Healthcare Analytics 3 (2023) 100135

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

Blockchain for medical collaboration: A federated learning-based approach for multi-class respiratory disease classification

Abdulla All Noman, Mustafizur Rahaman, Tahmid Hasan Pranto, Rashedur M. Rahman *

Department of Electrical & Computer engineering, North South University, Dhaka 1229, Bangladesh

ARTICLE INFO

Keywords:
Blockchain
Collaborative approach
Deep neural networks
Federated learning
Incentive mechanism
Respiratory disease
Weight manipulation technique

ABSTRACT

The scarcity and diversity of medical data have made it challenging to build an accurate global classification model in the healthcare sector. The prime reason is privacy concerns and legal obstacles which limit data-sharing scope among institutions in healthcare. On the other hand, data from a single source is hardly sufficient to develop a universal diagnosis model. While federated learning is a potential solution to privacy and data diversity concerns (allows distributed model training), an apt aggregation process for multi-class and heterogeneous medical data is still at the outset. This study aims to propose a federated learning mechanism that can effectively learn from multi-class and heterogeneous respiratory medical data. The proposed system trains and aggregates the local model by leveraging blockchain technology, ensuring privacy. While aggregating the local models, we introduced the weight manipulation technique that, unlike any other studies, uses the local model test accuracy as the principal parameter. The resulting metric scores show that learning from diverse and heterogeneous data, the performance of the proposed federated model is analogous to a single-source model (learning from single source data). Using the novel aggregation technique, the highest testing accuracy of 88.10% has been achieved for five classes, compared to the less complex single source model, which achieved 88.60% testing accuracy. A similar trend has been observed for models with three and four classes. For developing better synergy among organizations, this study introduces an incentive mechanism for the contributing institution while the blockchain stores the records to make the system transparent and trustworthy. The proposed mechanism has been implemented using a web system, which demonstrates how the weight manipulation technique can effectively learn from heterogeneous and multi-sourced data while preserving privacy.

1. Introduction

Respiratory diseases such as pneumonia, tuberculosis, and Covid-19 are often caused by highly contagious viruses, bacteria, and pathogens [1,2]. According to the World Health Organization (WHO), around one million individuals were infected by covid-19 in the United States within one week during the recent pandemic [3]. On the other hand, about ten million individuals will be infected with tuberculosis globally by 2020 [4]. A globally trusted and easily accessible machine learning model deployed across hospitals may aid in the detection of contiguous respiratory infections and prevent rapid spread. Unfortunately, the institutions' reluctance to share data due to privacy concerns [5], the high diversity of these data, and heterogeneity make it challenging to construct a global model, resulting in an expensive and time-consuming diagnostic procedure [6–8]. Although deep learning (DL) based medical chest X-ray image diagnosis (also known as classification/detection) models for respiratory diseases have demonstrated good accuracy and cost-effective solutions for medical diagnosis [9] yet, due to the lack of collaboration between medical institutions, research in this field are limited.

Although large institutions might construct models based on single-centered data (data at possession), their performance usually reveals a bias towards the data-dominated classes. It often fails to perform with data from different sources. Therefore, this type of single institutional instance could be merely presented as a global model [8]. On the contrary, data from various institutions might originate from distinct probability distributions in terms of numbers and identical perspectives, which are not independent and identically distributed (non-IID) [10]. For example, data from the exact geo-location are likely to be correlated to each other. If data has not been taken from sufficiently enough random order of time, it violates the independent property of data. Also, data from different regions could show biases towards the distribution of classes, disrupting the identical property of data. As a result of the non-IID property of medical image data, traditional distributed learning algorithms provide poor performance in medical image classification [11].

* Corresponding author.
E-mail address: rashedur@northsouth.edu (R.M. Rahman).

https://doi.org/10.1016/j.health.2023.100135
Received 5 September 2022; Received in revised form 14 November 2022; Accepted 9 January 2023

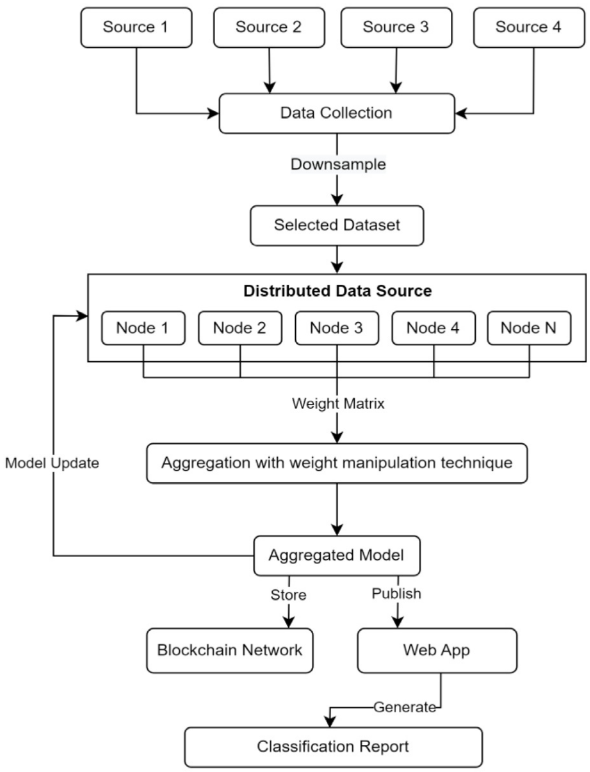

Fig. 1. Flow diagram of the overall system.

In the traditional distributed learning process, sensitive medical imagery data from different sources are collected to train a central model, which could trigger data privacy and security issues [12]. Therefore, a decentralized learning model is required to protect user privacy and other organizational attributes. On the other hand, the learning mechanism must be robust with heterogeneous (non-IID) data sources [13]. Federated Learning (FL) offers a learning approach that can learn from heterogeneous, decentralized data by aggregating the feature information from distributed edge nodes [14]. FL enables several institutions to pool collective knowledge and learn from one another's data without compromising the individual's data privacy [15,16]. Cetinkaya et al. demonstrated empirical evidence of the applicability of Federated Learning for multi-class respiratory disease detection from medical chest X-ray imaging of various institutional data in their study [17]. Although their work [17] demonstrates state-of-the-art performance on a concatenated dataset from several sources, it does not address the case of central server (known as the aggregator) failure or the server being compromised. There would be discontent from the contributing institution because all the prior model updated information is not kept.

Blockchain has recently become a key technology for improving system security, reliability and transparency [18]. The information within the blocks is distributed and decentralized, which is accessible and easily retrievable. One major drawback of the central aggregator of the federated learning technique is known as the single-point failures [19], which refer to the ultimate dependency on the central server. However, Federated Learning combined with blockchain can solve this problem effectively. With this advantage of both FL and blockchain, Kumar et al. proposed a medical Chest X-ray based diagnosis system for Covid-19 [20]. Their work [20] suggests blockchain secures the aggregator function and keeps the previous model information transparent within the blockchain. However, their work only shows evidence for a single-class (Covid-19) classification scenario, which does not advocate for multi-class cases in the real world. Another aspect being added is acute data heterogeneity (Also known as non-IID) in multi-class instances, which is infrequent in the case of binary classification and poses great difficulty for the model to learn. Qayyum et al. presented a DL-based covid-19 classification model that uses FL and blockchain [21]. Another principal aspect of the collaborative approach is ensuring the participation of the stakeholders. However, none of these studies [20,21] provides an incentive mechanism to attract the concerned stakeholders to provide their respective institution's data. This study proposes a global platform for medical institutions to collaboratively build a deep learning-based multi-class respiratory disease classification model utilizing FL and blockchain technology with an adaptive incentive mechanism for the participants.

This study focuses on answering these two fundamental research questions:

RQ1: How to handle data diversity, multiclass data, data heterogeneity, and model aggregation to train ML models using federated learning techniques?

RQ2: How to fairly incentivize the stakeholders based on their contribution to the system?

Research has already shown the viability of federated learning in medical use cases where the model learns from decentralized data sources with blockchain technology [26,27]. Most of the works in federated learning on respiratory disease diagnosis worked with binary classification problems like pneumonia, COVID-19, and some other studies have shown the potential of blockchain technology in the federated learning field [28].

In this study, to resemble a real-world scenario, we created a dataset (Fig. 1) by merging the dataset from various sources that are publicly available (Table 1). A deep neural network (DNN) architecture learns from multiple data sources using the federated learning technique. To aggregate local models, a modified FedAvg [14] algorithm is proposed with an external weight manipulation technique to deal with data heterogeneity (non-IID). This function manipulates the weights of the local models based on their performance on the local data. After each update, the learning history of the updated model is stored in the blockchain network. The models can be retrieved in case of single-point failure. An adaptive incentive system is introduced to entice the contributor to submit more accurate data. The incentive will change depending on the amount of shared data and a node's (institution's) model performance. Blockchain is used to increase transparency across the system and to create trust among organizations. Concretely, the objectives of this study can be summarized as follows:

- Developing a multi-class respiratory disease classification model by federated learning technique on digital chest X-ray imagery from a multi-center data source.
- Proposing a novel aggregation technique for the federated learning method called the "weight manipulation technique".
- Proposing and implementing a secure information-sharing medium in collaboration with federated learning and blockchain technology.
- Introducing an adaptive incentive mechanism inside the blockchain network to encourage the contributor to provide more authentic and diverse medical data.
- Presenting a web application to demonstrate the intended functionalities within the blockchain ecosystem.

The rest of the paper is organized as follows. The background study is presented in Section 2, which illustrates the previous knowledge on this topic and outlines the research trend. Section 3 presents our model architecture, dataset, and evaluation metrics. This section also describes the aggregation process of the local models and the incentive mechanism that follows the aggregation process. Section 4 presents the detailed architecture of our system; Section 5 advocates for our federated learning approach by carefully analyzing the result. Developing blockchain and web application environments are shown in Section 6. The study concludes in Section 8 with a discussion of the system implications in Section 7.

Table 1
Distribution of data in different classes.

Source	Class name	Number of images provided	Number of images selected for training	Number of images selected for testing	Total images selected
Rahman et al., [22] & Chowdhury et al., [23]	COVID-19	3616	800	200	1000
	Uninfected	10 192	800	200	1000
	Pneumonia	1340	800	200	1000
	Lung Opacity	6007	800	200	1000
Rahman et al., [24]	Uninfected	3500	0	0	1000
	Tuberculosis	700	541	123	
Jaeger et al., [25]	Uninfected	326	0	0	
	Tuberculosis	336	259	77	

2. Background study

2.1. Federated learning

Federated learning is a comparatively new technique that allows machine learning models to be developed using data from multiple participants while maintaining data privacy and security [29]. In the federated learning technique, multiple contributor nodes collaborate to share data among themselves to build one centralized model using the combined knowledge from these distributed data sources. Participating institutions shared their local parameters (weight) with a central model instead of sharing the actual data or attributes (to avoid privacy leakage), and the primary model combines all the parameters from different local nodes using the FedAvg [14] or its' derivative algorithm [30–32]. However, unlike a traditional training method on a single data source, federated learning has to deal with multiple data sources. Working with sophisticated sectors such as healthcare or finance comprises security and availability issues. For example, it can include edge devices like phones and IoT and make it trustable for high privacy demanding data like patient reports [13,33].

One of the significant concerns in distributed learning is the data's distinctive and non-identically distributed (Non-IID) nature. FedAvg [14] and several subsequent algorithms have shown promising performance in the case of non-IID data Fields [34], proving Federated Learning as a potential solution to other distributed fields. In the medical industry, the scarcity of data is another significant issue for machine learning applications because of privacy concerns and heterogeneous data distribution [35]. In this case, federated learning can aid in exchanging local data between hospitals without jeopardizing patients' privacy and construct a robust prediction model from a diversified and reasonably large distributed data source. A study by Rieke et al. described FL as a potential method for obtaining accurate, robust, safe, resilient, and impartial models [36,37] in the healthcare sector. However, even though Federated Learning has shown great prospects in several sectors, its implication in the healthcare industry still needs to be addressed.

2.2. Blockchain

Blockchain is a widely used technology that has emerged in recent years as a solution to various security challenges [38]. It is a collection of interconnected and distributed blocks containing data [18], where the first block is named the genesis block, and the other blocks are connected using cryptographic addresses. Bitcoin is one of the pioneer use cases of blockchain technology. Bitcoin was created in 2008 by Satoshi Nakamoto to address concerns such as double spending in the digital payment system [39]. In bitcoin, a tamperproof mechanism is implemented using a distributed consensus algorithm [39], in which most other users verify every transaction in the blockchain, and every block after the genesis block is added using a consensus algorithm with a cryptographic footprint and timestamp [40]. Everyone will have access to check the updated block information, ensuring that every transaction is transparent. Due to some unique properties, blockchain has shown enormous potential to be used in other fields besides digital payments [40,41]. A blockchain-based smart contract can replace any trust model that requires a third party. Ongoing blockchain research shows some potential use cases of blockchain technology in fields such as insurance [42] and banking [43].

Researchers worldwide are utilizing blockchain in a wide range of applications outside the digital payment industry [44–46]. For example, blockchain and smart contracts can be used in agriculture to track pre-harvest and post-harvest processing data using IoT devices, resulting in complete confidence and maximum profit for all parties involved, from farmers to consumers [47]. Li et al. [48] provided major work in file management systems with blockchain, demonstrating that the file loss rate can be reduced to a 0% rate which is exceptionally high in traditional cloud storage architecture. Nguyen et al. [49] demonstrated a considerably more secure deployment of mobile data for Electronic Health Record Systems (EHRs). So, blockchain implementation in real-world use cases shows tremendous future potential. According to cybersecurity firm Critical Insights research, cybersecurity breaches reached an all-time high in 2021, exposing a record quantity of patients' protected health information (PHI). Healthcare attacks harmed 45 million people in 2021, up from 34 million in 2020. The author examined breach data reported to the US Department of Health and Human Services (HHS) and showed that the number has tripled in only three years, up from 14 million in 2018 [50]. In recent years, blockchain has gained popularity in the healthcare industry as a secure medium for protecting patients' information [51].

As reported by IBM, on the one hand, 70% of healthcare executives believe blockchain will have the most significant effect in the health sector by improving clinical trial administration and, on the other hand, regulatory compliance and offering a decentralized foundation for exchanging electronic health records (HER) [52]. Furthermore, the worldwide blockchain technology market in healthcare is estimated to exceed $500 million by 2022 [53]. Zaabar et al. [54] demonstrated how crucial patient data could be kept in blockchain in a decentralized manner using a hashing technique. Chelladurai & Pandian also provided significant work in digitalized healthcare records (EHRs) [55] and utilized blockchain as a clinical data repository that offered patients a complete, distributed ledger record of all occurrences and seamless access to their electronic health records.

2.3. Federated learning with blockchain

In recent years, federated learning has emerged as one of the most popular techniques to build an aggregated model from multiple data owners without sharing sensitive raw data [56,57]. However, the intervention of a centralized accumulator in most data-sharing processes is particularly risky and often results in data leakage [28]. Geyer et al. presented a privacy-preserving federated optimization technique in which the client's contribution is hidden during the training process [58]. Although federated training methods ensure privacy, the two key challenges in federated learning are establishing trust amongst all parties involved in the training and aggregating large amounts of data from diverse sources, some of which may be new and unsuitable for other clients [28]. Blockchain is a potential solution to trust establishment, in which data is stored in a decentralized manner to establish confidence between all parties [28]. The collaboration between blockchain and federated machine learning approach shows great prospects in distributed machine learning techniques. Pokhrel and Choi showed a blockchain-based federated learning (BFL) for vehicular communication networks and a privacy-aware system in a decentralized training mechanism [59]. In the case of healthcare data, blockchain provides an additional degree of security, anonymity and transparency to prevent privacy leakage, which is a particular concern for client's shared medical information [57].

Machine learning has shown promising success for respiratory disease classification using CT scans and chest X-ray data [60]. Diseases like Pneumonia, COVID-19, and Tuberculosis can be detected by analyzing a patient's chest X-ray within a short time with trustable accuracy using a deep learning technique [60]. Because the medical data of the patients is confidential, the researcher experiences a data shortage to develop a more usable model. Blockchain-based federated learning can aid the secure collection of data to train reliable models without sharing it with others. Kumar et al. [20] illustrated an approach for detecting COVID-19 in which the model learns from distinct nodes (Hospitals) without sharing their raw data. In their work [20], a central model is sent to the edge nodes, where the model is trained by the data in the local node and the updated parameter is passed to the next node; the distance between those two nodes is considered the accepted measure of the data. All new weights are stored in the blockchain, so any node may view them and use them for training or testing on their local data without tampering with any existing nodes. Qayyum et al. [21] showed similar work to detect COVID-19 using chest X-ray and Ultrasound data. They emphasized the necessity of blockchain with federated learning due to the divergence posed by varying data distribution from different sources. However, only a limited number of studies in this field have worked with multi-class classification of respiratory diseases. As a result, a blockchain-based federated learning model for multi-class respiratory disease diagnosis could open up new possibilities.

3. Methodology

In this section, we will discuss the dataset, model architecture, aggregation technique, and evaluation metrics of our study. We have collected the data from the different publicly available datasets. Later, we downsampled and distributed the data among several nodes to mimic the real-world scenario. After a model is trained locally, weight matrices from individual local nodes are combined by a novel aggregation technique that is introduced in this study. The weight matrix of the aggregated model is stored in the blockchain as the cryptographic hash. At the same time, the model is also updated in our web platform for the users. The next training iteration in local nodes will commence using the latest aggregated weight matrix. In Fig. 1, we have shown our overall system's flow diagram.

3.1. Dataset

Due to the scarcity of medical datasets in the public domain, an aggregated dataset was created by combining data from various sources. Fig. 3 depicts our process of dataset accumulation and formation of the train and test sets from the aggregated dataset. We collected our initial data from four primary sources [22–25]. The data were then downsampled to remove biases. The downsampled dataset is the final aggregated dataset which contains 5000 X-ray image data (1000 in each class) of five respiratory diseases: COVID-19, Pneumonia, Tuberculosis, Lung Opacity, and Uninfected cases. We collected data on COVID-19, Viral Pneumonia, Uninfected, and Lung Opacity from the study by Rahman et al. [22] and Chowdhury et al. [23]. [22] created a large dataset by collecting data from different sources. To build a better-performing segmentation model, they used enhancement techniques, transfer learning, and lung segmentation to detect COVID-19 [22]. Most of their non-COVID data, such as pneumonia, uninfected, and lung abnormality, are collected from the pneumonia detection challenge dataset by the Radiological Society of North America (RSNA) [61]. Most of the COVID-19 data of Rahman et al. [22] are collected from Valencian Region Medical Image Bank (BIMCV) [62]. To evaluate the applicability of artificial intelligence (AI) in the rapid and accurate identification of COVID-19, Chowdhury et al. created a new chest X-ray image dataset which consists of 423 COVID-19, 1485 viral pneumonia, and 1579 uninfected chest X-ray images. Chowdhury et al. [23] also used RSNA [61] data bank as the source of data. On the other hand, datasets of tuberculosis and viral pneumonia are limited compared to the availability of the other three classes [23]. Data on tuberculosis are extracted from two separate datasets presented in the study of Rahman et al. [24] and Jaeger et al. [25]. Another study by Rahman et al. [22] used the public dataset of chest X-ray images of tuberculosis patients to precisely diagnose TB utilizing a combination of image pre-processing, data augmentation, image segmentation, and classification using a deep learning technique. Nine distinct deep CNN architectures were trained, validated, and tested for classifying TB and non-TB cases [24]. Rahman et al. (2020) collected tuberculosis data from 3 different sources such as the National Library of Medicine (NLM) [25], the National Institute of Allergy and Infectious Diseases Belarus [24], and the RSNA CXR dataset [61,63]. Further data on tuberculosis was collected from the study of Jaeger et al. [25]. This dataset has 662 frontal chest X-rays, of which 326 images are of uninfected cases and 336 are cases with indications of TB [25]. These X-rays were collected from The Third People's Hospital of Shenzhen. It also contains 138 frontal chest X-ray images accumulated from Montgomery County's Tuberculosis screening program, of which 80 are uninfected cases, and 58 are cases with indications of TB [25]. Fig. 2 depicts sample images of all the classes from our aggregated dataset.

From these four sources [22–25], a total of 26 017 images were initially collected, among which the highest number of images (14 018) belonged to the uninfected class, and the lowest (1036) belonged to the tuberculosis class. To avoid class biases in our final aggregated dataset, all the types were down-sampled to 1000 as the lowest data available was close to 1000 for the tuberculosis class. Finally, our dataset to be used in this study consisted of 5000 chest X-ray images and 1000 images belonging to each class. Of these 5000 images, 4000 are utilized during training (dataset D1 in Fig. 3), and 1000 are kept for testing both global and local model performances (global testing set in Fig. 3). Each image is reshaped into $150 * 150$ pixels and further converted into grayscale. This dataset is the aggregated final dataset used in our study to train, test, and validate our federated learning method. This aggregated dataset is diverse and represents the actual data distribution in the real world. The test set is global so that the participating actors can test their local models on this more diversified set. Table 1 shows the data distribution in different categories. In addition, Fig. 3 illustrates the collection, division, and distribution of our dataset that will be used in different stages within our system.

3.2. Model architecture

Federated Learning (FL) is a collaborative learning technique that can utilize different neural network architectures to learn from diverse data sources. CNN and DNN-based architectures are usually employed to train federated models collaboratively. The study demonstrates an out-and-out system design and future application of federated learning with blockchain in multi-class medical image classification [13,21,58,59]. We introduced a novel federated learning approach for improving the performance of deep classification architecture in the case of heterogeneous data (non-IID) sources. To make the model and the training process more straightforward, a DNN-based architecture has been used as the classification architecture.

To train the DNN model, input images are first converted into grayscale and then reshaped to an identical shape of $150 * 150$ pixels.

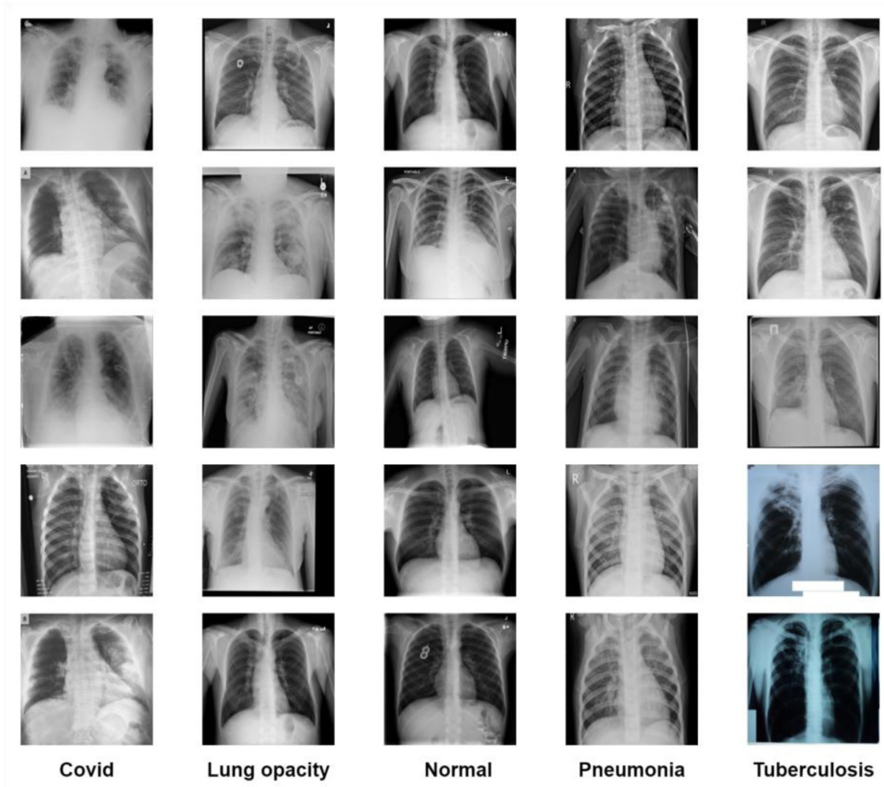

Fig. 2. Sample images of dataset.

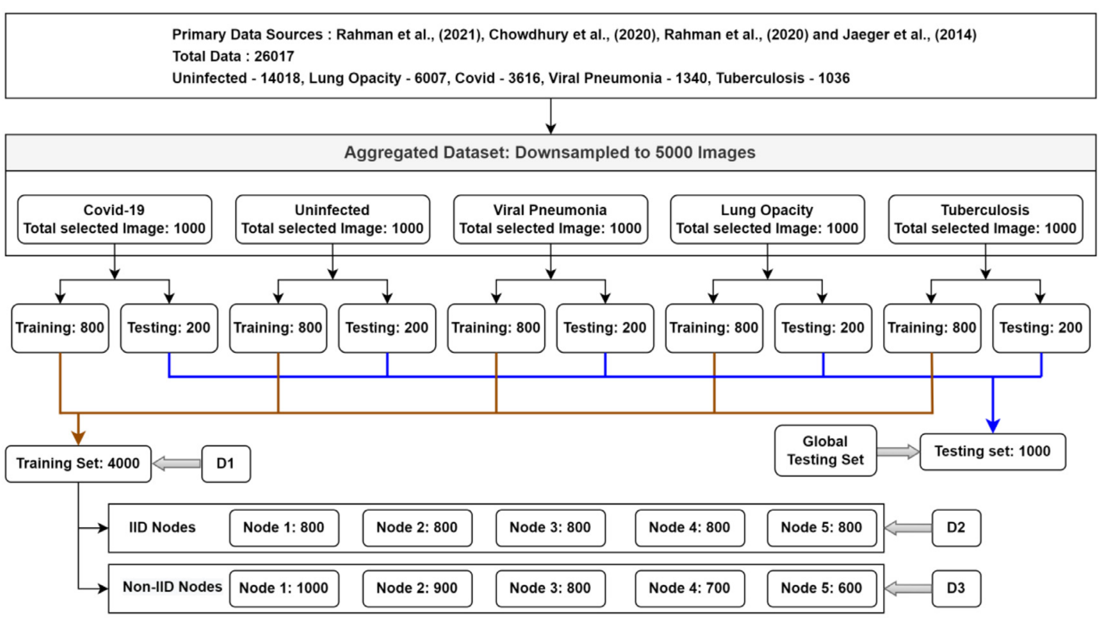

Fig. 3. Collection, selection, and distribution of dataset.

Afterwards, the images are flattened to a one-dimensional input, resulting in 22 500 input values for the input layer of the DNN architecture. Fig. 4 presents the network architecture along with the input, hidden, and output layers of our DNN model. The first, second, and third hidden layers consist of 2048, 512, and 128 neurons, respectively. The final hidden layer (fourth layer), composed of 32 neurons, is connected to the output layer. The output layer has five neurons as the architecture should learn to classify among five classes. All the layers are fully connected, and the output layer uses one hot encoding to predict the probable class of an image.

3.3. Weight manipulation technique

Finding IID data on each node is implausible in a real-world scenario. Data distribution may vary substantially across different contributors, and the underlying patterns of data from each node may

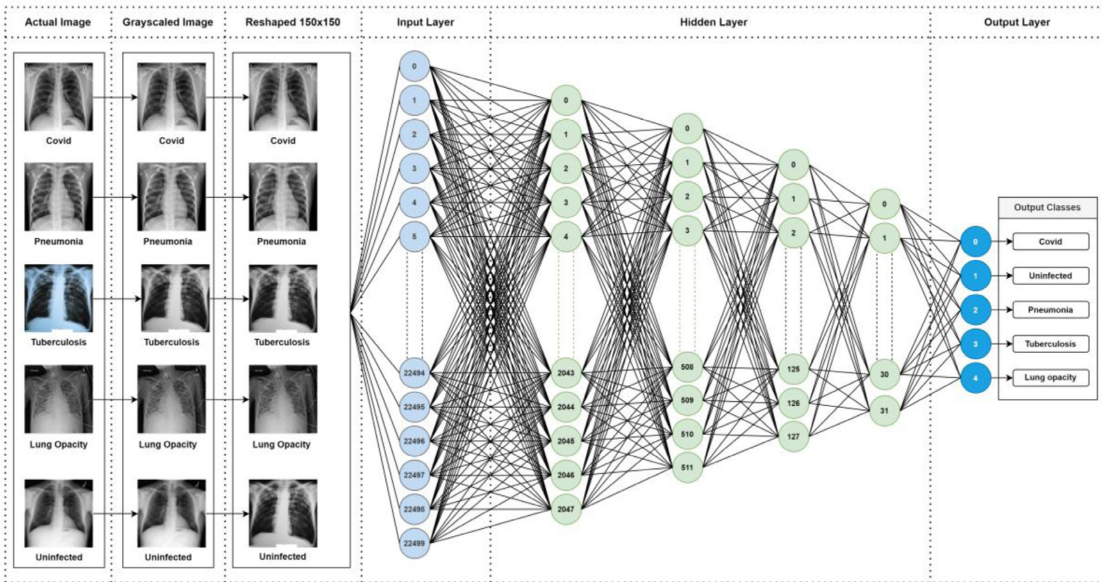

Fig. 4. Images pre-processing & model architecture.

Algorithm 1. Proposed modified FedAvg algorithm. The M number of nodes are indexed by m; B is the local minibatch size, E is the number of local epochs, η is the learning rate, n is the total number of data instances, and a is the cumulative test accuracy of all nodes on a global test set.

Initialize w_0 ;
For each round t = 0, 1,. . . . **do**
  For each node m = 0, 1,.....M-1 in parallel
    w_{t+1}^{i} NodeUpdate(m, w_t)
    $w_{t_n+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^{k}$
    $w_{t_a+1} \leftarrow \sum_{k=1}^{K} \frac{a_k}{a} w_{t+1}^{k}$
    $w_{t+1} \leftarrow \frac{(w_{t_n+1})+(w_{t_a+1})}{2}$
NodeUpdate(m, w) : //Run on node m
For each local epoch i from 0 to E-1 **do**
  For each minibatch of b of size B **do**
    $w \leftarrow w - \eta \Delta l(w; b)$
return w

differ significantly. The traditional aggregator functions used in various studies consider the “amount of data” as one of the most significant contributors to the aggregate process [14,64]. If we only consider the “amount of data” possessed by a node to aggregate the local models, the aggregated model may not obtain the optimum efficiency because of data heterogeneity. Our objective for this study is to achieve the best performance of the aggregated model from the limited and varied data across different data providers. In light of this, we attempted to develop a technique in which the model aggregation process is dependent not only on the quantity of data each node has but also on how well the local model captures the underlying pattern of these data. Since test accuracy indicates the model’s capability on unseen data, we also prioritize test accuracy as a parameter to the aggregator function. Since both our local and global models have the same architecture, we can assume that if the local model improves its accuracy on a subset of data, that will also affect the global model.

We modified the FedAvg [14] algorithm (Algorithm 1) to consider the amount of data and the local test accuracy of a particular model during the federated aggregation. The modification is based on the node’s weight matrix and will be triggered whenever the aggregation function attempts to combine the nodes through the aggregation process. Eq. (1) illustrates how the manipulation function is established based on different variables.

$$m = \frac{\alpha_n + \beta_n}{2} \tag{1}$$

$$where,\ \alpha_n = \frac{D_n}{\sum_{i=1}^{n} D_i},\ and\ \beta_n = \frac{T_n}{\sum_{i=1}^{n} T_i}$$

Here, α_n and β_n are the manipulation variable for training data and test accuracy. D_n and T_n denote the amount of data and the test accuracy of the nth node. Here, $\sum_{i=1}^{n} D_i$ calculates the total number of images used across local devices, and, $\sum_{i=1}^{n} T_i$ calculates the cumulative test accuracy of those corresponding nodes during the federated learning process. The value of α_n and, β_n will be multiplied each time with the weight matrix of the local model. The aggregated model will be constructed in the federated learning process by combining several local models. Combining these local models involves a unifying function that merges the weights from local models to construct the aggregated model. Finally, the weight matrix of the aggregated model will be formed by taking the statistical mean of the manipulated weights based on “data used” and “test accuracy”. This process is repeated throughout the federated learning process, where a local model that is trained with more data and has higher testing accuracy will contribute more to the aggregated model and vice versa.

3.4. Incentive mechanism

A successful contribution of a new model trained with new image data from the contributor will automatically add an incentive to their balance. The data provided by the contributor will be stored in the blockchain, and the incentive will be determined using the formula below, depending on the model’s data.

In Eq. (2), NI represents the number of images used in the model. There will be a fixed price for each image. In our formula, we assigned

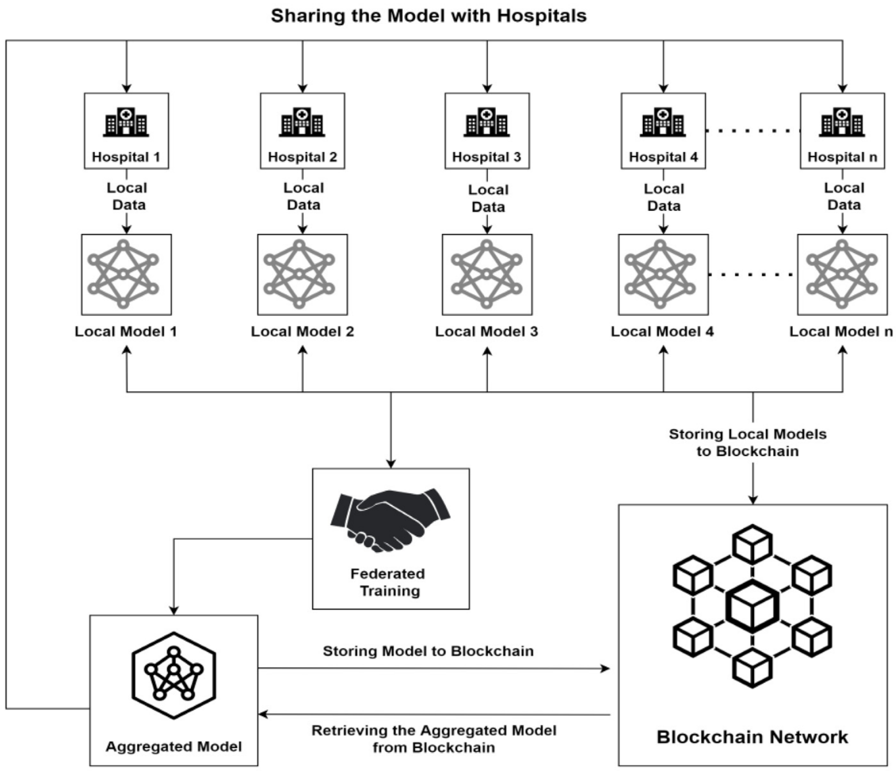

Fig. 5. Overview of the system.

each image a price of 0.001 units. The notations R, P, and F1 are abbreviations for recall, precision, and f1 score, respectively. As we have five classes for the classification task, we will get a total of five values for all the metrics (precision, recall, and F1 score). The summation function will add the five values, and the average will be taken.

$$I = \frac{NI}{1000} + \frac{\left(\frac{2\sum_{i=1}^{n} R_i}{n} + \frac{\sum_{i=1}^{n} F_i^1}{n}\right)^2}{1 + \left(1 - \frac{\sum_{i=1}^{n} P_i}{n}\right)} \tag{2}$$

In medical use cases, the recall value plays a significant role in correctly identifying the true positive cases and generating fewer false negatives. Hence, extra emphasis on recall has been given by multiplying it by an integer value of two. In the denominator, the average precision value will be subtracted from integer value one (the maximum of metrics can be one). So, the more precision is generated, the smaller the denominator becomes. Consequently, as the denominator decreases, the incentive will increase proportionally to a much more significant number. The opposite case is for the lowest recall and precision values, where the incentive will be reduced immensely. Therefore, the total incentives rely on the number of images used by a model and its accuracy, f1, and recall values on a global test set.

3.5. Evaluation metrics

Precision and recall are widely used metrics in medical cases [65]. False-positive and false-negative are treacherous as the medical sector requires absolute certainty when reaching a decision. For multi-class classification for a particular class λ, the following evaluation metrics are used:

$$Precision_\lambda = \frac{TP_\lambda}{TP_\lambda + FP_\lambda} \tag{3}$$

$$Recall_\lambda = \frac{TP_\lambda}{TP_\lambda + FN_\lambda} \tag{4}$$

$$F1 - Score_\lambda = \frac{2 \times Precision_\lambda \times Recall_\lambda}{Precision_\lambda + Recall_\lambda} \tag{5}$$

TP, FP, and FN represent True Positive, False Positive, and False Negative, respectively. Based on Eqs. (3) and (4), it is clear that high FP and FN values will be represented by lower precision and recall values. The trade-off between precision and recall exhibits the true performance of a classification algorithm. As TP increases, TN values will decrease. Thus, precision and recall are inversely related to each other. Therefore, a balance between precision and recall must be found to preserve the accuracy of predictions. As a result, F1-Score is also used as an evaluation measure to check optimization between Precision and Recall.

4. System design

The collaborative approach of medical institutions worldwide to build a global AI model to classify respiratory diseases seems impractical as the personal data vulnerability concern is related to this. Due to privacy concerns and the organization's reputation, the records are rarely made public. A global AI model with a firm grasp of global data may be able to resolve some severe difficulties with efficiency, time, and cost. However, considering the fundamental problems of transparency and privacy, it is challenging to train any global model without sharing data. As data sharing in the FL system is insignificant, a Federated learning-assisted system could be a promising solution to the abovementioned problem.

Our system implements a federated learning and blockchain-based platform where the institutions can share their information (model) without disclosing the actual data. The complete model's information will also be stored on the blockchain network to create a firm trust boundary between the institutions. Fig. 5 depicts the overview of our system, where it has been seen that each institution or hospital can share its trained model to participate in building a decentralized global model. The institution's model information and the final aggregated model information will be stored in blockchain through a web application. As a result, other institutions may have a better overview of these shared and global models. It will help them to trust one another. On top of that, blockchain is an effective system to keep the records immutable, tamperproof, and transparent, making the whole process more trustworthy to the stakeholders.

We proposed a deep neural architecture (DNN) for training the local models in the federated learning process. The DNN model comprises four hidden layers, one input layer, and one output layer. The dataset we collected has been categorized into five classes (COVID-19, Regular, Viral Pneumonia, Lung Opacity, and Tuberculosis). From a dataset of 5000 images constructed using data from various sources (Described in Table 1), each class contains 1000 images individually. Among those 1000 images, 800 images are used for training and 200 for testing. A universal test set was created by combining those 200 test images from five classes, and the rest of the 800 images (Described in Fig. 2) were selected for training the centralized (model trained in a single center with all data) and local models (Model trained locally during federated session). We repeatedly trained and evaluated a model for the initial model parameters and architecture selection by modifying parameters like image reshapes, learning rates, momentum, number of neurons in the hidden layers, etc., and then noted the overall performance. Finally, we used the best hyperparameters combination for the centralized model training. The detailed system design, roles of the other actors connected to our system, and a comprehensive workflow of our proposed approach have been presented in the following subsections.

4.1. System actors

Three different actors are involved within our system: Regulator, Contributor, and User. The regulators initiate the system with required actions, such as initial training and uploading the aggregated model to our system. In the later phase, the aggregator function will be invoked from a contributor node. They are also responsible for managing the federated learning process across the contributing nodes (Off-chain) and uploading the latest aggregated model in the web application by fulfilling some additional information.

The second type of actors is the contributors, who are the hospitals or organizations that will contribute to the federated learning process by providing the latest and original data from their different institutions. The incentive provided to a contributor node will be provided upon successfully uploading a local model and its report to the web application. Based on the information, the model contributing the most to the aggregated model will get the highest amount of incentive and vice versa. The last and final type of actor in our system is the users. Any hospital or institution not contributing to the system can register as a user and use the updated aggregated model to instantly diagnose and generate a chest X-ray image classification report using our web application.

4.2. Overview of the proposed architecture

The system architecture comprises four different components: actors, DNN model/federated learning, web application, and finally, the blockchain network. Section 4.1 explains the role of different actors across our system. Initially, the contributor node will train the model on their respective devices using the initial aggregated dataset. After the training procedure, the contributors will upload the result and metrics of the trained model into the blockchain through the web application. After uploading, the model weight matrix will be converted into a cryptographic hash and stored on the blockchain. An incentive mechanism is designed so that the contributor node that will provide the model with excellent performance and comparatively a more significant number of images to the system will get more rewards compared to the models trained with a limited number of pictures and have low performance. When several contributors submit their models into the blockchain network through the web application, the federated learning process will be evoked by the regulator to aggregate those models with the previous aggregated model. A modified FedAvg [14] algorithm will be used to aggregate the local models. Unlike FedAvg [14], our modified version takes the testing accuracy of the local models into account. This modification has been introduced to ensure that the global aggregated model's performance is not substantially hindered by a local model with an unfavorable performance. Due to the distributed training (FL) process with weight manipulation, the contributor's model with fewer data and worse parameter values would contribute less to the aggregated model compared to the model with more data and better parameters (discussed in Section 3.3).

Afterwards, the aggregated model will be uploaded from a regulator node in the web application with the metric evaluation details. The model's parameters will be stored in the blockchain network for further usage by the users. And the aggregated model will be replaced as the latest model within the system. Suppose a user attempts to get a diagnosis result via the web application by uploading a chest X-ray image. In that case, the web application will use the most recent stored model to generate the diagnosis report.

Moreover, the weight matrix of the current aggregated model will be converted into a cryptographic hash. It will be stored in the blockchain network along with the model's details for better transparency and stakeholder trust. Any institution or hospital worldwide logged in as a user can upload chest X-ray images and check the disease in real-time using the latest aggregated model. Fig. 6 describes the entire system architecture, where different components of the architecture and the connection between them have been vividly portrayed. The worldwide contributors can contribute to the aggregated model by providing new and unique data. The aggregation is a continuous and incremental process of improving the existing aggregated model. When a certain number of models are credited to the system, the aggregation process will be evoked from a regulator. After the completion of the aggregation process, the updated aggregated model will be deployed to our system. The contributing nodes will be rewarded with incentives based on improved performance on the global test set. Repeating the process with different data variations, the model becomes versatile and accurate in classifying chest diseases.

5. Result & analysis

The federated learning starts with distributing the dataset for five different classes into five IID (D2 in Fig. 2) and five non-IID (D3 in Fig. 2) nodes. As mentioned in Section 3.1, after choosing 800 images for training and 200 images for testing for each class, we first experimented with the model training with variable classes (two, three, four & five classes) centrally using the architecture described in Section 4.2. Different sets of experiments were carried out to demonstrate our weight manipulation technique is adaptable to data heterogeneity even with varying classes.

First, the centralized model was trained using dataset D1 (Fig. 2). It solidifies the classification capability of our model architecture for different classes; hence, the model is ideal for our following experiments with federated learning. The models were trained up to 120 epochs. However, the performance of the two-class model (non-infected and pneumonia) was relatively stable after five epochs. To prevent model overfitting, we terminated training the model after thirty epochs. Fig. 7 depicts the loss vs. epoch graph, accuracy vs. epoch graph, and confusion matrix for different classes of the centrally trained five-class, four-class, three-class, and two-class models. Fig. 7 shows that after around 60 epochs, the performance of the five-class, four-class, and three-class models was relatively steady with slight variations, and the test accuracy and loss did not significantly improve. The best test

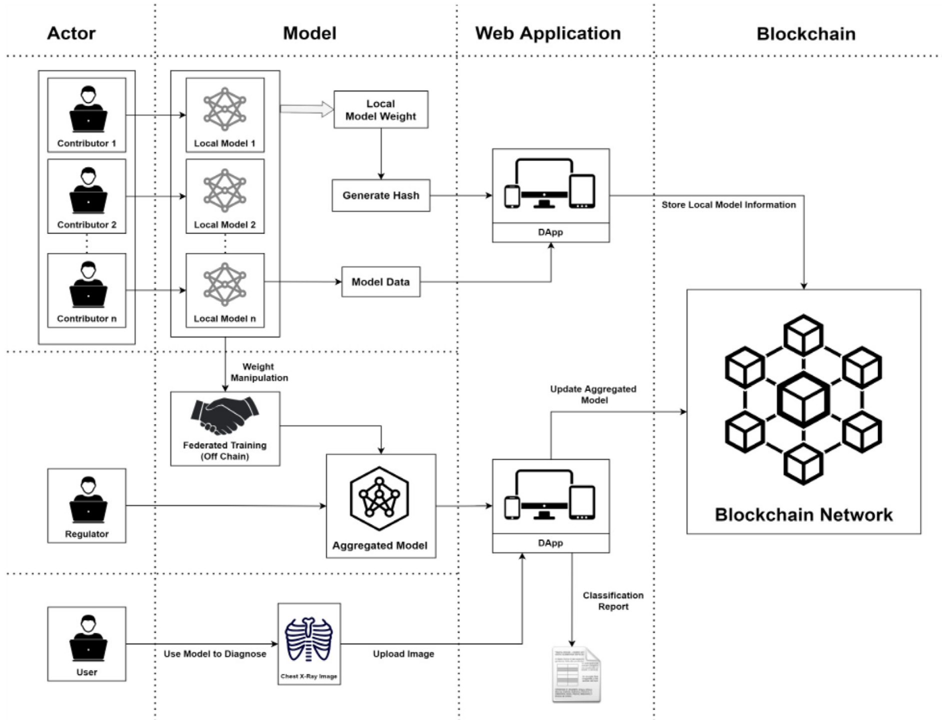

Fig. 6. Detailed system architecture.

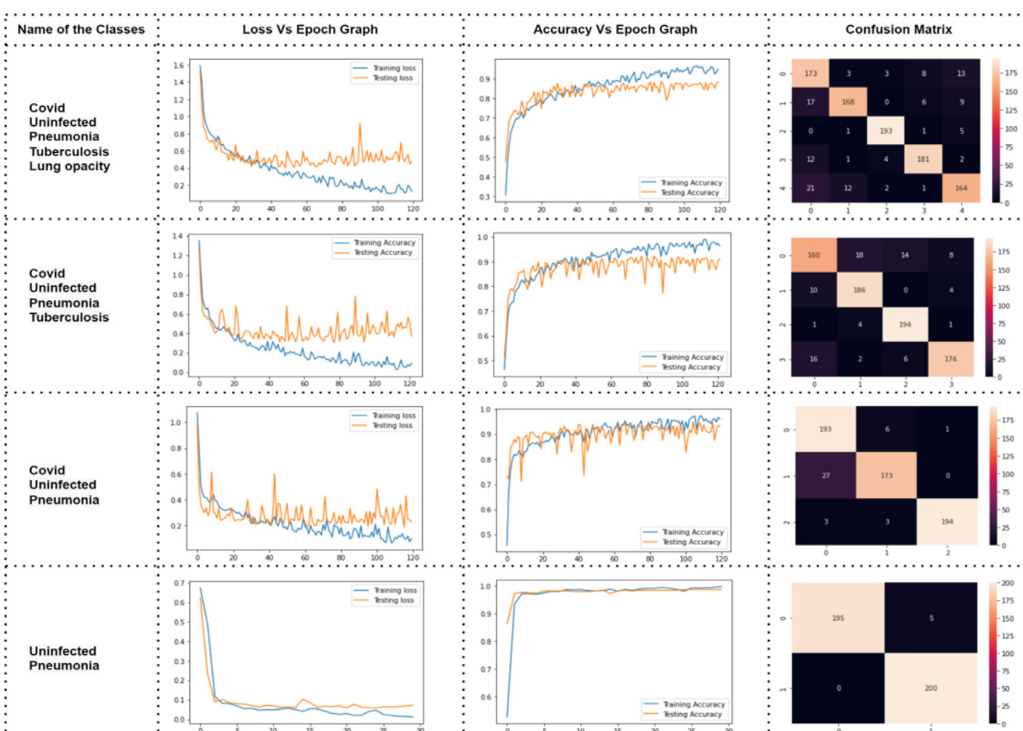

Fig. 7. The result of centrally trained models for different classes.

accuracy on the global test set for five classes was 88.60%, for four-class 91.13%, for three-class 95.17%, and for two-class 98.75%, respectively.

After observing the centrally trained model's performance on the D1 dataset, we created two subsets of the dataset from the D1 dataset for simulating the IID and non-IID node scenarios, mimicking a real-world data distribution. Each group of datasets contains five sets of data, and each set is considered a node for federated learning. The first group includes the same number of images for each node (dataset D2). The second group contains a different number of images (dataset D3). The whole scenario is depicted in Table 2 and Fig. 2, where the first, second,

Fig. 8. Individual node's training result.

Table 2
Distribution of data on different nodes.

Node	Number of data for every class	
	Group 1 (All nodes with variable data)	Group 2 (All nodes with same data)
1	200	160
2	180	160
3	160	160
4	140	160
5	120	160

Table 3
The result of federated learning for different category.

Number of classes	Data on each node	Weigh manipulation function	Maximum test accuracy of aggregated model
3	Same	✘	93.67%
	Different	✘	93.50%
	Different	✔	94.00%
4	Same	✘	90.62%
	Different	✘	90.50%
	Different	✔	91.63%
5	Same	✘	87.80%
	Different	✘	87.60%
	Different	✔	88.10%

third, fourth, and fifth nodes of group 1 (D3 in Fig. 2) have 200, 180, 160, 140, and 120 images for each class. On the other hand, every node of group 2 (D2 in Fig. 2) has the same 160 images for each class.

Before the federated aggregation starts, each node locally trains the model with the local data D2 and D3 in Fig. 2. The reason behind training separately on each node with smaller data sets is to discover the data pattern on the individual node. Test accuracy of the individual node will be used in the weight manipulation function for the federated learning process. Fig. 8 depicts the details of the individual node's training result. From Fig. 8, it can be seen that as the number of data was limited on each node, the overall performance was comparatively poor. The test accuracy graph was unstable in both cases, and the test loss graph showed some noticeable fluctuations

We repeated the process for five-class, four-class, and three-class classification and discovered that the aggregated model performed better when we used the weight manipulation function. For three-class classification, the maximum test accuracy without weight manipulation on D2 and D3 datasets were 93.67% and 93.50%, respectively. However, with the weight manipulation function, the maximum test accuracy (94%) on dataset D3 was achieved. The maximum testing accuracy for four-class classification without employing the weight manipulation function on D2 and D3 was 90.62% and 90.50%, respectively. Using weight manipulation, the accuracy was increased to 91.63%. Finally, for the five-class classification, the overall test accuracy was slightly lower than the three- and four-class classifications. Still, the accuracy was higher with the weight manipulation function (88.10%) than in the other two categories (87.80% and 87.60%). The comprehensive data of all three cases have been depicted in Table 3.

Fig. 9 shows the test accuracy graph for different data categories (IID, non-IID, and non-IID with weight manipulation) and variable classes. From the graph, it can be seen that whenever the weight manipulation function was used, there was a tendency to have a head start on the test accuracies in the starting epochs. The average test accuracy after 120 epochs was also higher in this category compared to the other two types. Another critical outcome we encountered was that each node with the same data (IID) performed better than the nodes with variable data (non-IID) without using the weight manipulation technique. In the traditional federated aggregation process, the weighted average of the total data point among the nodes is taken to build the weight matrix of aggregated model. However, in practical use cases, different nodes may have different statistical distributions of data, and the local accuracy of those nodes can be a substantial contributing factor in building the aggregated model. Our non-IID dataset (D3) has variable data on each node, simulating a real-world data distribution. So, a mechanism was well required to tackle the abovementioned case on models trained using non-IID data with different local accuracies. Our study introduces the weight manipulation technique as a solution to this problem. Our experiments using the weight manipulation technique based on the number of images and individual test accuracy of local models showed better results in the case of non-IID aggregation. Since all three categories were evaluated on the same percentage of data, it has been experimentally established that the FedAvg [14] algorithm in assistance with weight manipulation function is a preferable choice in the case of multi-class classification on non-IID data.

One of the objectives of this study was to train a five-class classification model in a federated way. As we introduced weight manipulation into our research, it was essential to check whether this technique works appropriately for class variability. So, we trained three- and four-class classification models using the federated learning process with the help of the weight manipulation strategy. Every time, it performed better than the standard FedAvg [14] algorithm for three, four, and five-class classifications. Finally, based on the experiment described in this section, we selected the weight-manipulated five-class model as the final model to deploy into our system.

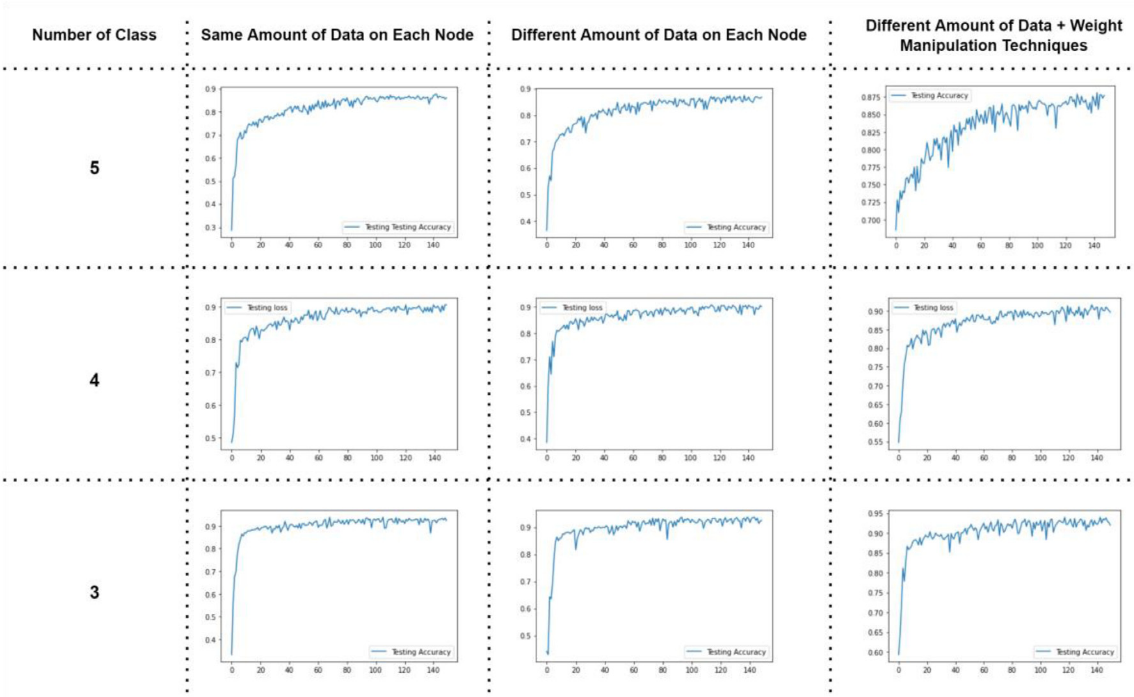

Fig. 9. Test accuracy graph for a different category (IID and non-IID and non-IID with weight manipulation).

6. Deploying blockchain & web application environment

This section discusses our system implementation in detail. To implement four different levels of our system, we used the windows ten operating system, Intel(R) Core(TM) i7-7700HQ CPU clocked at 2.80 GHz, four cores, eight logical processors, and a total of 8 GB of RAM. The test environment was coded using the VSCode IDE. The architecture, including the federated learning, blockchain, and experiment stages, was implemented using Python. The web application was developed using HTML5, CSS, and Bootstrap on top of the Python flask library. Python version 3.9.6 and flask version 2.0.2 were used for coding implementation.

6.1. Blockchain development

The blockchain was implemented as a python class called ‘Blockchain’, and it inherits another class called ‘Block’. The Block class takes block data, block number, block hash, preceding block hash, timestamp, and nonce as input to create a new block object. When a contributor contributes to the system, a new block is created with the data provided by the contributor. The data has been represented in JSON format throughout the system. The hashing function employs the imported SHA256 hash algorithm, which creates a 64-bit hash function using a block’s contents. The Blockchain class contains five functions to add, delete, mine, and validate a block and another function to check the current blockchain’s integrity and return the current chain. The function statements of ‘add’ and ‘mine’ are shown in Fig. 10.

6.2. Web application development

Initially, a regulator uploads the initial federated aggregated model in our system. The modified aggregation process proposed in this study, designed specifically for federated learning using heterogeneous medical data, has been chosen based on the experiments described in Section 5. Fig. 11(a) shows the model update procedure for a regulator. After uploading the aggregated model, the information and the hash created using the model’s weight matrix will be stored in the blockchain. The successful creation of a block by a regulator in the blockchain network is depicted in Fig. 11(b).

The contributors can exhibit their willingness to contribute to the final model at any time, but the regulators are responsible for initiating a federated learning session. To contribute to the model, contributors must upload their information to the web application through the contribution page. Fig. 12(a) shows the contribution page where the contributor has to fill in some information required to proceed. The information will be stored automatically in the blockchain and will be used to calculate the incentive described in Section 3.4. A successful contribution will create a block, and this block will later be added to the blockchain network. Fig. 12(b) depicts the successful creation of a block by the contributors.

The users will use the updated model for real-time classification. The model’s weight matrix is stored in the blockchain to ensure transparency across the system. After the federated-trained model has been deployed into the system, any hospital or institution with an active user account can use the latest aggregated model to generate classification results. When users upload an image, the web application will adjust it (resize, reshape) to fit into the model as per our initial training size ($150 * 150$, grayscale). The image uploaded for classification will not be stored in our database or shared with any other users in our system. After classification, the image will be discarded from the system, and only the result and metrics will be used in further steps ensuring privacy for the individual provider. Fig. 13 demonstrates the usage of the aggregated model to form classification results

Fig. 14 shows the dashboard for different actors. The contribution option is solely visible to the contributor, whereas the regulator can access the update model option. The federated learning-related operations (“Contribute a model” and “Update model”) are not visible to the user actors. All actors can use the “Diagnosis Chest X-ray” option to get a classification report anytime. The actors can also find the username, role, and amount deposited to their account in the dashboard. The history of how many models are aggregated and the total number of images used in those models can be seen from the dashboard. We aggregated five models in our federated learning session

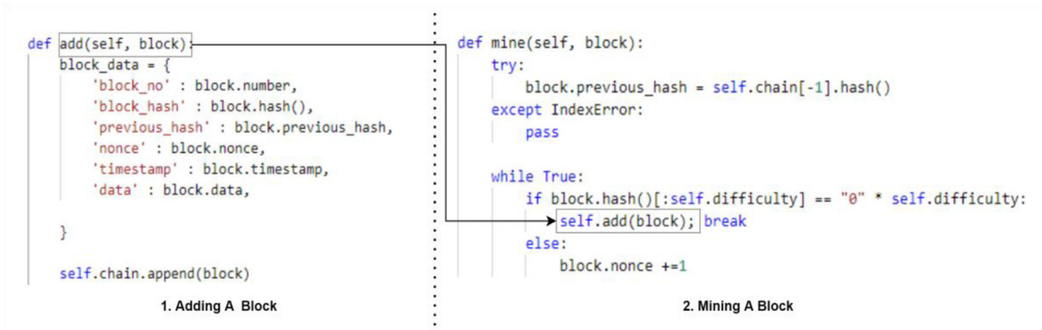

Fig. 10. Block adding & mining function.

Fig. 11. Model update page of the web application & successful creation of block for regulators.

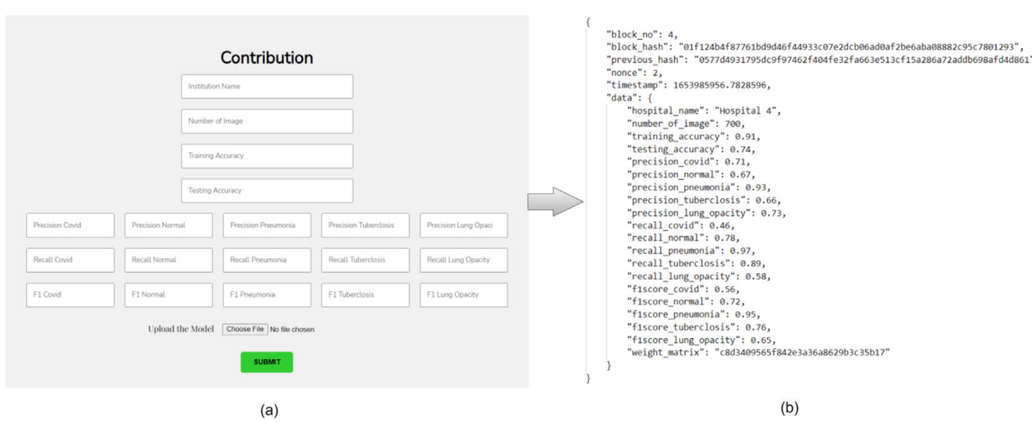

Fig. 12. Contribution page of the web application & successful creation of a block for contributors.

and used 4000 images on those five models, which can be seen in the dashboard. Algorithm 2 depicts an overall representation of the flow of our federated learning approach, from selecting the model to deploying the federated aggregated model in a web application and the further usage of the model.

7. Discussion

Deep Learning algorithms for medical image diagnosis have established their efficiency and reliability on a single data-centered model. To learn and perform, these deep learning models require a representation of real-world data. However, medical Data are highly confidential, and the leakage of medical data privacy may result in legal complications and damage the institution's reputation. Our proposed system will protect data privacy by establishing a platform where medical institutions cooperate securely without disclosing their patients' private information using the blockchain-enabled federated learning technique.

Federated learning has made it possible to gain collective knowledge for deep learning from various data sources without sharing data among institutions. The data is not shared with a central server; instead, it is used within a node, with only the weight matrix aggregated centrally. Our weight manipulation technique applied as an aggregated function yields a promising accuracy result while learning from various data sources. This research's primary objective is to improve the federated learning technique by employing a novel aggregation function (wight manipulation); it has been experimentally proven to work in the case of multi-class respiratory diseases using federated learning and blockchain.

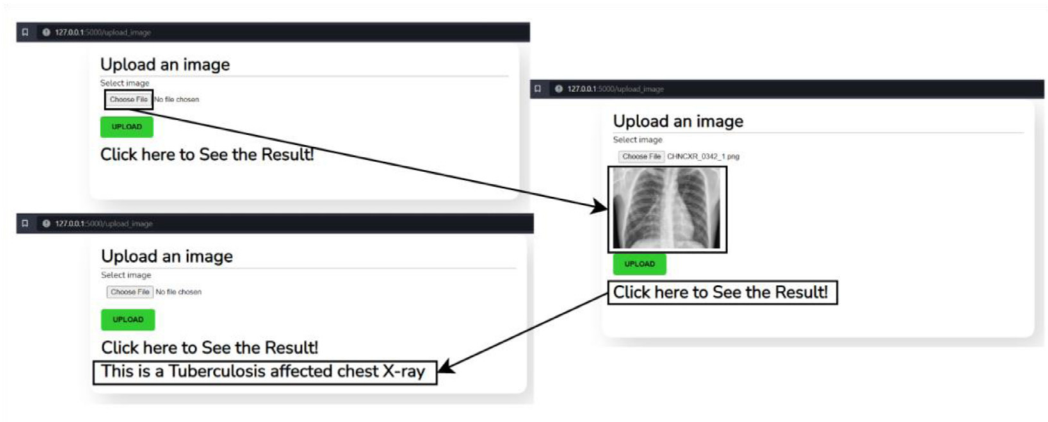

Fig. 13. Example of classification diagnosis.

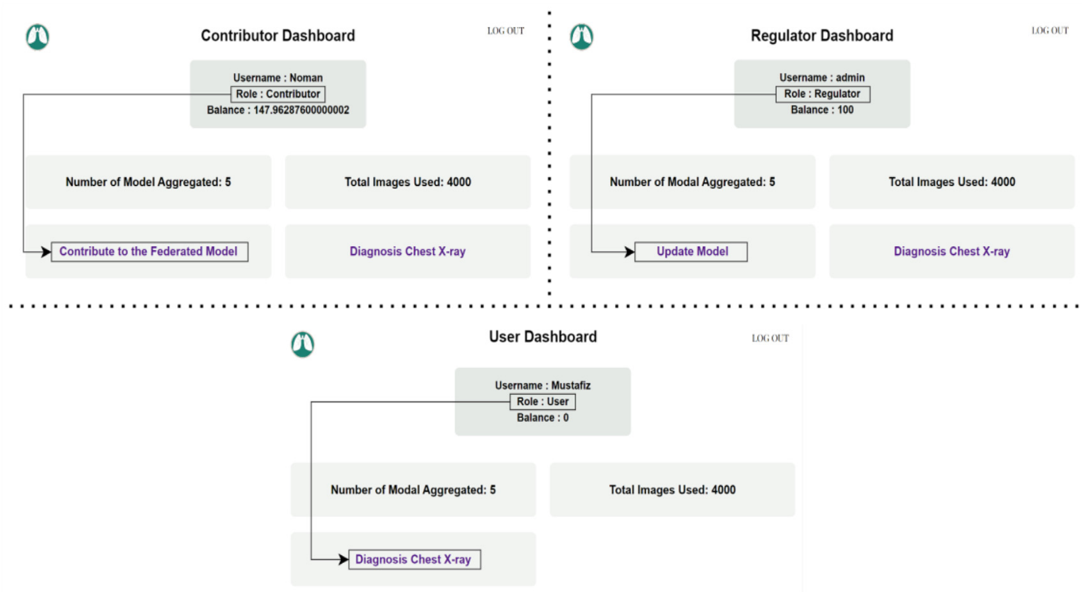

Fig. 14. System dashboard for different actors.

Table 4
Comparison of the proposed system and existing works.

Paper name	Number of classes	Working prototype	Weight manipulation	Fair incentivization	Dealing with heterogeneous data	Off-chain implementation	Blockchain integration
Yan et al., [66]	3	×	×	×	×	×	×
Qayyum et al., [21]	2	×	✔	×	×	×	×
Ślazyk et al., [67]	4	×	×	×	✔	×	×
Cetinkaya et al., [17]	4	×	✔	×	×	×	×
Kumar et al., [20]	2	×	✔	×	✔	✔	✔
Ho et al., [68]	3	×	✔	×	✔	×	×
Feki et al., [69]	2	×	×	×	✔	×	×
Our proposed method	5	✔	✔	✔	✔	✔	✔

Our method offered an incentive mechanism that has been particularly designed for medical use cases. In the case of blockchain, all prior updating weight and accuracy, which are essential for deep learning models, will be stored in the blockchain, which is immutable and accessible to all contributors at any time. Therefore, the stakeholders will gain confidence in the model's performance by witnessing the learning history of the model. Our proposed approach provides the platform for institutions to collaborate in the collective interest of building a universal model. Our architecture eliminates barriers to collaboration between medical institutions by preserving data privacy. Table 4 contrasts our proposed strategy with previously published studies on applying federated learning in the classification of medical chest diseases. The table shows that no other prior work has been published with a functional prototype and a fair incentivization policy for the contributors. Additionally, relatively few studies have included blockchain integration to enhance the trust boundary among the participants. Given all of these, the contribution of our research can be summed up as follows:

Algorithm 2: Pseudo Steps of the federated learning approach

Result:

Selecting the suitable DL model for central and federated training.

Building ten nodes where five of them follow IID characteristics and the rest five follow Non-IID.

Developing weight manipulation techniques to learn from non-IID data.

Building a global blockchain-based platform to update federated aggregated models from time to time.

Implementing an adaptive incentive mechanism to increase the synergy between contributing institutions.

01 **Pre-deploy DL experiment:**

The initial model and architecture have been selected by experiments with different parameters. The best model is selected based on its performance.

02 **Collection of the Dataset:**

The dataset has been chosen, which consists of merging different existing datasets. Initially, the dataset contains 5000 images, where 4000 images are selected for the training set and 1000 for the test set. The training set is further divided into smaller subsets to simulate the IID and Non-IID nodes.

03 **Pre-deploy Federated Learning Technique**

The federated aggregated model has been merged from both IID and Non-IID nodes. The Non-IID nodes with the weight manipulation function performed better. So, the model and architecture setting has been selected to deploy in the web application.

04 **Deploy the selected model:**

The information of those individual node's models (Non-IID Nodes) has been included in the blockchain as an initial contributing node, and the aggregated model from these nodes is also initially deployed into the system. The information of the initial aggregated model is also stored in the blockchain.

05 **Model Usage:**

Organizations worldwide can contribute to the federated aggregated model by registering through the web application. They will automatically receive incentives based on their model performance. The regulators will start federated learning among the nodes after some organizations show their willingness by registering in the web application. After the end of the training session, the previous aggregated model will be replaced by the newly aggregated model by the regulators. Any organization can use the model from the web application for diagnosis purposes or call the model's API to use it on the back end of their website.

- Introducing a global platform for medical institutions to improve a respiratory disease classification model collaboratively.
- Training multi-class respiratory disease classification model using federated learning.
- Introducing the weight manipulation technique as an aggregator that improves model learning by considering the test accuracy of local models.
- Increasing synergy among the medical institution to work together by providing privacy-preserving among them.
- Creating a new dataset by collaboration among medical organizations.
- Utilizing blockchain in the middle creates transparency and trust boundaries among the actors.

8. Conclusion

Classification models that are trained on single-source data often fail to perform in real-world cases where data comes from heterogeneous sources representing various distributions. To create robust models, authentic data from diversified sources is essential. However, sharing private medical data between medical institutions are hindered by privacy concerns and legal obligations imposed on medical institution. In this study, we presented a modified aggregation process for federated learning to build a universal multi-class classification model that learns from a heterogeneous medical data source. Our approach uses a novel aggregation technique, namely the "weight manipulation technique", to aggregate local models from distributed data providers to train a multi-class respiratory disease classification model. The aggregated models trained with the proposed weight manipulation technique performed approximately similarly to a less complex model trained on single-sourced data.

A cutting-edge blockchain platform has been developed to implement the privacy-preserving, immutable and transparent application that accommodates all our system requirements. The regulator node provides an initial baseline model, which is then locally trained across distributed contributor nodes (medical institutions). The local models are aggregated using the proposed aggregation procedure and merged into a single model deployed into our application to be used by the user nodes. Using the proposed aggregation procedure, the model became increasingly robust and precise to accurately classify respiratory diseases with a new batch of data provided by the contributor nodes.

None of the learning history stored in the blockchain can be changed due to immutability, ensuring absolute trust in the platform while sharing data containing sensitive personal information. An adaptive incentive mechanism incentivizes organizations to provide real-world data. As authentic data is more likely to influence a model's performance, providing accurate data is highly rewarded by our incentive mechanism. Our proposed architecture is generic for applications in similar settings in which data privacy barriers limit collaboration. The study builds a competitive universal model from non-IID diverse data using the federated learning technique, which employs the proposed weight manipulation technique. Contrary to the traditional aggregation method, the proposed weight manipulation technique uses the test accuracy of local nodes, which fences the central model from an overfitted local model, eventually ensuring that the central model only aggregates local models that actually learned from new data.

The proposed aggregation method does not compromise privacy, and medical institutions can collaboratively develop robust ML models for real-world data. The method ensures transparency, availability and immutability using blockchain, which will enhance the stakeholder's trust in the overall system. Apart from that, as no data is being shared with the system, participants can, without hesitancy, contribute their data, thus, creating synergy in industrial collaboration. Experimenting with our approach for different DL-based architectures is a future perspective of our work. Additionally, the proposed method can be utilized in other non-medical but sophisticated use cases like FinTech, IoT, and other industries.

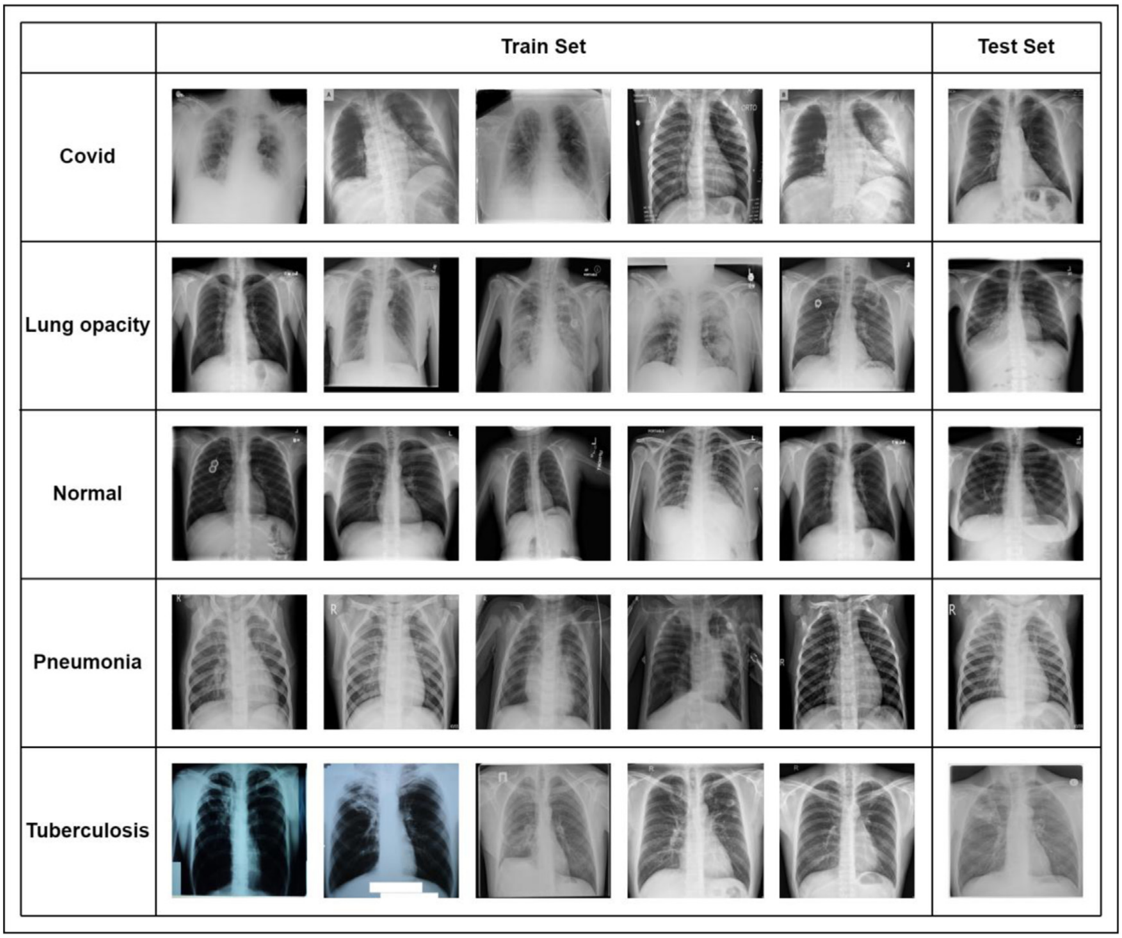

Fig. A.1. Figure of dataset containing 25 train and 5 test data.

Data availability

Data will be made available on request.

Acknowledgments

Supported by North South University, Faculty Research Grant, Grant/Award Number: CTRG-21-SEPS-19.

Appendix

See Fig. A.1.

References

[1] Respiratory diseases - genes and disease - NCBI bookshelf, 2022, https://www.ncbi.nlm.nih.gov/books/NBK22167/ (accessed Jun. 17, 2022).

[2] Respiratory disease | definition, causes, & major types | britannica, 2022, https://www.britannica.com/science/respiratory-disease (accessed Jun. 17, 2022).

[3] United States of America: WHO coronavirus disease (COVID-19) dashboard with vaccination data | WHO coronavirus (COVID-19) dashboard with vaccination data, 2022, https://covid19.who.int/region/amro/country/us (accessed Jun. 17, 2022).

[4] Tuberculosis, 2022, https://www.who.int/news-room/fact-sheets/detail/tuberculosis (accessed Jun. 17, 2022).

[5] M. Gabelloni, et al., Bridging gaps between images and data: a systematic update on imaging biobanks, Eur. Radiol. 32 (5) (2022) 3173–3186, http://dx.doi.org/10.1007/S00330-021-08431-6, 2021 325.

[6] A. Rimmer, Radiologist shortage leaves patient care at risk, warns royal college, BMJ 359 (2017) j4683, http://dx.doi.org/10.1136/BMJ.J4683.

[7] X. Rong, et al., Effect of delay in diagnosis on transmission of COVID-19, Math. Biosci. Eng. 17 (3) (2020) 2725–2740, http://dx.doi.org/10.3934/MBE.2020149, 2020 32725.

[8] M. Xie, et al., Multi-center federated learning, 2020, http://dx.doi.org/10.48550/arxiv.2005.01026.

[9] S.K. Zhou, et al., A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises, Proc. IEEE 109 (5) (2020) 820–838, http://dx.doi.org/10.1109/JPROC.2021.3054390.

[10] H. Zhu, J. Xu, S. Liu, Y. Jin, Federated learning on non-IID data: A survey, Neurocomputing 465 (2021) 371–390, http://dx.doi.org/10.1016/J.NEUCOM.2021.07.098.

[11] Z. Sun, S.-B. Lin, Distributed learning with dependent samples, IEEE Trans. Inform. Theory (2022) 1, http://dx.doi.org/10.1109/TIT.2022.3175761.

[12] S. Mavandadi, et al., Distributed medical image analysis and diagnosis through crowd-sourced games: A malaria case study, PLoS One 7 (5) (2012) e37245, http://dx.doi.org/10.1371/JOURNAL.PONE.0037245.

[13] T. Li, A.K. Sahu, A. Talwalkar, V. Smith, Federated learning: Challenges, methods, and future directions, IEEE Signal Process. Mag. 37 (3) (2020) 50–60, http://dx.doi.org/10.1109/MSP.2020.2975749.

[14] H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson, B. Agüera y Arcas, Communication-efficient learning of deep networks from decentralized data, in: Proc. 20th Int. Conf. Artif. Intell. Stat. AISTATS 2017, 2016, http://dx.doi.org/10.48550/arxiv.1602.05629.

[15] M. Aledhari, R. Razzak, R.M. Parizi, F. Saeed, Federated learning: A survey on enabling technologies, protocols, and applications, IEEE Access 8 (2020) 140699–140725, http://dx.doi.org/10.1109/ACCESS.2020.3013541.

[16] K.M. Jawadur Rahman, et al., Challenges, applications and design aspects of federated learning: A survey, IEEE Access 9 (2021) 124682–124700, http://dx.doi.org/10.1109/ACCESS.2021.3111118.

[17] A.E. Cetinkaya, M. Akin, S. Sagiroglu, A communication efficient federated learning approach to multi chest diseases classification, in: Proc. - 6th Int. Conf. Comput. Sci. Eng. UBMK 2021, 2021, pp. 429–434, http://dx.doi.org/10.1109/UBMK52708.2021.9558913.

[18] Z. Zheng, S. Xie, H.N. Dai, X. Chen, H. Wang, Blockchain challenges and opportunities: A survey, Int. J. Web Grid Serv. 14 (4) (2018) 352–375, http://dx.doi.org/10.1504/IJWGS.2018.095647.

[19] Z. Wang, Q. Hu, Blockchain-based Federated Learning: A Comprehensive Survey.

[20] R. Kumar, et al., Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging, IEEE Sens. J. 21 (14) (2021) 16301–16314, http://dx.doi.org/10.1109/JSEN.2021.3076767.

[21] A. Qayyum, K. Ahmad, M.A. Ahsan, A. Al-Fuqaha, J. Qadir, Collaborative federated learning for healthcare: Multi-modal COVID-19 diagnosis at the edge, 2021, http://dx.doi.org/10.48550/arxiv.2101.07511.

[22] T. Rahman, et al., Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images, Comput. Biol. Med. 132 (2021) http://dx.doi.org/10.1016/J.COMPBIOMED.2021.104319.

[23] M.E.H. Chowdhury, et al., Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8 (2020) 132665–132676, http://dx.doi.org/10.1109/ACCESS.2020.3010287.

[24] T. Rahman, et al., Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization, IEEE Access 8 (2020) 191586–191601, http://dx.doi.org/10.1109/ACCESS.2020.3031384.

[25] S. Jaeger, S. Candemir, S. Antani, Y.-X.J. Wáng, P.-X. Lu, G. Thoma, Two public chest X-ray datasets for computer-aided screening of pulmonary diseases, Quant. Imaging Med. Surg. 4 (6) (2014) 475, http://dx.doi.org/10.3978/J.ISSN.2223-4292.2014.11.20.

[26] S. Aich, et al., Protecting personal healthcare record using blockchain federated learning technologies, in: Int. Conf. Adv. Commun. Technol. ICACT, Vol. 2021-February, 2021, pp. 109–112, http://dx.doi.org/10.23919/ICACT51234.2021.9370566.

[27] M. Adnan, S. Kalra, J.C. Cresswell, G.W. Taylor, H.R. Tizhoosh, Federated learning and differential privacy for medical image analysis, Sci. Rep. 12 (1) (2022) 1–10, http://dx.doi.org/10.1038/s41598-022-05539-7, 2022 121.

[28] Y. Lu, X. Huang, Y. Dai, S. Maharjan, Y. Zhang, Blockchain and federated learning for privacy-preserved data sharing in industrial IoT, IEEE Trans. Ind. Inform. 16 (6) (2020) 4177–4186, http://dx.doi.org/10.1109/TII.2019.2942190.

[29] H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson, B. Agüera y Arcas, Communication-efficient learning of deep networks from decentralized data, in: Proc. 20th Int. Conf. Artif. Intell. Stat. AISTATS 2017, 2017.

[30] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, M. Jirstrand, A performance evaluation of federated learning algorithms, in: Proc. Second Work. Distrib. Infrastructures Deep Learn., 2018, http://dx.doi.org/10.1145/3286490.

[31] W. Huang, T. Li, D. Wang, S. Du, J. Zhang, Fairness and accuracy in federated learning, 2020, http://dx.doi.org/10.48550/arxiv.2012.10069.

[32] T. Li, A.K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, V. Smith, Federated optimization in heterogeneous networks, 2020.

[33] M.H. Rahman, M.A. Rahman, S. Talapatra, The bullwhip effect: causes, intensity, and mitigation, 8, (1) (2021) 406–426. http://dx.doi.org/10.1080/21693277.2020.1862722. http://mc.manuscriptcentral.com/tpmr.

[34] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, V. Chandra, Federated learning with non-IID data, 2018, http://dx.doi.org/10.48550/arxiv.1806.00582.

[35] T.S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I.C. Paschalidis, W. Shi, Federated learning of predictive models from federated electronic health records, Int. J. Med. Inform. 112 (2018) 59–67, http://dx.doi.org/10.1016/J.IJMEDINF.2018.01.007.

[36] M. Shohel Parvez, N. Tasnim, S. Talapatra, A. Ruhani, A.S.M.M. Hoque, Assessment of musculoskeletal problems among Bangladeshi university students in relation to classroom and library furniture, J. Inst. Eng. Ser. C 103 (3) (2022) 279–292, http://dx.doi.org/10.1007/S40032-021-00792-7, 2021 1033.

[37] N. Rieke, et al., The future of digital health with federated learning, Npj Digit. Med. 3 (1) (2020) 1–7, http://dx.doi.org/10.1038/s41746-020-00323-1, 2020 31.

[38] M.A. Khan, K. Salah, IoT security: Review, blockchain solutions, and open challenges, Futur. Gener. Comput. Syst. 82 (2018) 395–411, http://dx.doi.org/10.1016/J.FUTURE.2017.11.022.

[39] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, SSRN Electron. J. (2008) http://dx.doi.org/10.2139/ssrn.3977007.

[40] M. Nofer, P. Gomber, O. Hinz, D. Schiereck, Blockchain, Bus. Inf. Syst. Eng. 59 (3) (2017) 183–187, http://dx.doi.org/10.1007/S12599-017-0467-3, 2017 593.

[41] T.H. Pranto, A.A. Noman, M. Rahaman, A.K.M.B. Haque, A.K.M.N. Islam, R.M. Rahman, A. Blockchain, Smart contract and data mining based approach toward the betterment of E-commerce, 2021, http://dx.doi.org/10.1080/01969722.2021.2018545.

[42] M. Raikwar, S. Mazumdar, S. Ruj, S. Sen Gupta, A. Chattopadhyay, K.Y. Lam, A blockchain framework for insurance processes, in: 2018 9th IFIP Int. Conf. New Technol. Mobil. Secur. NTMS 2018 - Proc., Vol. 2018-January, 2018, pp. 1–4, http://dx.doi.org/10.1109/NTMS.2018.8328731.

[43] I. Eyal, Blockchain technology: Transforming libertarian cryptocurrency dreams to finance and banking realities, Computer (Long. Beach. Calif.) 50 (9) (2017) 38–49, http://dx.doi.org/10.1109/MC.2017.3571042.

[44] M. Pournader, Y. Shi, S. Seuring, S.C.L. Koh, Blockchain applications in supply chains, transport and logistics: a systematic review of the literature, 58 (7) (2019) 2063–2081. http://dx.doi.org/10.1080/00207543.2019.1650976.

[45] P. Ratta, A. Kaur, S. Sharma, M. Shabaz, G. Dhiman, Application of blockchain and internet of things in healthcare and medical sector: Applications, challenges, and future perspectives, J. Food Qual. (2021) 2021, http://dx.doi.org/10.1155/2021/7608296.

[46] A. Tandon, P. Kaur, M. Mäntymäki, A. Dhir, Blockchain applications in management: A bibliometric analysis and literature review, Technol. Forecast. Soc. Change 166 (2021) 120649, http://dx.doi.org/10.1016/J.TECHFORE.2021.120649.

[47] T.H. Pranto, A.A. Noman, A. Mahmud, A.B. Haque, Blockchain and smart contract for IoT enabled smart agriculture, PeerJ Comput. Sci. 7 (2021) 1–29, http://dx.doi.org/10.7717/PEERJ-CS.407/SUPP-1.

[48] J. Li, J. Wu, L. Chen, Block-secure: Blockchain based scheme for secure P2P cloud storage, Inf. Sci. (NY) 465 (2018) 219–231, http://dx.doi.org/10.1016/J.INS.2018.06.071.

[49] D.C. Nguyen, P.N. Pathirana, M. Ding, A. Seneviratne, Blockchain for secure EHRs sharing of mobile cloud based E-health systems, IEEE Access 7 (2019) 66792–66806, http://dx.doi.org/10.1109/ACCESS.2019.2917555.

[50] 2021 H2 healthcare data breach report | critical insight, 2022, https://cybersecurity.criticalinsight.com/2021_H2_HealthcareDataBreachReport (accessed Jun. 05, 2022).

[51] M. Hölbl, M. Kompara, A. Kamišalić, L.N. Zlatolas, A systematic review of the use of blockchain in healthcare, Symmetry 10 (10) (2018) 470, http://dx.doi.org/10.3390/SYM10100470, 2018, 10, Page 470.

[52] I. Institute for Business Value, Healthcare rallies for blockchainskeeping patients at the centerhealthcare rallies for blockchainskeeping patients at the centerhealthcare rallies for blockchains, 2017.

[53] Global blockchain technology market in the healthcare industry, 2018–2022, 2022, https://www.marketresearch.com/Frost-Sullivan-v383/Global-Blockchain-Technology-Healthcare-12732534/ (accessed Jun. 05, 2022).

[54] B. Zaabar, O. Cheikhrouhou, F. Jamil, M. Ammi, M. Abid, HealthBlock: A secure blockchain-based healthcare data management system, Comput. Netw. 200 (2021) 108500, http://dx.doi.org/10.1016/J.COMNET.2021.108500.

[55] U. Chelladurai, S. Pandian, A novel blockchain based electronic health record automation system for healthcare, J. Ambient Intell. Humaniz. Comput. 13 (1) (2021) 693–703, http://dx.doi.org/10.1007/S12652-021-03163-3, 2021 131.

[56] M.S. Parvez, S. Talapatra, N. Tasnim, T. Kamal, M. Murshed, Anthropomorphic investigation into improved furniture fabrication and fitting for students in a Bangladeshi university, J. Inst. Eng. Ser. C 103 (4) (2022) 613–622, http://dx.doi.org/10.1007/S40032-022-00857-1, 2022 1034.

[57] J. Konečn, H. Brendan McMahan, F.X. Yu, A. Theertha Suresh, D. Bacon Google, P. Richtárik, Federated learning: Strategies for improving communication efficiency, 2016, Accessed: Mar. 25, 2022. [Online]. Available: http://arxiv.org/abs/1610.05492.

[58] R.C. Geyer, T. Klein, M. Nabi, S. Se, E. Zurich, Differentially private federated learning: A client level perspective, 2017, http://dx.doi.org/10.48550/arxiv.1712.07557.

[59] S.R. Pokhrel, J. Choi, Federated learning with blockchain for autonomous vehicles: Analysis and design challenges, IEEE Trans. Commun. 68 (8) (2020) 4734–4746, http://dx.doi.org/10.1109/TCOMM.2020.2990686.

[60] E. Mekov, M. Miravitlles, R. Petkov, Artificial intelligence and machine learning in respiratory medicine, 14 (6) (2020) 559–564. http://dx.doi.org/10.1080/17476348.2020.1743181.

[61] RSNA pneumonia detection challenge | kaggle, 2022, https://www.kaggle.com/c/rsna-pneumonia-detection-challenge (accessed Jul. 03, 2022).

[62] M. de la I. Vayá, et al., BIMCV covid-19+: a large annotated dataset of RX and CT images from COVID-19 patients, 2020, http://dx.doi.org/10.48550/arxiv.2006.01174.

[63] RSNA pneumonia detection challenge | kaggle, 2022, https://www.kaggle.com/competitions/rsna-pneumonia-detection-challenge/overview (accessed Jul. 03, 2022).

[64] A. Fallah, A. Mokhtari, A. Ozdaglar, Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach, 2020.

[65] I. Rizwan, I. Haque, J. Neubert, Deep learning approaches to biomedical image segmentation, Inform. Med. Unlocked 18 (2020) 100297, http://dx.doi.org/10.1016/J.IMU.2020.100297.

[66] B. Yan, et al., Experiments of federated learning for COVID-19 chest X-ray images, Commun. Comput. Inf. Sci. 1423 (2021) 41–53, http://dx.doi.org/10.1007/978-3-030-78618-2_4/COVER.

[67] F. Ślazyk, P. Jabłecki, A. Lisowska, M. Malawski, S. Płotka, CXR-FL: Deep learning-based chest X-ray image analysis using federated learning, 2022, pp. 433–440, http://dx.doi.org/10.1007/978-3-031-08754-7_50.

[68] T.T. Ho, K.D. Tran, Y. Huang, FedSGDCOVID: Federated SGD COVID-19 detection under local differential privacy using chest X-ray images and symptom information, Sensors 22 (10) (2022) http://dx.doi.org/10.3390/S22103728.

[69] I. Feki, S. Ammar, Y. Kessentini, K. Muhammad, Federated learning for COVID-19 screening from chest X-ray images, Appl. Soft Comput. 106 (2021) http://dx.doi.org/10.1016/J.ASOC.2021.107330.

Update

Healthcare Analytics

Volume 4, Issue , December 2023, Page

DOI: https://doi.org/10.1016/j.health.2023.100237

Healthcare Analytics 4 (2023) 100237

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

Erratum to "Blockchain for medical collaboration: A federated learning-based approach for multi-class respiratory disease classification" [Healthc. Anal. 3 (2023) 100135]

Abdulla All Noman, Mustafizur Rahaman, Tahmid Hasan Pranto, Rashedur M. Rahman *

Department of Electrical & Computer Engineering, North South University, Dhaka 1229, Bangladesh

The publisher regrets to inform that Declaration of Competing Interest statement was not included in the published version of this article. The publisher would like to apologize for any inconvenience caused.

The appropriate Declaration/Competing Interest statements, provided by the Authors, is included below.

- All authors have participated in (a) conception and design of the experimentation (b) analysis and interpretation of the data (c) drafting the manuscript and revising it critically and (d) approval of the final version.
- This manuscript has not been published anywhere else or is in consideration for publication elsewhere.
- The authors have no other affiliation with any other organization with a direct or indirect financial interest in the subject matter of the manuscript.
- The following authors have a direct or indirect financial interest and affiliation with the following organizations.

Name	Affiliation
Abdulla All Noman	Department of Electrical and Computer Engineering, North South University, Dhaka-1229, Bangladesh
Mustafizur Rahaman	Department of Electrical and Computer Engineering, North South University, Dhaka-1229, Bangladesh
Tahmid Hasan Pranto	Department of Electrical and Computer Engineering, North South University, Dhaka-1229, Bangladesh
Rashedur M. Rahaman	Department of Electrical and Computer Engineering, North South University, Dhaka-1229, Bangladesh

DOI of original article: https://doi.org/10.1016/j.health.2023.100135.
* Corresponding author.
E-mail address: rashedur@northsouth.edu (R.M. Rahman).

https://doi.org/10.1016/j.health.2023.100237