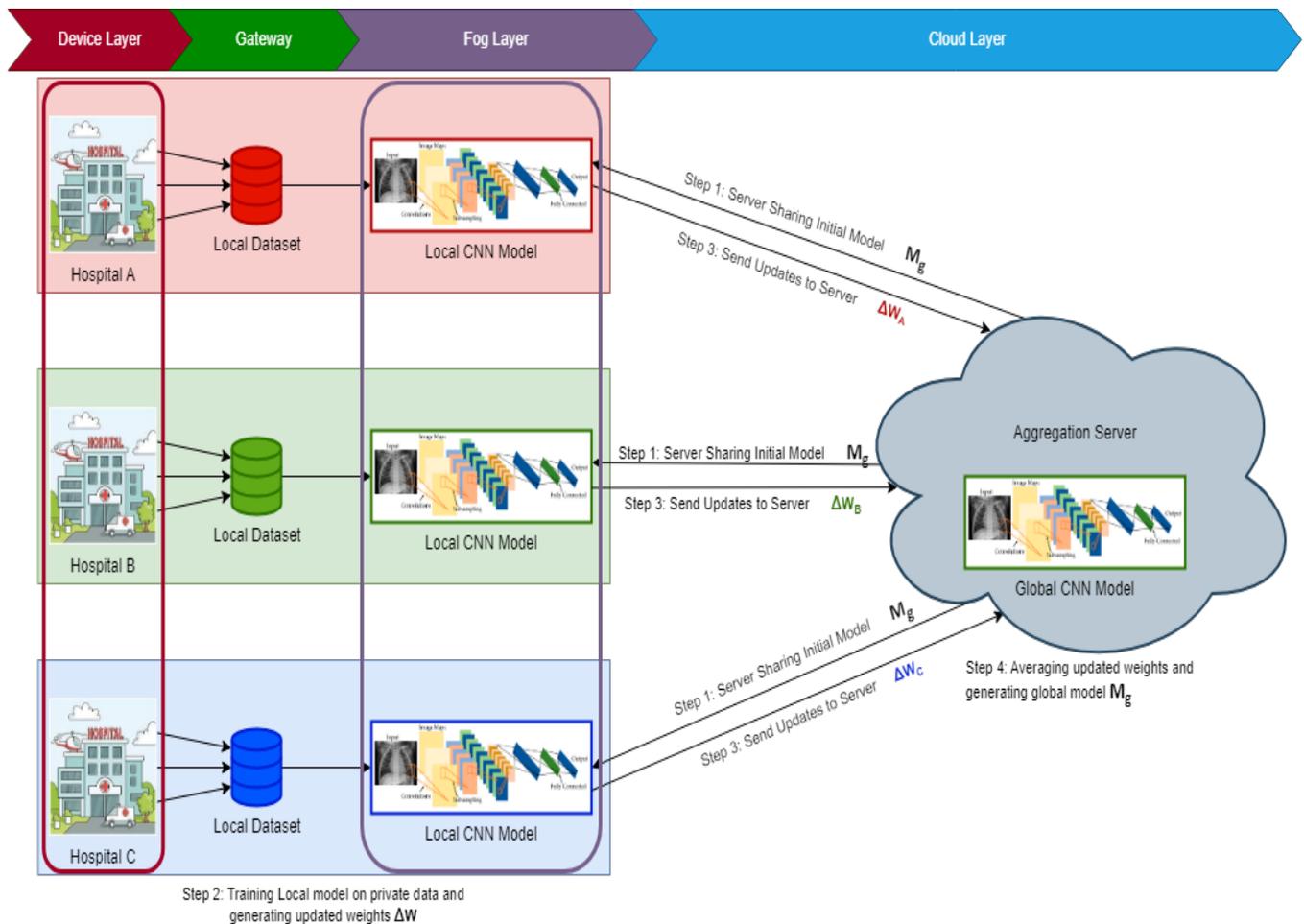
---

### 3.4. Proposed Framework Overview

To address the task of differentiating images capturing distinct health conditions, our approach revolves around exploring an FL structure rooted in a client–server setup (depicted in Figure 1). This architecture is underpinned by the FedAvg algorithm, which is renowned for its pivotal role. FedAvg operates as a communication-efficient mechanism within distributed training, seamlessly integrating with localized models. In this context, localized models maintain their data locally to uphold privacy, while a central server facilitates communication between these localized counterparts. This framework hinges upon a central parameter server tasked with orchestrating client alterations and overseeing a standardized global model. Through collaborative efforts, clientele leverage their distinctive datasets to formulate a robust model. Our proposition entails the design of an intricate CNN model primed to undertake tasks encompassing classification and feature extraction, particularly within the realm of chest imagery analysis. The model's capacity to gauge the likelihood of infection is grounded in scrutinizing input images. Elaborated specifications regarding this CNN model's architecture are presented in Section 3.5. The learning phase of this CNN model encompasses multiple communication cycles between clients and the central server. Before the commencement of training sessions, the CNN model is initialized with random base weights. We assume that each privately stored local X-ray image is accessible to clients. Each communication cycle comprises four distinct steps:

- Step 1: A global central architecture is maintained on the central aggregation server, which is initially endowed with weights. These weights are then broadcast to a subset of client devices selected at random (e.g., hospitals).
- Step 2: Subsequent to receiving the initial parameters, each client embarks on training using a subset of its data via mini-batch stochastic gradient descent (SGD) over multiple epochs. This endeavor aims to minimize local gradients. Notably, the convergence of each local model transpires at varying epoch counts. Clients iteratively refine their models by minimizing classification loss via categorical cross-entropy.

- Step 3: Users stationed at the client nodes update the server upon completion of their respective local training iterations (involving SGD across epochs on local data points).
- Step 4: After collecting data from all participating nodes and calculating an average model, the central server updates the global model parameters.



**Figure 1.** Proposed model architecture.

A single Federated Learning (FL) round for our CNN model entails these four phases. This sequence of actions is then repeated iteratively (across rounds). The server dispatches the updated global model parameters from the preceding round during each new round. Furthermore, if an abundant number of clients are available, the subset of participating clients can be modified between rounds.

### 3.5. Preprocessing, Model Training, and Postprocessing

The Hounsfield units for each volume cut for preprocessing before their values show at [1.0, 1.0]. Experimental observations show that standardizing the data to have a mean of zero and unit variance and using individual volume statistics rather than global dataset statistics helps improve the model’s generality. Furthermore, using three neighboring slices as input for convolutional neural models after normalizing the data yields better results. Every local client used the dataset to improve his or her model for one epoch during each round of FL. All local clients used the Adam optimizer, which had a learning rate of  $1 \times 10^{-4}$ , 0.9, 0.999, and  $1 \times 10^{-7}$  for beta1, beta2, and epsilon. Following the analysis of the biases of the dataset, we observed that the classification algorithms are mainly based on the pixels at the edge of the image to make their decision. Therefore, we aim to remove these edges and focus only on the lungs with the information needed for the classification. This required

using a U-Net neural network pretrained on CXR images and developed especially for lung segmentation. After segmentation, the images were cropped around the lungs with a margin of 10 pixels. The homogenized and cropped images were then used to build our new dataset for training and testing the DL model [30].

A variety of data augmentation techniques, such as random horizontal and vertical flips (with a probability of 50%), lucky clockwise rotations (with an amplitude ranging from 0.1 to 0.1), random horizontal and vertical translations (with a range of 0.1 to 0.1 of the input image's length or width), and random shear and scaling were applied to the training data (with a range of 0.1 to 0.1). Different Python libraries were used for data preprocessing before giving the data to the model for training. In this research work, we also utilized some of those libraries. Panda and Numpy libraries were used for data cleaning. Panda, Numpy, and Matplotlib libraries were used for data processing. SKlearn was also used for data augmentation purposes. The Numpy and TensorFlow programming frameworks were used in order to put these strategies into action. The Keras library facilitated accessing different layers of models for classification. A relatively small piece of the training data was extracted at each process step to assess the model's performance. If the global performance of the model on the local test dataset did not improve for five federated rounds in a row, we considered the training to have collapsed and the FL to have come to an end. The deep CNNs were built using a single NVIDIA TitanXp GPU (graphics processing unit).

During the post-processing phase, we used non-maximum suppression<sup>30</sup>, a method prevalent in image processing, to separate overlapping elements. We separated an overlapping series of frames using non-maximum suppression to create bounding boxes with the greatest anticipated probability. For more accuracy, we eliminated any bounding boxes in an image with a chance lower than a certain threshold, which was determined by looking at a collection of projected bounding boxes. Out of the remaining bounding boxes, we chose the one with the greatest probability and eliminated those with an Intersection over Union (IoU) value larger than 0.5. In addition, we used an open-source AI model for lung segmentation<sup>31</sup> to eliminate any false-positive detections that occurred outside of the lung area.

### 3.6. Model Architecture

This research suggests a collaborative and decentralized architecture for chest X-ray imaging-based COVID-19 screening. We aim to demonstrate that a deep CNN model can perform federated learning to benefit from rich private data interchange while upholding privacy. The specific design of the CNN is not the focus of this work because many different architectural solutions could only slightly improve or worsen the overall performance. The information exchange flow between the central server and federated clients is shown in Figure 2. The parts that follow give information on these two parties as discussed subsequently:

#### 3.6.1. Client-Side Model Update

On the client side, training is carried out since every federated client has access to the same dataset and has the computing capacity to carry out mini-batch Stochastic Gradient Descent (SGD). Clients whose CNN architecture and loss functions are the same are ignored. Algorithm 2 provides the steps used in training. At the beginning of the round, a copy of the global model is sent to each client's computer so that it may be initialized. Once the client has finished the same number of iterations as there are local epochs, they compute a gradient update to generate the new model, which is then shared with the aggregation server. During this training process, the local data for each customer are treated as private information and are never discussed with anyone else. This describes the procedure done on the client side. The client side can be stated as local algorithms working on Hospitals A, B, and C with local datasets. On the client side, the model is trained, and weights are sent to the global model. During training, dataset  $P_k$  is divided into equal sizes of batch  $B$ . For each batch  $B$ , the gradient is computed, and the weight of the local model is updated.

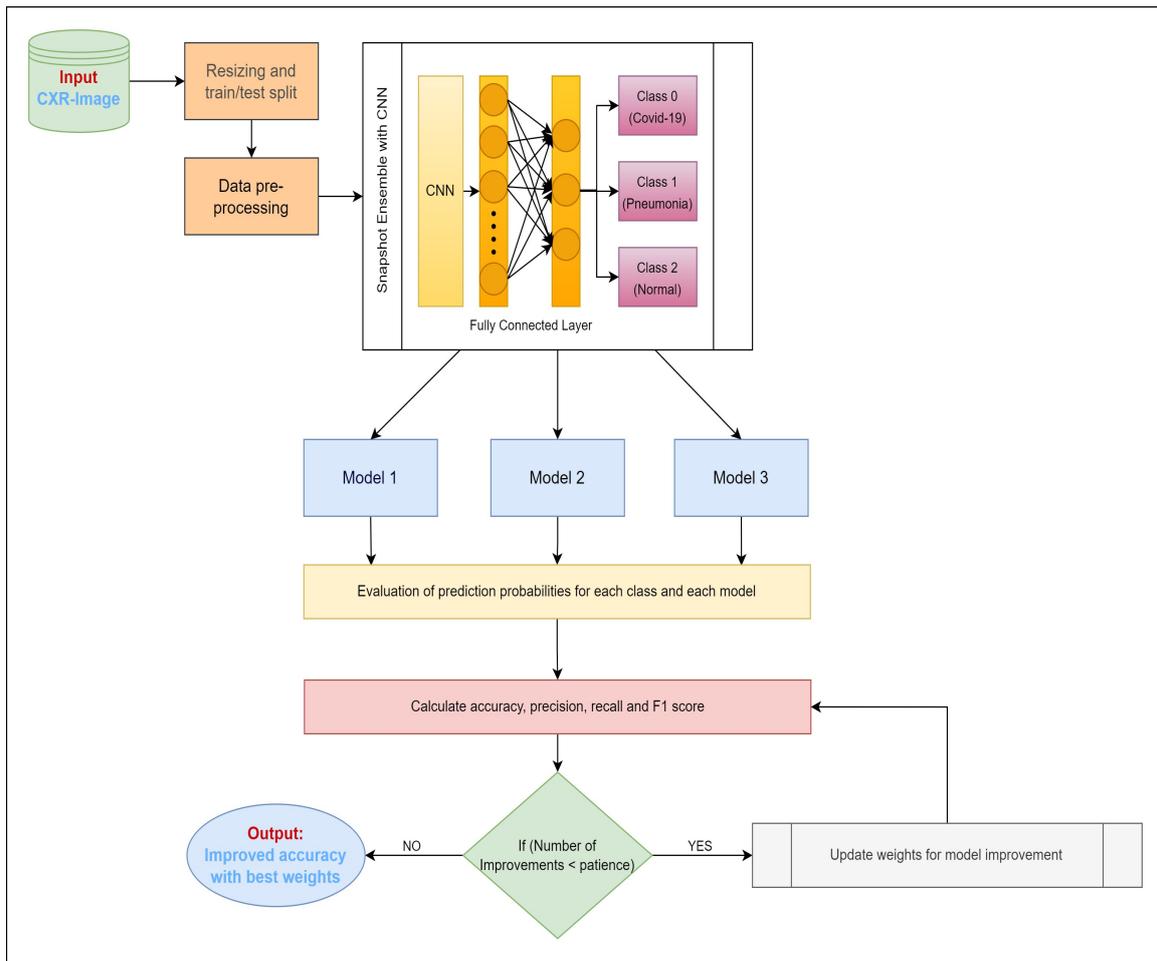


Figure 2. Proposed model flow.

**Algorithm 2:** Client-side algorithm.

```

Input: loss function  $l_s$  and local learning rate  $\eta$ 
Result: number of local epochs and local training data
1 begin
2   Algorithm
3   Client update ( $w^t$ )
4    $w \leftarrow w^t$ 
5    $B_t \leftarrow$  Split  $k$  into batches of size  $B_t$ 
6   for each local epoch  $j$  from 1 to  $E$  do
7     for each batch  $bt$  in  $B_t$  do
8       compute gradient  $g_i^b \leftarrow \Delta l(w; bt)$ 
9       Update local model  $w \leftarrow w - \eta g_i^b$ 
10    end
11  end
12 end

```

3.6.2. Server-Side Model Aggregation

An aggregation server is responsible for managing the global model. This server also ensures that each participating client receives the initial model and tracks the progression of the training. The server receives synchronized updates from all clients at each federated round and utilizes these changes to generate a new model with updated parameters following Equation (2). Algorithm 3 describes the procedure done on the server side. The server side can be stated as a global algorithm working on Hospitals A, B, and C with the

global dataset. The model is trained on the server side, and weights are sent to the local models. During training, weights are updated after each round  $t$ , i.e., 1, 2, 3, and so on. At the end, the gradient value for the global model is computed.

---

**Algorithm 3:** Server-side algorithm.

---

**Input:** loss function  $l_s$  and local learning rate  $\eta$   
**Result:** number of federated rounds

```

1 begin
2   Algorithm
3   Aggregating ( $C, K$ )
4   Initialize global model  $\mathbf{w}^0$ 
5   for each round  $r = 1, 2, 3, \dots, R$  do
6      $m_c \leftarrow \max(C \times K, 1)$ 
7      $S_t \leftarrow$  random set of  $m_c$  clients
8     for each client  $k$  belonging to  $S_t$  do
9       Send  $\mathbf{w}^{t-1}$  to client  $k$ 
10      Compute gradient  $\mathbf{g}_i^b \leftarrow \Delta l(\mathbf{w} : \mathbf{b})$ 
11    end
12  end
13 end
```

---

#### 4. Findings and Discussion

This section presents an evaluation, and the outcomes of our suggested paradigm are described. The performance of our model was evaluated by simulation. The experimentation setup used to evaluate and train the model is also elaborated on here. Results show that our model is better regarding accuracy, precision, and recall.

##### 4.1. Description of the Dataset

The dataset used in this study comprises 2473 X-ray images from the COVID-19 CXR dataset. This dataset is sourced from the COVID-19 Radiography Database, available on Kaggle. It includes 1345 viral pneumonia images, 219 COVID-19 images, and 1341 regular CXR images. The detailed dataset composition is presented in the following table.

##### 4.2. COVID-19 Chest X-ray Database

A collaborative effort between researchers from the University of Qatar, Doha, Qatar; Dhaka University, Bangladesh; Pakistan; and Malaysia resulted in a chest X-ray image dataset. This dataset includes positive samples of COVID-19, regular X-rays, and viral and bacterial pneumonia images. The dataset was released in multiple phases. The initial release contained 219 COVID-19, 1341 regular, and 1345 viral chest X-ray images. The COVID-19 class has since been expanded to 1200 CXR images. In a subsequent update, the dataset was expanded further with the addition of 3616 COVID-19-positive cases, 10,192 normal cases, 6012 lung opacity non-COVID infections, and 1345 viral influenza images. Regular updates are performed to incorporate new X-ray images of COVID-19 pneumonia cases as they become available. The details of datasets are provided in Table 1, and the dataset distribution is outlined in Table 2.

**Table 1.** Dataset details for chest X-ray images.

Image Type	Data Source	Number of Images
COVID-19	PadChest database [31], Germany [32], SIRM, Kaggle, GitHub, Twitter [33–35]	2473
Normal	RSNA dataset [36]	10,192
Lung Opacity	Radiological Society of North America CXR dataset [37]	6012
Viral Pneumonia	CXR (pneumonia) dataset [38]	1345

**Table 2.** Dataset and number of images.

Dataset	Number of Images
Dataset division	Training (70%) and Testing (30%)
COVID-19 images	3616
Normal images	10,192
Lungs images	6012
Influenza images	1345
Total images	21,165
Images per hospital	7055

#### 4.3. Experimentation Setup

In opposition to previous models built using TensorFlow version 1.8, the suggested model was established with open-sourced TensorFlow federation version 2.1.0, with Keras providing the backend in both cases. The use of TensorFlow 1.8 for previous models was primarily based on the existing infrastructure and legacy models already built using that version. However, TensorFlow 2.1.0, which we used for our proposed model, offers several improvements over the older versions. It brings tighter Keras integration, easier model exporting for deployment, and enhanced performance optimizations. These features aligned well with the requirements of our proposed model, prompting us to transition to the newer version. Different Python libraries were used for data preprocessing before giving the data to the model for training. In this research work, we also utilized some of those libraries. Panda and Numpy libraries were used for data cleaning. Panda, Numpy, and Matplotlib libraries were used for data processing. SKlearn was also used for data augmentation purposes. The Numpy and TensorFlow programming frameworks were used in order to put these strategies into action. The Keras library facilitated accessing different layers of models for classification. The experiment was run on a PC desktop with an Intel i5 6th generation CPU, an NVIDIA Titan GPU, 16 GB RAM, and a 3.5 GHz speed control. Each dataset was split into three portions: 70% for instruction, 20% for evaluation, and 10% for validation. The proposed system was trained using TensorFlow federation version 2.1.0, and the other classic model was trained using TensorFlow version 1.8; both used 70% of the three databases: 23,804, 31,501, and 2443 pictures. Python 3.9.1 was used to create these models. Table 3 summarizes the parameters and their values used in this experiment.

**Table 3.** Experiment parameters and their values.

Simulation Parameters	Value
CPU	Intel i5 6th generation
RAM	16 GB
Processor frequency	3.5 GHz
GPU	NVIDIA Titan
Federated Learning version	2.1.0
TensorFlow version	1.8
Python version	3.9.1

#### 4.4. Evaluation Metrics

In this part, the discussion of the evaluation metrics is given and offers a thorough analysis of the outcomes. The classified accuracy uses statistics for evaluating the classification of the model. It is calculated by dividing the total number of examples (images) in the database under examination by the set of instances (images) that were correctly categorized. The numerical equivalent is:

$$Accuracy(A_c) = \frac{T_{+ve} + T_{-ve}}{T_{+ve} + T_{-ve} + F_{+ve} + F_{-ve}} \quad (5)$$

Here,  $T_{+ve}$  stands for true positive, and the number of medical images properly defined as true is counted.  $F_{+ve}$  stands for false positive and is a count of medical photos wrongly defined as false.  $T_{-ve}$  stands for true negative and is the count of correct medical images identified as true, and  $F_{-ve}$  stands for false negative. It is the count of medical images incorrectly identified as false.

In image classification systems, recall and precision are used for performance measurement. The proportion of correct positive images to all classified images is known as precision. The formula for precision ( $P_r$ ) is:

$$Precision(P_r) = \frac{T_{+ve}}{T_{+ve} + F_{+ve}} \quad (6)$$

Along with precision ( $P_r$ ), performance measurement is accomplished by *Recall* ( $R_c$ ). The proportion of correctly categorized images to all images in the database is known as recall. It can be summed up as follows in formula form:

$$Recall(R_c) = \frac{T_{+ve}}{T_{+ve} + F_{-ve}} \quad (7)$$

The Harmonic mean of the recall and precision is the *F-score*; a greater number indicates that the system has superior predictive capacity. The performance of systems cannot be judged just based on precision or recall. The *F-score* can be calculated as follows:

$$F\text{-score} = 2 \left( \frac{P_r * R_c}{P_r + R_c} \right) \quad (8)$$

#### 4.5. Accuracy

Figure 3 explains the COVID-19 image classification by the two different models, i.e., the local and the global model. The accuracies of both software configurations varies for different hospital models. As the figure explains, the accuracy of the local model was moderate throughout all hospital models. In contrast, the FL model surpassed the local model in all three hospitals in patient detection regarding accuracy and precision. The figure also shows that the accuracy of the FL model constantly increased in the respective hospitals, and the local model has the lowest accuracy in Model B hospitals.

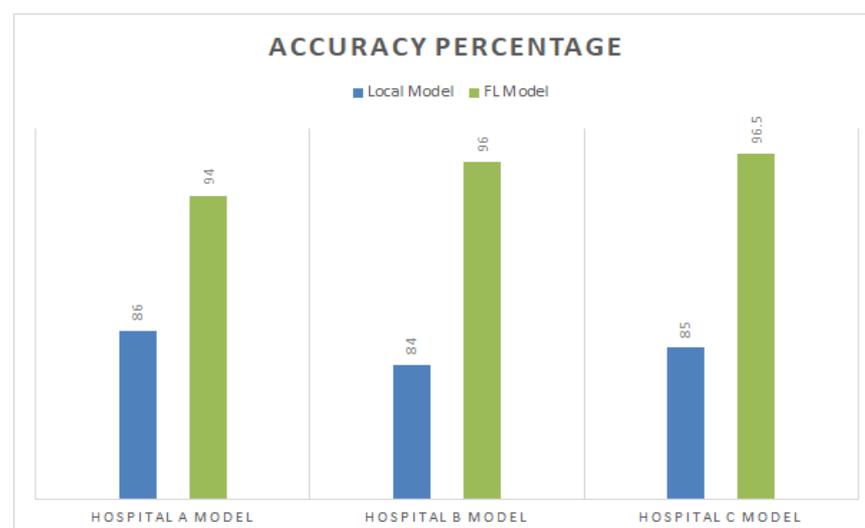
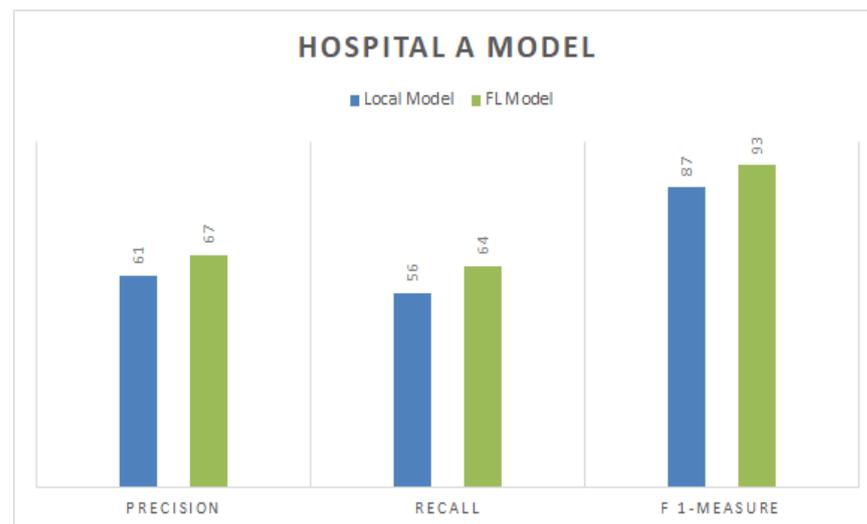


Figure 3. Accuracy.

#### 4.6. Precision, Recall, and F1-Measure for Hospital A

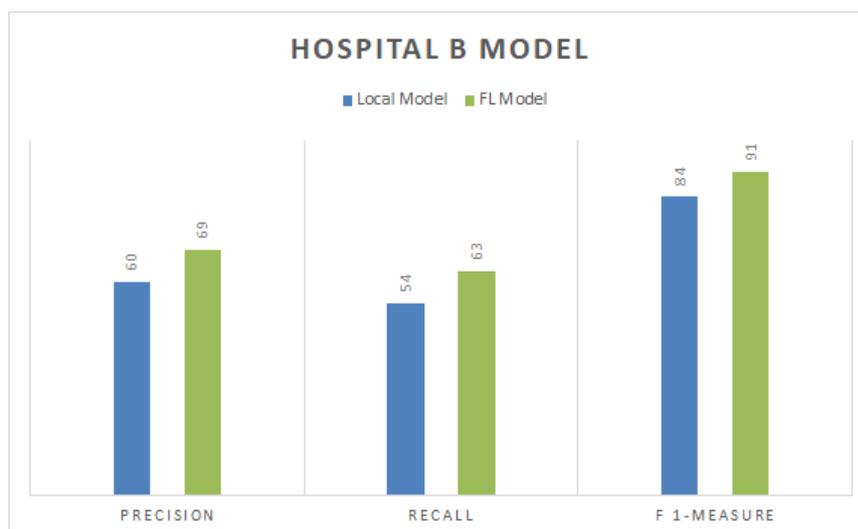
Figure 4 represents the precision, recall, and F1-measure values for the local and FL-based models for a particular dataset of Hospital A. The figure shows that the FL-based model scored the best precision, recall, and F1-score, with respective values of 67%, 64%, and 93%. At the same time, other CNN-based local models gained 61%, 56%, and 87% precision, recall, and F1-measure scores, respectively. Models that are based locally often necessitate longer training periods compared to those that employ FL. This is primarily due to the distributed nature of local models, where the data and the computational resources are scattered across multiple devices or locations. This distribution can complicate the process and extend the time required for model training. On the other hand, Federated-Learning-based models have proven to be more efficient. FL allows a model to be trained across multiple decentralized devices or servers, which hold local data samples without exchanging them. It allows for faster, more efficient training because it eliminates the need for data centralization and, therefore, speeds up the entire process. Comparatively, when it comes to efficiency and speed, FL-based models outperform local CNN-based models. The latter often rely on local resources and data, which can limit their performance and extend their training time. Hence, strategies that leverage FL are generally considered superior to those based on a local CNN.



**Figure 4.** Precision, recall, and F1-measure for Hospital A.

#### 4.7. Precision, Recall, and F1-Measure for Hospital B

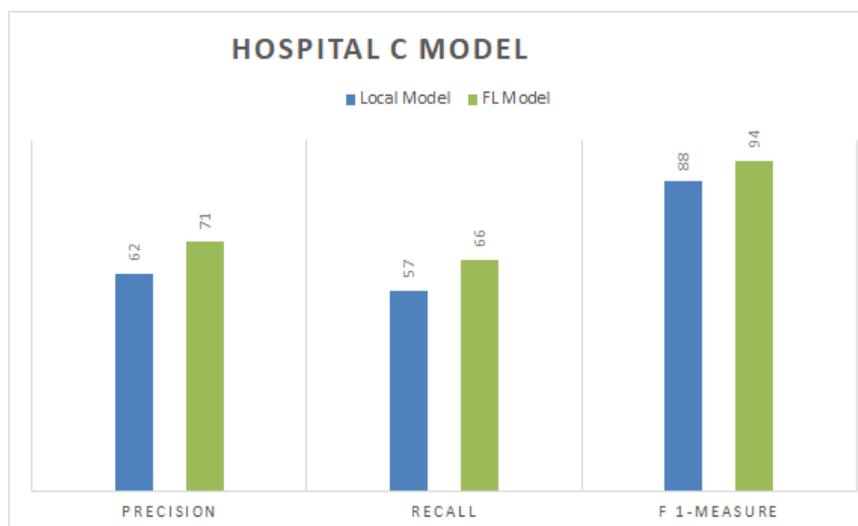
Figure 5 represents the precision, recall, and F1-measure values for the local and FL-based models for a particular dataset of Hospital B. The Figure shows that the FL-based model scored the best precision, recall, and F1-score, with respective values of 69%, 63%, and 91%. In contrast, other CNN-based local models gained 60%, 54%, and 84% precision, recall, and F1-measure scores, respectively. Local models also require more time for training than FL-based models due to their distributed nature. This demonstrates that approaches based on FL are more efficient than those based on a local CNN.



**Figure 5.** Precision, recall, and F1-measure for Hospital B.

#### 4.8. Precision, Recall, and F1-Measure for Hospital C

Figure 6 represents the precision, recall, and F1-measure values for the local and FL-based models for a particular dataset of Hospital C. The figure shows that the FL-based model scored the best precision, recall, and F1-score, with respective values of 71%, 66%, and 94%. In comparison, other CNN-based local models gained 62%, 57%, and 88% precision, recall, and F1-measure scores, respectively. Local models also require more time for training than FL-based models due to their distributed nature. This demonstrates that approaches based on FL are more efficient than those based on a local CNN.



**Figure 6.** Precision, recall, and F1-measure for Hospital C.

#### 4.9. Confusion Matrix

The confusion matrix for the training phase is shown in Figure 7. The results show that there were 10,946 true positives and 7860 true negatives. On the other hand, there were 291 false positives and 624 false negatives. In contrast, the confusion matrix for the validation phase is shown in Figure 8. The results show that there were a total of 608 true positives and 2005 true negatives. On the other hand, there were 32 false positives and 87 false negatives. The analysis of our confusion matrix results reveals notable percentage differences between the training and testing (validation) phases, shedding light on how our COVID-19 screening model using chest X-ray images generalizes to unseen data. In the training phase, we achieved a substantial count of 10,946 true positives, indicating the

model’s proficiency in correctly identifying COVID-19 cases. However, during the testing phase, the true positive count dropped significantly to 608, representing a percentage difference of approximately 94.44%. This decline suggests a challenge in the model’s ability to generalize effectively to new data, particularly in identifying COVID-19 cases in real-world scenarios.

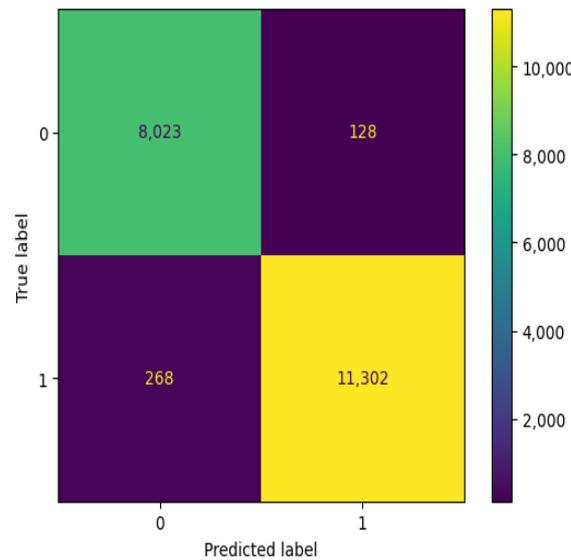


Figure 7. Confusion matrix for training.

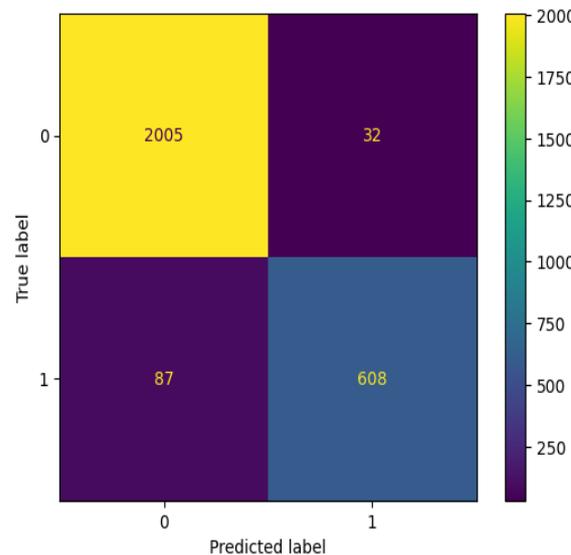


Figure 8. Confusion matrix for testing.

Conversely, there was a noticeable decrease in false positives during testing, with a percentage difference of around 89.02%. This is a positive development, as it indicates a lower rate of false alarms when applying the model to unseen data. Additionally, the false negatives decreased by approximately 86.10% during testing, implying a reduction in missed diagnoses, which is crucial for patient care. The most pronounced percentage difference is in the true positive rate, underscoring the need for further optimization to enhance the model’s performance on unseen data. These findings highlight the critical importance of rigorous model evaluation and validation, as they offer insights into the model’s generalization capabilities and potential areas for improvement. Future work should focus

on fine-tuning the model to achieve more consistent performance between the training and testing phases, ultimately increasing its reliability in real-world healthcare applications.

#### 4.10. Discussion

This research shows that it is possible to aggregate COVID-19 data across detailed listings while respecting patient privacy using FL. During COVID-19, when no time is available to set up complex material around institutions and countries, this decentralized training strategy is a crucial scaling enabler of AI-based technologies. The application of FL has recently been shown in different industries, such as edge computing for digital equipment. Still, diagnostic imaging is getting more complicated. It involves common challenges, such as data with high dimensionality and with improper cohort sizes, that exert untested influences on the current FL methods.

This is the first study showing FL's feasibility and efficiency for COVID-19 image processing. In this field, teamwork is vital during crises around the world. According to our experimental findings, FL outperformed single-site models and their ensembles regarding generalization ability. This is an example of successful decentralized optimization using various training with multidimensional data. In a solid evidence study, we demonstrate the viability of a convolution neural federated DL system for identifying chest CT problems in COVID-19 disease patients. Significantly, the AI model developed using cohorts of Hong Kong performed well on outer, hidden, independent datasets gathered by institutions in Europe and Asia in addition to internal testing cases.

These international centers used different scanner brands, imaging protocols, and groups of patients, and the study participants' COVID-19 pneumonia severity varied. The diversity of the data in this comprehensive study shows that it is possible to create reliable and generalized AI techniques for battling COVID-19 through image processing in diverse healthcare situations. Our model may allow proper use in reality as it usually requires 40 ms to assess one CT volume. The unreliability of AI techniques was expected to be improved by multicenter training, which has been proven in various medical rummaging scenarios 25–27. The Individual-2-model trained on 4146 photos of CT slices outperformed the two single-site systems trained on 958 and 660 data points, respectively, in our studies on all three external testing sites. As a result, it was discovered that more extensive training databases could enhance the model's performance on new datasets.

Furthermore, merging all three main training sites might improve the test accuracy, even if two areas contributed fewer cases. This indicates that larger data scales and the diversity of richer data through imaging scanners and protocols are also crucial for lowering model error and enhancing generalizability. In this view, a collaboration between various clinical institutions is a necessary step in creating AI for broader adoption, particularly in the case of the COVID-19 pandemic, when global cooperation is vital. Despite enrolling patients from seven clinical centers in various regions, each participating institution only had a modest amount of patients. There was an imbalance between the locations due to the pandemic performance management system receiving more cases than other hospitals. This reflected the real-world circumstance that most single sites had difficulty finding enough COVID-19 patients to train their internal AI systems.

Therefore, multicenter research with collaborative data merging efforts is crucial and helpful to handle the COVID-19 long-tail dispersion. More patients and institutions will be involved in FL in upcoming work. However, it is essential to note that the current solid evidence study's tiny patient population cannot prevent the creation of a DL model due to a sizable number of lesions identified by the CT volumes. The German cohort's model extended less effectively compared to those of other foreign affiliates. The patient populations' varied ethnicities could be one factor. The cohort of hand annotations is not directly compatible through training examples due to the diverse lesion annotation processes used at various clinical sites, known in computer vision as idea shift. For instance, the training data noted floor occlusion and consolidating.

In contrast, the German cohort had a small number of pleural effusion lesions—atypical in COVID-19—that contrasted only slightly with healthy lung tissue. There were 15 cases with mild lesions (difficult to visualize in the window of the lung), 5 cases with a diffuse lesion that was inappropriate for detection processes that anticipate lesion bounding boxes, and 1 case was disqualified due to no CT findings. Image characteristic of individuals eliminated from the German cohort are shown in Figure 3. This allowed the model to maintain 35 examples while achieving an AUC of 88.15% (95% CI 86.38–89.91). The model obtained an AUC of 77.15% (95% CI 72.84–81.47) when tested on all 56 cases of the German data. The tool for CT abnormality detection is meant to be applicable in conjunction with the typical visual assessment performed by qualified radiologists. By providing quantifiable data during clinical decision making, the AI tool helps the expert. The loop expert also serves as a buffer against incorrect predictions, such as false positives in the scanning process, that are not worrisome.

It is crucial to remember a shift in concepts, such as that seen in the cohort of Germans, will prevent upcoming multinational trials that employ FL if each site follows a set protocol. Following a set protocol was not feasible due to the random design of the study, which comprised data from different websites and each cohort's independent collection of data. Despite these drawbacks, our multinational verification highlights the practical challenges of conducting such research while demonstrating our strategy's potential. In conclusion, the CNN-based AI model is efficient for detecting CT anomalies of COVID-19 disease patients after being trained using the confidential FL model. The work of AI in delivering less cost and scalable methods for the burden of lesion estimation for enhanced treatment of clinical diseases is demonstrated by the overall unreliability of regional and worldwide cohorts as external, which is beneficial by adding various datasets.

Despite the significant contributions made with the fog-based FL model introduced in this study, there are a few constraints to consider. Primarily, the model's effectiveness is heavily reliant on both the quality and quantity of data provided by the participating nodes. If the data are insufficient or skewed, this could impact the model's accuracy. Another constraint lies in the decentralized nature of the FL model itself. While this decentralization offers enhanced data privacy, it also introduces complexity in terms of coordination and synchronization among nodes. These complexities could result in increased computational time and resource usage. It is also worth noting that the model's performance may vary based on different datasets and real-world scenarios, which further underlines the importance of robust and diverse data sources. Future research should aim to address these limitations to further improve the performance and applicability of the FL model.

## 5. Conclusions and Future Work

In order to construct predictive models using existing AI techniques, central data storage and training are frequently required. This increases computational complexity and threatens privacy. In order to solve the issue, this research suggests using a system based on FL to improve the interpretation of heterogeneous CT scanners from many sources. While preserving security and privacy, this method distributes the data among the hospitals. Additionally, extreme learning machines and capsular networks are used to successfully extract and classify characteristics from a number of publicly viewable heterogeneous CT imaging datasets to identify COVID-19. Additionally, FL is used to train collaborative institutions. Further, using chaotic encryption keys increases faith in maintaining confidentiality and privacy during data retrieval and sharing. The algorithm was tested extensively and compared with other DL methods. The findings reveal that the suggested model performs better than the competition regarding accuracy, precision, recall, and high F1-scores. It also beats the competition in terms of trust and data relevance. The proposed model performed better than previous approaches but still requires innovation when managing clinical datasets. Future efforts must be made to combine FL with blockchain to optimize the solution's cost effectiveness.

**Author Contributions:** Conceptualization, N.T. and M.B.; Methodology, N.T. and M.B.; Software, M.B.; Validation, S.A.M., M.A., H.S.A. and N.T.; Formal Analysis, N.T., S.A.M., H.A.A.A. and M.A.; Investigation, M.B.; Resources, H.A.A.A. and Y.A.A.; Data Curation, H.A.A.A. and Y.A.A.; Writing—Original Draft Preparation, N.T. and M.B.; Writing—Review and Editing, S.A.M., M.A., H.S.A. and Y.A.A.; Visualization, N.T. and M.B.; Supervision, N.T.; Project Administration, N.T., S.A.M. and M.A.; Funding Acquisition, H.S.A., H.A.A.A. and Y.A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-RG23148).

**Data Availability Statement:** The datasets used in this study are publicly available, and their links can be found in the references section.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Li, H.; Li, C.; Wang, J.; Yang, A.; Ma, Z.; Zhang, Z.; Hua, D. Review on security of federated learning and its application in healthcare. *Future Gener. Comput. Syst.* **2023**, *144*, 271–290. [[CrossRef](#)]
2. Farooq, U.; Tariq, N.; Asim, M.; Baker, T.; Al-Shamma'a, A. Machine learning and the Internet of Things security: Solutions and open challenges. *J. Parallel Distrib. Comput.* **2022**, *162*, 89–104. [[CrossRef](#)]
3. Ardabili, S.F.; Mosavi, A.; Ghamisi, P.; Ferdinand, F.; Varkonyi-Koczy, A.R.; Reuter, U.; Rabczuk, T.; Atkinson, P.M. COVID-19 outbreak prediction with machine learning. *Algorithms* **2020**, *13*, 249. [[CrossRef](#)]
4. Hui, D.S.; Azhar, E.I.; Madani, T.A.; Ntoumi, F.; Kock, R.; Dar, O.; Ippolito, G.; Mchugh, T.D.; Memish, Z.A.; Drosten, C.; et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int. J. Infect. Dis.* **2020**, *91*, 264–266. [[CrossRef](#)]
5. Moqurrab, S.A.; Anjum, A.; Tariq, N.; Srivastava, G. An efficient framework for semantically-correlated term detection and sanitization in clinical documents. *Comput. Electr. Eng.* **2022**, *100*, 107985. [[CrossRef](#)]
6. Kuzinkovas, D.; Clement, S. The detection of covid-19 in chest x-rays using ensemble cnn techniques. *Information* **2023**, *14*, 370. [[CrossRef](#)]
7. Albahli, S.; Nawaz, M. TSM-CV: Twitter Sentiment Analysis for COVID-19 Vaccines Using Deep Learning. *Electronics* **2023**, *12*, 3372. [[CrossRef](#)]
8. Rasha, A.H.; Li, T.; Huang, W.; Gu, J.; Li, C. Federated learning in smart cities: Privacy and security survey. *Inf. Sci.* **2023**, *632*, 833–857.
9. Hassan, M.; Tariq, N.; Alsirhani, A.; Alomari, A.; Khan, F.A.; Alshahrani, M.M.; Ashraf, M.; Humayun, M. GITM: A GINI Index-Based Trust Mechanism To Mitigate And Isolate Sybil Attack In RPL-Enabled Smart Grid Advanced Metering Infrastructures. *IEEE Access* **2023**, *11*, 62697–62720. [[CrossRef](#)]
10. Dayan, I.; Roth, H.R.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.Z.; Liu, A.; Costa, A.B.; Wood, B.J.; Tsai, C.S.; et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **2021**, *27*, 1735–1743. [[CrossRef](#)]
11. Ahmed, S.; Yong, J.; Shrestha, A. The Integral Role of Intelligent IoT System, Cloud Computing, Artificial Intelligence, and 5G in the User-Level Self-Monitoring of COVID-19. *Electronics* **2023**, *12*, 1912. [[CrossRef](#)]
12. Moqurrab, S.A.; Anjum, A.; Tariq, N.; Srivastava, G. Instant\_Anonymity: A Lightweight Semantic Privacy Guarantee for 5G-enabled IIoT. *IEEE Trans. Ind. Inform.* **2022**, *19*, 951–959. [[CrossRef](#)]
13. Riedel, P.; von Schwerin, R.; Schaudt, D.; Hafner, A.; Späte, C. ResNetFed: Federated Deep Learning Architecture for Privacy-Preserving Pneumonia Detection from COVID-19 Chest Radiographs. *J. Healthc. Inform. Res.* **2023**, *7*, 203–224. [[CrossRef](#)] [[PubMed](#)]
14. Dou, Q.; So, T.; Jiang, M.; Liu, Q.; Vardhanabhuti, V.; Kaissis, G.; Li, Z.; Si, W.; Lee, H.; Yu, K.; et al. Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study. *NPJ Digit. Med.* **2021**, *4*, 60. [[CrossRef](#)] [[PubMed](#)]
15. Lee, T.F.; Chang, I.P.; Su, G.J. Compliance with HIPAA and GDPR in Certificateless-Based Authenticated Key Agreement Using Extended Chaotic Maps. *Electronics* **2023**, *12*, 1108. [[CrossRef](#)]
16. Moqurrab, S.A.; Tariq, N.; Anjum, A.; Asheralieva, A.; Malik, S.U.; Malik, H.; Pervaiz, H.; Gill, S.S. A deep learning-based privacy-preserving model for smart healthcare in Internet of medical things using fog computing. *Wirel. Pers. Commun.* **2022**, *126*, 2379–2401. [[CrossRef](#)] [[PubMed](#)]
17. Malik, H.; Naeem, A.; Naqvi, R.A.; Loh, W.K. DMFL\_Net: A Federated Learning-Based Framework for the Classification of COVID-19 from Multiple Chest Diseases Using X-rays. *Sensors* **2023**, *23*, 743. [[CrossRef](#)]
18. Flores, M.; Dayan, I.; Roth, H.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.; Liu, A.; Costa, A.; Wood, B.; et al. Federated Learning used for predicting outcomes in SARS-COV-2 patients. *Res. Sq.* **2021**, *1*.

19. Mosaiyebzadeh, F.; Pouriye, S.; Parizi, R.M.; Sheng, Q.Z.; Han, M.; Zhao, L.; Sannino, G.; Ranieri, C.M.; Ueyama, J.; Batista, D.M. Privacy-Enhancing Technologies in Federated Learning for the Internet of Healthcare Things: A Survey. *Electronics* **2023**, *12*, 2703. [CrossRef]
20. Bian, G.; Qu, W.; Shao, B. Blockchain-Based Trusted Federated Learning with Pre-Trained Models for COVID-19 Detection. *Electronics* **2023**, *12*, 2068. [CrossRef]
21. Kairon, P.; Bhattacharyya, S. COVID-19 outbreak prediction using quantum neural networks. In *Intelligence Enabled Research*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 113–123.
22. Kalliatakis, G.; Stamatiadis, G.; Ehsan, S.; Leonardis, A.; Gall, J.; Sticlaru, A.; McDonald-Maier, K.D. Evaluating deep convolutional neural networks for material classification. *arXiv* **2017**, arXiv:1703.04101.
23. Kandati, D.R.; Gadekallu, T.R. Federated learning approach for early detection of chest lesion caused by COVID-19 infection using particle swarm optimization. *Electronics* **2023**, *12*, 710. [CrossRef]
24. Kanne, J.P.; Little, B.P.; Chung, J.H.; Elicker, B.M.; Ketai, L.H. Essentials for radiologists on COVID-19: An update—Radiology scientific expert panel. *Radiology* **2020**, *296*, E113–E114. [CrossRef] [PubMed]
25. Feki, I.; Ammar, S.; Kessentini, Y.; Muhammad, K. Federated learning for COVID-19 screening from Chest X-ray images. *Appl. Soft Comput.* **2021**, *106*, 107330. [CrossRef] [PubMed]
26. Annavarapu, C.S.R. Deep learning-based improved snapshot ensemble technique for COVID-19 chest X-ray classification. *Appl. Intell.* **2021**, *51*, 3104–3120.
27. Naz, S.; Phan, K.T.; Chen, Y.P.P. A comprehensive review of federated learning for COVID-19 detection. *Int. J. Intell. Syst.* **2022**, *37*, 2371–2392. [CrossRef]
28. Durga, R.; Poovammal, E. FLED-Block: Federated Learning Ensembled Deep Learning Blockchain Model for COVID-19 Prediction. *Front. Public Health* **2022**, *10*, 892499. [CrossRef]
29. Schwab, P.; Schütte, A.D.; Dietz, B.; Bauer, S. Clinical predictive models for COVID-19: Systematic study. *J. Med. Internet Res.* **2020**, *22*, e21439. [CrossRef]
30. Vaid, S.; Kalantar, R.; Bhandari, M. Deep learning COVID-19 detection bias: Accuracy through artificial intelligence. *Int. Orthop.* **2020**, *44*, 1539–1542. [CrossRef]
31. Bustos, A.; Pertusa, A.; Salinas, J.M.; De La Iglesia-Vaya, M. Padchest: A large chest x-ray image dataset with multi-label annotated reports. *Med. Image Anal.* **2020**, *66*, 101797. [CrossRef]
32. Available online: <https://github.com/ml-workgroup/covid-19-image-repository> (accessed on 1 December 2022).
33. Available online: <https://github.com/ieee8023/covid-chestxray-dataset> (accessed on 1 December 2022).
34. COVID-19 Radiography Database, 2022. Available online: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database> (accessed on 1 December 2022).
35. COVID-19 Radiography Database, 2020. Available online: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database/versions/1> (accessed on 1 December 2022).
36. RSNA Pneumonia Detection EDA, 2024. Available online: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data> (accessed on 1 December 2022).
37. RSNA Pneumonia Detection Challenge. Available online: <https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pneumonia-detection-challenge-2018> (accessed on 1 December 2022).
38. CRX-Pneumonia. Available online: <https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia> (accessed on 1 December 2022).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.