Chest Radiographs Based Pneumothorax Detection Using Federated Learning

Ahmad Almadhor[1,*], Arfat Ahmad Khan[2], Chitapong Wechtaisong[3,*], Iqra Yousaf[4],
Natalia Kryvinska[5], Usman Tariq[6] and Haithem Ben Chikha[1]

[1]Department of Computer Engineering and Networks, College of Computer and Information Sciences, Jouf University, Sakaka, 72388, Saudi Arabia
[2]Department of Computer Science, College of Computing, Khon Kaen University, Khon Kaen, 40002, Thailand
[3]School of Telecommunication Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand
[4]Department of Computer Science and Software Engineering, International Islamic University, Islamabad, Pakistan
[5]Information Systems Department, Faculty of Management, Comenius University in Bratislava, Odbojárov, Bratislava, 440, Slovakia
[6]Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
*Corresponding Authors: Ahmad Almadhor. Email: aaalmadhor@ju.edu.sa; Chitapong Wechtaisong.
Email: chitapong@g.sut.ac.th

Abstract: Pneumothorax is a thoracic condition that occurs when a person's lungs collapse, causing air to enter the pleural cavity, the area close to the lungs and chest wall. The most persistent disease, as well as one that necessitates particular patient care and the privacy of their health records. The radiologists find it challenging to diagnose pneumothorax due to the variations in images. Deep learning-based techniques are commonly employed to solve image categorization and segmentation problems. However, it is challenging to employ it in the medical field due to privacy issues and a lack of data. To address this issue, a federated learning framework based on an Xception neural network model is proposed in this research. The pneumothorax medical image dataset is obtained from the Kaggle repository. Data preprocessing is performed on the used dataset to convert unstructured data into structured information to improve the model's performance. Min-max normalization technique is used to normalize the data, and the features are extracted from chest X-ray images. Then dataset converts into two windows to make two clients for local model training. Xception neural network model is trained on the dataset individually and aggregates model updates from two clients on the server side. To decrease the over-fitting effect, every client analyses the results three times. Client 1 performed better in round 2 with a 79.0% accuracy, and client 2 performed better in round 2 with a 77.0% accuracy. The experimental result shows the effectiveness of the federated learning-based technique on a deep neural network, reaching a 79.28% accuracy while also providing privacy to the patient's data.

Keywords: Privacy preserving; pneumothorax disease; federated learning; chest x-ray images; healthcare; machine learning; deep learning

1 Introduction

Pneumothorax appears to be a thoracic sickness disease that occurs when a person's lungs collapse, enabling air to flow into the pleural cavity, which is the region near the lungs and chest wall [1,2]. The lungs collapse due to the air leak pushing against the lung's outer wall. The lungs may partially or entirely collapse in this circumstance [2–4]. Pneumothorax can happen as a result of a chest injury that tears the lung's surface, letting air become trapped in the pleural cavity, as a result of impact lung conditions like pneumonia or chronic obstructive pulmonary disease, etc., or when the air in the pleural cavity does not flee and expands [3,5]. A severe or extensive pneumothorax may become deadly and result in shock, dyspnea, or other life-threatening symptoms. To eliminate the excess air, a large pneumothorax is considered a significant anomaly [6]. It must be treated immediately with a chest tube or needle. According to [7,8], 99.9 unexpected pneumothorax cases are for every 100,000 hospitalizations annually. According to Martinelli et al. [9], pneumothorax is being identified as one of the crucial characteristics that aggravate the cases of COVID-19 and increase the hospitalization rate.

Chest radiography, which is quick and inexpensive to obtain, is frequently used in hospitals to test for and diagnose pneumothorax. Radiologists, however, find it challenging to diagnose pneumothorax using chest X-rays because images may have superimposed formations [10]. Patterns of different thoracic disorders have varying appearances, dimensions, and positions on chest X-ray images, and patient poses might distort the image during the image recording [2]. For instance, the disease area is primarily made up of extra air, which has poor contrast and is frequently obscured inside the X-ray image by various thoracic components. Therefore, the competence of the treating radiologists is crucial for the accurate diagnosis of pneumothorax [3]. Furthermore, prompt assessment of the obtained images is essential for successful medication. The proper treatment and diagnosis of pneumothorax are frequently delayed in clinical settings where there is a shortage of radiologists with the necessary training, which can cause patients severe injury and even death. The computerized screening tool is critically needed to help medical professionals diagnose and correctly identify pneumothorax [2,3].

Artificial intelligence-based computerized disease detection has become a hot research topic in medical diagnosis over the past few years [1,11,12]. By reducing the number of errors made during the image interpretation process, AI systems can improve the efficiency of every disease detection method [13]. For image categorization and segmentation issues, deep learning-based methods are frequently used. These methods already have tremendous success in the fields of comprehending natural scene images [14,15], medical image processing [16,17], and imagery from remote sensing analysis [18]. To correctly diagnose diseases, seasoned radiologists must carefully alter image display parameters, including window width, level, and intensity. Chest X-rays, unlike raw images, have a low resolution and little contrast, making them particularly difficult to evaluate. Additionally, because of these unique qualities, it is particularly challenging to automatically detect and segment lesions from chest X-ray images [3]. Data mining, machine learning, and deep learning are employed in the information era to turn data into knowledge for the early diagnosis of diseases [19,20]. Lack of data, however, makes it challenging to use artificial intelligence models in the medical world due to privacy concerns [21,22]. Private data cannot be guaranteed to be secure using machine learning algorithms since it is so simple for attackers to steal or change information, whether on purpose or accidentally [23].

Federated learning (FL), which protects data privacy, was created to deal with these issues because previous techniques are not considering this. Artificial Intelligent systems can learn about individual data through federated learning without compromising privacy [24–27]. By keeping their data private, end devices can take part in learning and disseminating the prediction model while disclosing their

information. Currently, researchers are focusing on using federated learning with health records. The study's main contributions are given below:

- This research proposes an approach that detects, and diagnoses pneumothorax disease based on federated learning using the pneumothorax (Chest X-ray) images data.
- The in-depth features are extracted from Chest X-ray images using the Xception model while normalizing the data using the min-max normalization technique.
- The proposed model trains the local model first and sends the model's parameters to the global model. Then the global model aggregates the updated parameters and trains the process by ensuring the privacy of each client's data.
- The experiment shows that the federated learning model based on an Xception neural network model improves accuracy for pneumothorax disease by 79.28% compared to conventional methods and offers client data protection.

The paper is organized into the following sections: Section 2 discusses the related work of machine learning, deep learning, and federated learning techniques for diagnosing pneumothorax disease. The proposed framework, which covers dataset selection, data preprocessing, feature extraction, and model architecture, is described in Section 3. Section 4 explains the results, and Section 5 presents the work's conclusion and recommendations for future research.

2 Literature Review

This section provides the history of Pneumothorax disease using machine learning and deep learning techniques. The background of the healthcare devices using federated learning is also provided.

2.1 Machine and Deep Learning Techniques

Machine learning has been used for various healthcare applications such as tumor detection, activity recognition, health assessment, dementia detection, and many others [28–33]. Authors in this research [34] verify a machine learning model developed on an open-source dataset and then optimize it for chest X-rays of patients with significant pneumothorax. The research is in the form of a retrospective. Pneumothorax cases and all the other cases (non-Pneumothorax) are separated from the open-source chest X-ray dataset, providing 41,946 non-Pneumothorax and 4696 Pneumothorax cases for the training set and 11,120 non-Pneumothorax and 541 Pneumothorax cases for the testing set. The construction of a restricted supervision machine learning model included localized and un-localized pathologies. Then, from 2013 to 2017, cases from the healthcare system of the institutions are analyzed. The training set included 682 non-pneumothorax instances and 159 pneumothorax records. The validation set included 48 pneumothorax instances and 1287 non-pneumothorax records. The proposed model performed well on the open-source dataset with 0.90% AUC.

The author proposed the automatic method used for image multi-scale intensity texture analysis and segmentation in this research [35]. Support Vector Machine (SVM) is used first to detect frequent pneumothorax. Utilizing the Local Binary Pattern (LBP), features are retrieved from lung images. The SVM is utilized to classify pneumothorax. Then next, the proposed automatic pneumothorax detection approach utilizes multi-scale intensity texture segmentation to separate aberrant lung regions from backgrounds and noise in chest pictures. For texture transformation from calculating numerous overlapping blocks, anomalous regions are segmented. Using the technique of Sobel edge detection, the rib margins are located. The ribbed border is filled between the abnormal sections to obtain an entire disease region. The proposed approach obtained 85.8% highest accuracy with five by five patch

size. They have labeled and evaluated a massive dataset of chest radiographs to encourage innovation in this field. They looked into the value of utilizing AI models to produce comments for review because labeling images is difficult and time-consuming. This research's authors proposed a machine learning annotation (MLA) approach to quicken the annotation procedure and apply sensitivity and specificity as evaluation measurements [36].

Deep learning has been utilized recently to analyze clinical data across several sectors [37,38], and it excels in tasks like image segmentation and recognition [39–41]. The author of [2] presents the Deep Neural Network (DNN) model to detect pneumothorax regions in chest X-ray images. The model uses transfer learning and a Mask Regional Convolutional Neural Network (Mask RCNN) structure, with ResNet101 serving as the Feature Pyramid Network (FPN). The proposed model is developed using a pneumothorax dataset created by the American College of Radiology and Society for Imaging Informatics in Medicine (SIIM-ACR).

A deep-learning-based AI model is proposed in [1] and is used as a treatment method created by skilled thoracic surgeons to predict the amount of pneumothorax visible on a chest radiograph. U-net carries out semantic separation and categorization of pneumothorax and non-pneumothorax regions. Chest computed tomography (standard of gold and ratio of volume) and chest radiography (accurate labels and ratio of area) are being used to measure the quantity of pneumothorax, and the artificial intelligence model is being used to calculate it (label prediction and ratio of area). Based on clinical results, each value is compared and examined. The research examined 96 patients, 67 of whom made up the training set and the remaining patients the test set. The artificial intelligence model demonstrated a 61.8% dice similarity coefficient, a 69.2% sensitivity, a 99.1% negative predictive value, and 97.8% accuracy. A pneumothorax detection methodology was proposed to examine the capability of transfer learning (TL). For this study, the CheXpert dataset was chosen. The trained model CheXNet was applied to initialize the model weights in the 122-layer deep neural network (DNN). Just 13911 of the 94,948 CXRs used to train the model (out of 94,948) corresponded to the pneumothorax class. They used the weighted binary cross-entropy loss function to resolve this class-imbalance issue. The framework demonstrated an AUC of 70.8% when tested on a batch of 7 pneumothorax images and 195 normal CXR [42].

The author [43] recommends a feature-based neural network for pneumothorax diagnosis combining lateral X-ray and frontal data. There are two inputs and three outputs in this network. The frontal and lateral chest X-ray images serve as the two inputs. The three outputs are the frontal chest X-ray image classification outcomes, the lateral chest X-ray image classification results, and the classification performance incorporating the properties of the lateral chest and fused frontal X-ray images. The proposed technique incorporates the residual block to address the pneumothorax recognition model's vanishing gradient issue. The work uses channel attention strategies to enhance the model's performance due to the vast number of channels in the system. The proposed model yields the best results compared to other methods, with an area under the curve score of 0.979. A deep learning-based method was presented to segregate the pathologic region autonomously. These attention masks provided a rough estimate of the area of pathology after training an image-level classifier that not only predicted the class of the CXR pictures but also generated the attention masks. The segmentation model was developed with some images with good pixel-level annotations and the sparsely annotated masks produced by the image-level classifier. Spatial label smoothing regularisation (SLSR) was used to fix several problems in these masks. The ResNet101 model was used to classify images at the image level, while the Guided Attention Inference Network created masks with weak annotations (GAIN). Three distinct types of architectures, including U-Net, LinkNet, and Tiramisu (FCDenseNet67), were tested while maintaining an input resolution of $256 * 256$ for segmentation purposes. The experimental

data, which included 5400 CXRS, including 3400 pneumothorax instances and 2000 normal CXRs, was gathered from a Chinese Hospital. The dataset was uniformly and randomly split into training and testing sets, with 4000 CXRS in the training set and 1400 CXRS in the test set. The model's performance was assessed using a test set, and the Intersection over Union (IoU) score was reported as 66.69% [44].

2.2 Pneumothorax Detection Privacy

The author of this study describes non-invasive sensing-based diagnostics of pneumonia disease that maintain security while utilizing a deep learning model to make the procedure non-invasive [45]. The proposed approach is split into two parts. The first part uses the proposed approach based on a chaotic CNN to encode primary data in the shape of images. Additionally, numerous chaotic maps are combined to construct a random number generator. The resulting random series is then used for pixel permutation and replacement. The second part of the proposed approach proposes a unique DL method for diagnosing pneumonia employing X-ray images as a dataset. Machine learning classifiers such as RF, DT, SVM, and NB are also applied to compare the proposed approach CNN model. The CNN-based proposed technique achieved good performance at 97% compared to ML techniques.

2.3 Federated Learning Techniques

The authors in [46] present a Federated Edge Aggregator (FEA) system with Distributed Protection (DP) using IoT technologies. This study's performance of the suggested technique from a healthcare perspective is demonstrated using data from the MNIST database, CIFAR10, and the COVID-19 dataset. An iteration-based converged Convolutional Neural Network (CNN) algorithm at the Edge Layer (EL) is presented to balance the privacy protection provided by federated learning with algorithm performance throughout an IoT network. After a predetermined number of iterations, the results outperform other baseline methodologies, showing a 90% best accuracy. Furthermore, the suggested strategy more rapidly and successfully satisfies the privacy and security notion. An edge-assisted framework using federated learning is suggested in [47] for training local models using data generated by nearby healthcare IoT devices while protecting user data privacy and utilizing the least resources. The suggested framework consists of three modules: cloud, edge, and application. A three-part edge module will download the initial trained model from the cloud module after it has been transmitted there. The federated learning server will determine the edge device's eligibility and the local storage controller's local data. The cloud modules aggregated and controller components will notify the owner of the updated global model after model training using an edge device. The application module is in charge of producing data and assembling medical equipment. The architecture currently used for disease monitoring, tracking, and prevention has some limitations, like assaults.

Wearable sensors/devices are becoming well known for health monitoring and activities, including heart rate monitoring, medication timing, pulse rate monitoring, sleep, walking, etc., with the rise of IoT technology [48–54]. Without preserving the privacy of the data, machine learning models continue to enhance the functionality of such devices. To solve this problem, a Fed Health framework is put out [55]. Fed Health employs federated learning for data gathering and transfer learning for model construction. While maintaining data privacy, Fed Health's accuracy increased by 5.3% compared to other models. This article analyses various FL approaches and proposes a real-time distributed networking framework based on the Message Queuing Telemetry Transport (MQTT) protocol. They specifically design various ML network strategies based on FL tools utilizing a parameter server (PS) and completely decentralized paradigms propelled by consensus techniques. The suggested method is validated in brain tumor segmentation, employing a refined version of the well-known U-NET model

and realistic clinical datasets gathered from everyday clinical practice. Several physically independent computers spread across several nations and connected over the Internet are used to execute the FL procedure. The real-time test bed is employed to quantify the trade-offs between training accuracy and latency and to emphasize the crucial operational factors that influence effectiveness in actual implementations [56]. Another research used a federated learning framework in which authors propose a decentralized SDN controller to operate as an agent, provide a softwarization mechanism, and communicate with the virtualization based EFL environment [57]. Adopting the architecture, the agent controller and the orchestrator enable a centralized perspective for deploying virtual network services with effective virtual machine mapping to carry out edge modification processes and effectively build the forwarding path. The proposed agent valuation and reward output are the foundations for the proposed NFVeEC for local-EFL upgrades.

In conclusion, many machine and deep learning methods have been investigated to develop pneumothorax disease prediction methods. However, they are limited in performance, which is mentioned in Table 1. They also did not offer reassuring evidence of improved accuracy and did not prioritize data protection. To preserve data privacy that ML and DL models do not account for, we proposed a federated learning framework in this study that is built. The federated learning research that was previously mentioned centered on resolving various healthcare challenges [47,48,55].

Table 1: Literature review

Author	Techniques	Dataset	Limitation
Filice et al. [36]	Machine learning annotations	NIH chest X-ray dataset	Low performance
Maziad et al. [37]	Deep transfer learning	Chest X-ray	Limited accuracy
Kim et al. [1]	Deep-learning based AI model	Chest radiograph	Low sensitivity
Dey [39]	Supervised segmentation framework	Chest X-ray	Low performance

3 Proposed Framework

This section explains the concepts of federated learning, data preprocessing, and model architecture as they are essential to the proposed framework. The experiment of the proposed methodology is carried out on the pycharm IDE, and python language is used. In the proposed federated learning framework, the local model of each client is trained before receiving the model parameters for the global model. After aggregating the new parameter, the training process is initiated using the global model, which guarantees the privacy of the client's data.

Fig. 1 shows the steps performed in the proposed framework. The first step is to obtain a pneumothorax (Chest X-ray) image dataset from the Kaggle repository. The second step performs data preprocessing by extracting the in-depth features using Xception neural network model and normalizing the data into the range [0, 1] using the min-max normalization technique. The preprocessed dataset is divided into two panes for two clients in the last step. The preprocessed dataset is then used to train the local model using a deep neural network. Results from both clients are combined and updated in the global model.
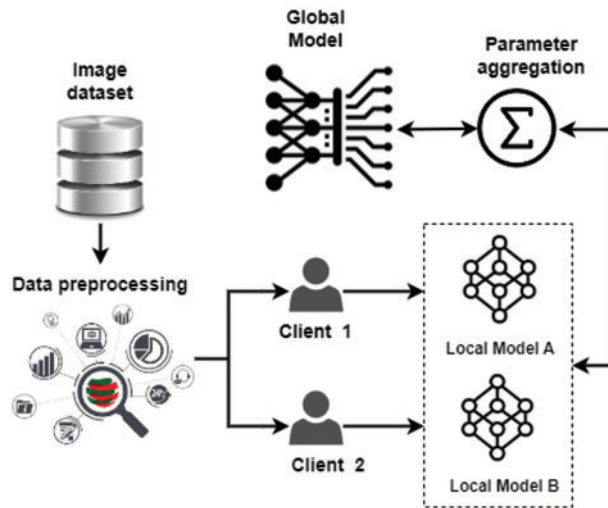
Figure 1: Proposed federated learning framework

According to the target application, the cloud-based federated learning server analyses essential data types and trains hyper-parameters, including learning rate, number of epochs, activation function, and Adam optimizer [58,59]. The federated learning model involves three essential steps. The first stage is initializing training [60]. Additionally, the federated learning server first develops a global model [61,62]. The Xception neural network model has specified client specifications and multiple hyper-parameters. It is worth noting that the federated learning server chooses the model's epoch and learning rate.

The Xception neural network model needs to be trained at the second level. Every client begins by collecting new information and updating the local model's (M^y_x) parameter, which relies on the global model (G^y_x), where y is the index for the subsequent iteration. Every client searches for the ideal scenario to reduce the loss. Finally, periodically submit the updated settings to the federated learning server. The third level is the integration of the global model. Deliver the updated parameters to every client after combining the results from multiple clients on the server side. The federated learning server uses Eq. (1) to focus on decreasing the global mean loss function.

$$Loss\,(G^y) = \frac{1}{M}\sum\nolimits_{x=M}^{x=1} Loss\,(M^y_x) \tag{1}$$

In Algorithm 1, the *Ds* denoted the dataset split into two parts after preprocessing for two clients. Then extract the features *fe* by using the Xception model. At the server side in the algorithm, initialize the model weight *w*0. which is expressed as w0. The algorithm will run for a specific amount of rounds. *ti* denotes the current round of the model, and T is the total round of the local model. The number of clients in each round is decided *ci* is the current client, and C is the total number of clients. A local model upgrade receives the current weight as input and delivers the new weight value. The server learning rate is multiplied by the received updates and added together to get the new weight value. On the Client side, data is split into two windows. The weights should be modified for every client after a set number of local epochs and batch size. Update the local model by calculating the loss function Fi(w). The procedure is repeated until the requisite accuracy is attained or the loss function is constantly minimized. Table 2 represents all the symbols used in the proposed framework.

Table 2: Description of the symbols

Symbols	Symbols description
(M^y_x)	Local model
(G^y_x)	Global model
$\sum$	Summation
D_s	Dataset
f_e	Feature extraction
w_0	Model weight
t_i	Model current round
T	Total rounds
ci	Current client
C	Total clients
$F_i(w)$	Loss function
E_p	Epochs size
B_s	Batch size
Z_{norm}	Normalization

3.1 Experimental Dataset

This research uses the pneumothorax medical image dataset from the Kaggle repository. The small pneumothorax dataset consists of 2027 medical images of lungs collected by the radiologist during the chest X-ray of patients. Pneumothorax small dataset (2027 images) used for binary classification task either Pneumothorax or not. The total training data is 2027. Fig. 2 represents the samples of the Pneumothorax dataset in the form of chest X-ray images. The dataset is divided into two dormers because this study works with two clients, each with its data. The dataset was split into training and testing data, with 75% data for training models and 25% for testing the models.


Algorithm 1: Federated Learning Framework

1: Dataset D_s (Pneumothorax/Chest X-ray image)
2: Feature extraction f_e (Xception model)
3: function global model upgrade
4:      weight initialization w_0
5:      for $(t_i = 1)$ to T do
6:          for $(ci = 1)$ to C do
7:          $w^{ci}_{ti+1} =$ local model upgrade (c_i, w_{ti})
8:      $w_{ti+1} = \sum_{ci=1}^{C} w^{ci}_{Ds} * w^{ti}_{ci+1}$
9: function local model upgrade (ci, w_{ti})
10:      for $(Epoch = 1)$ to E_p do
11:          for $(B_s \in BatchSize)$ do
12:          $w = w - \nabla F_i(w)$
13:      return w
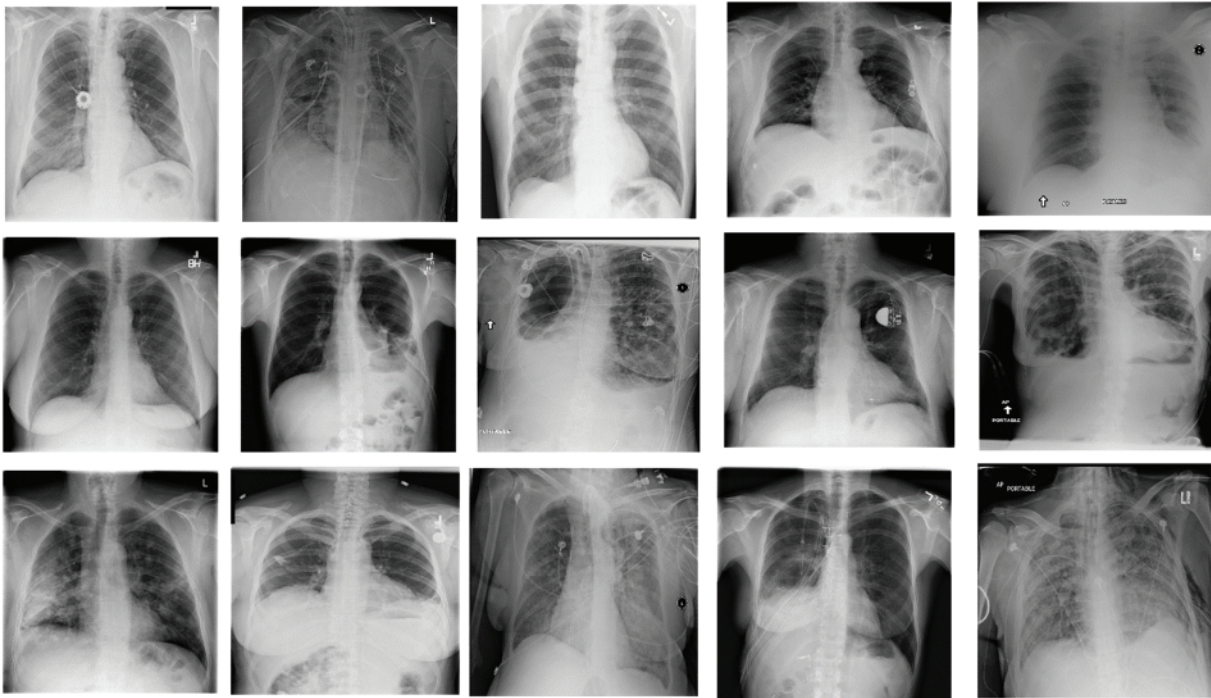
Figure 2: Samples of pneumothorax dataset (chest X-ray images)

3.2 Data Preprocessing

Data preprocessing in machine learning involves transforming unstructured data into a format used to develop and improve machine learning models. Data preprocessing is the first step in machine learning before creating a model. Data preprocessing is a vital process that raises data reliability to extract valuable information. Actual data is typically inconsistent, erroneous (contains errors or outliers), partial, and lacking specific attribute values or trends. Data preprocessing is crucial because it makes it simpler to arrange, filter, and raw format data for machine learning models to use. The min-max normalization technique is used for data preprocessing. The min-max scaling for normalizing the features is used to structure the low variance and ambiguous dataset and maintain the data integrity. The input attributes scaling is essential for a specific model that depends on the magnitude of values. As a result, normalizing refers to the discrete range of real-valued numerical characteristics between 0 and 1. For data normalization, Eq. (2) is being used.

$$Z_{norm} = \frac{Z_i - Z_{min}}{Z_{max} - Z_{min}} \tag{2}$$

After Preprocessing, the dataset is divided into two parts: a training dataset and a testing dataset, where 75% of the dataset is used to train the proposed model, and 25% is used as testing data to evaluate the proposed model.

3.3 Feature Extraction

After preprocessing the data, the pneumothorax (chest X-ray) image dataset is trained. Features are extracted that represent each pneumothorax. The Xception neural network model extracts the significant features without any human help. The advantage of deep learning is that it decides how

to use convolutional filters for extracting the features from the training dataset. We use the Xception deep neural network model to classify pneumothorax types and extract the in-depth features.

3.4 Model Architecture

The Xception deep neural network model is used for pneumothorax prediction [63]. The neurons and the number of layers are crucial when modeling the structure of neural networks. In a deep neural network model, the size of the training set determines how many neurons are used as input and output. An inception network is a deep neural network with a structure of recurring modules known as inception components. The Xception model generally consists of three layers: an input layer, different hidden layers (dropout, dense, etc.), and the output layer. This study uses a sequential Xception model with one input layer. The shape of the input layer is 2048 with the relu activation function. The hidden layer is the next: the hidden layer comprises one dropout layer and one dense layer. To prevent the model from overfitting, dropout layers are being used. The value of the dropout layer is 0.3. The dense layer comprises 2 units, and the activation function is sigmoid. The next is the output layer, which is the fully connected layer used for binary classification problems. To address the binary classification problem, every dense layer uses the activation functions relu along with a fully connected layer. The Xception model has utilized Adam as an optimizer to calculate and reduce the loss; the Xception model uses categorical cross-entropy.

4 Experimental Results and Analysis

The experimental results and an assessment of the suggested approach are presented in this section. The study also explores the impacts of several methodology-proposed parameters. The experiment is conducted on a pneumothorax medical image (chest X-ray) dataset from the Kaggle repository. In the experiment, one server and two clients are engaged. The two clients use the chest X-ray image dataset to train the Xception neural network model. The Xception neural network model is trained using the pneumothorax (chest X-ray) image dataset and establishes the random weight. The specifics of the Xception neural network model are given in the section on model architecture. We initially collected 2027 chest X-ray images from Kaggle, and then we extracted the detailed features that indicate whether a pneumothorax is present. The dataset is split into two portions: 25% is utilized for model testing, and 75% is used for model training. All numbers are normalized using the min-max normalization to fit into the range [0, 1]. The experiment employs a variety of evaluation criteria, including recall, f1-score, precision, and accuracy. For loss prevention, the results for each client are checked three times. The client's results are aggregated at the server end. The label is transformed into a machine-readable format using the label encoder approach. The accuracy of the Xception neural network model is 79.28% following server-side aggregation.

4.1 Server-Based Training Using Log Data

The Xception neural network model is trained using data logged on servers. Two clients comprise a federated learning system, with one serving as the primary parameter. At an introductory level of the model training stage, the server selects which client or node is used and gathers any changes received. Logs are anonymized and stripped of any personally identifiable details prior to training. After the federated learning starts, there are three rounds. We set the number of rounds to 3 and evaluated the tests three times after initializing some server-side parameters. The clients send the learning outcomes to the server during the fit-round. In the assessment phase, both clients send the test results to the server when the findings are aggregated. According to the server's analysis of the data from N clients, the pneumothorax neural network model has the maximum accuracy, scoring 79.28%.

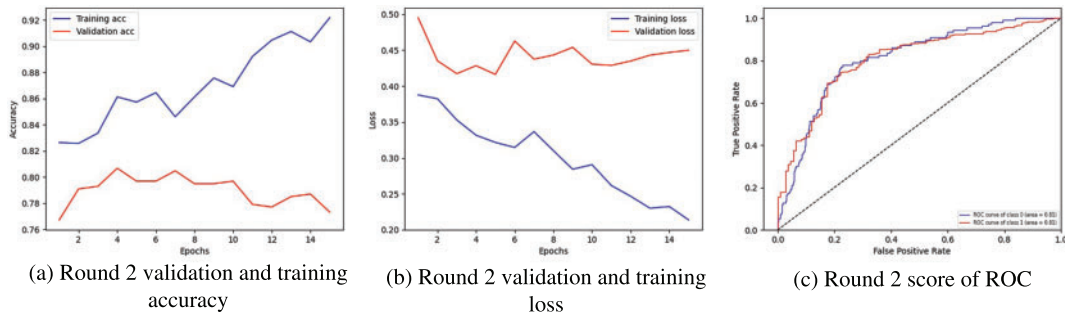### 4.2 Federated Training and Testing Using Client 1

Client 1 conducts the experiment using a sequential Xception neural network model with two dense and one dropout layer as hidden layers. The input shape of the input layer is 2048 with the relu activation function. Adam optimizer calculates and reduces the loss; the Xception neural network model uses categorical cross-entropy. Client 1 conducted the experiments three times and returned the results in two rounds: the evaluation round and the fit round. The experiment is conducted on Xception neural network model that runs three test rounds. The experiment uses different evaluation metrics: precision, accuracy, f1-score, and recall in three rounds. Table 3 represents the experimental results of client 1. The study runs the Xception neural network model for round 1, and the result obtains in terms of accuracy, recall, precision, and F1-score. In the first round, the Xception neural network model provides 77.0% accuracy, 82.0% precision, 77.0% recall, and 79.0% F1 score. Analyze the results to prevent the model from overfitting once again. In round 2, the Xception neural network model achieved an accuracy of 79.0%, a precision of 82.0%, a recall of 77.0%, and an F1-score of 79.0%. In round 3 Xception neural network model obtained an accuracy of 77.0%, a precision of 80.0%, a recall of 80.0%, and an F1-score of 80.0%. From client one, round 2 provides the highest result.

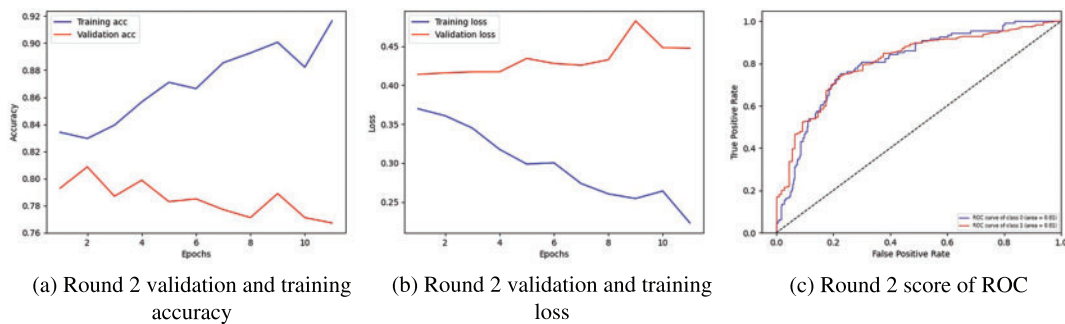**Table 3:** Proposed model result of client 1

Rounds	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Round 1	77.0	82.0	77.0	79.0
Round 2	79.0	82.0	77.0	79.0
Round 3	77.0	80.0	80.0	80.0

In Fig. 3, the highest result is visualized. In Fig. 3a, the graph represents the training and validation accuracy, in which round 2 of client 1 obtained the highest validation accuracy than training accuracy. The training accuracy curve is represented by a blue line, while an orange line represents the validation accuracy curve. Training accuracy started from 0.0.82% at 0th and, after some fluctuation between drops and gains, reaches 0.92% percent accuracy at the 14th epoch. Validation accuracy is 0.76% at 0th and, after some fluctuation between drops and gains, reaches 0.77% percent accuracy at the 14th epoch. The graph in Fig. 3b shows the training and validation loss; in the training process, the loss is reduced on each epoch, which increases the model's performance. The loss curve in which the blue line represents the training loss curve, and the orange line curve shows the validation loss. Training loss initiated from 0.40% at 0th epoch and decreased to 0.20 at 14th epoch. Validation loss initiated from 0.50% at 0th epoch and decreased to 0.45% at 14th epoch.

The graph in Fig. 3c indicates the Receiver Operating Characteristic (ROC). Two classes are used in the experiment, and both received receiver operating characteristic (ROC) scores of 0.81%, representing better performance on the used dataset for the proposed model. Better performance is represented by the receiver operating characteristic (ROC) curves adjacent to the top-left corner. The receiver operating characteristic (ROC) curves close to the top-left corner demonstrate better performance. The confusion matrix for client one in the proposed methodology is graphically depicted in Fig. 4a, which provides a broad idea of how a classification algorithm works. The proposed method works better because it has more continuous, better valid positive and negative values and fewer false positive and negative values.

(a) Round 2 validation and training
accuracy

(b) Round 2 validation and training
loss

(c) Round 2 score of ROC

Figure 3: Representation of client one highest-scoring outcomes

(a) Round 2 validation and training
accuracy

(b) Round 2 validation and training
loss

(c) Round 2 score of ROC

Figure 4: Representation of client 2 highest-scoring outcomes

4.3 Federated Training and Testing Using Client 2

The experiment is conducted on Xception neural network model that runs three test rounds. Client 2 conducts the experiment using a sequential Xception neural network model with two dense and one dropout layer as hidden layers. The input shape of the input layer is 2048 with the relu activation function. Adam optimizer calculates and reduces the loss; the Xception neural network model uses categorical cross-entropy. Client 2 conducts the experiments three times and returns the results in the evaluation and fit rounds. The experiment uses different evaluation metrics: precision, accuracy, f1-score, and recall in three rounds.

Table 4 represents the experimental results of client 2. The study runs the Xception neural network model for round 1, and the result obtains accuracy, precision, recall, and f1-score. In the first round, the Xception neural network model provides an accuracy of 76.0%, 77.0% precision, 85.0% recall, and 81.0% F1-score. Analyze the results to prevent the model from overfitting once again. In round 2, the Xception neural network model achieved an accuracy of 77.0%, 81.0% precision, 79.0% recall, and 80.0% F1-score. In round 3, we obtained an accuracy of 79.0%, 78.0% precision, 80.0% recall, and 79.0% F1-score. From client two, round 2 provides the highest result.

Table 5 demonstrate the comparison of proposed technique with previous technique that used to predict pneumothorx. Fig. 4 visualized the highest result. The graph in Fig. 4a represents the training and validation accuracy. The validation and training accuracy graph represent that round 2 of client 2 obtained the highest validation accuracy compared to training accuracy. The training accuracy curve is represented by a blue line, while an orange line represents the validation accuracy curve. Training accuracy started from 0.0.84% at 0th and, after some fluctuation between drops and gains, reaches 0.92% percent accuracy at the 12th epoch. Validation accuracy is 0.79% at the 0th and, after some fluctuation between drops and gains, reaches 0.77% percent accuracy at the 12th epoch.

Table 4: Proposed model result of client 2

Rounds	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Round 1	76.0	77.0	85.0	81.0
Round 2	77.0	81.0	79.0	80.0
Round 3	76.0	78.0	80.0	79.0

Table 5: Comparison with previous approach

Author	Techniques	Dataset	Performance
Filice et al. [36]	Machine learning annotations	NIH chest X-ray dataset	Low performance
Maziad et al. [37]	Deep transfer learning	Chest X-ray	70.08%
Dey [39]	Supervised segementation framework	Chest X-ray	66.69%
Our approach	Xception based federated framework	Chest X-ray	79.28%

In Fig. 4b, the graph shows the training and validation loss; in the training process, the loss is reduced on each epoch, which increases the model's performance. The loss curve in which the blue line represents the training loss curve, and the validation loss is shown by the orange line curve. Training loss initiated from 0.37% at 0th epoch and decreased to 0.20 at 12th epoch. Validation loss initiated from 0.41% at 0th epoch and increased to 0.44% at 12th epoch. The third graph indicates the Receiver Operating Characteristic (ROC) in Fig. 4c. Two classes are used in the experiment, and both received receiver operating characteristic (ROC) scores of 0.81%, representing better performance on the used dataset for the proposed model. Better performance is represented by the receiver operating characteristic (ROC) curves adjacent to the top-left corner. The ROC curve represents better performance near the corner of the top-left. The confusion matrix for client two in the proposed methodology is graphically depicted in Figs. 5a and 5b, which provides a broad idea of how a classification algorithm works. The proposed method works better because it has more continuous, better valid positive and negative values and fewer false positive and negative values.

(a) Confusion matrix of client 1   (b) Confusion matrix of client 2
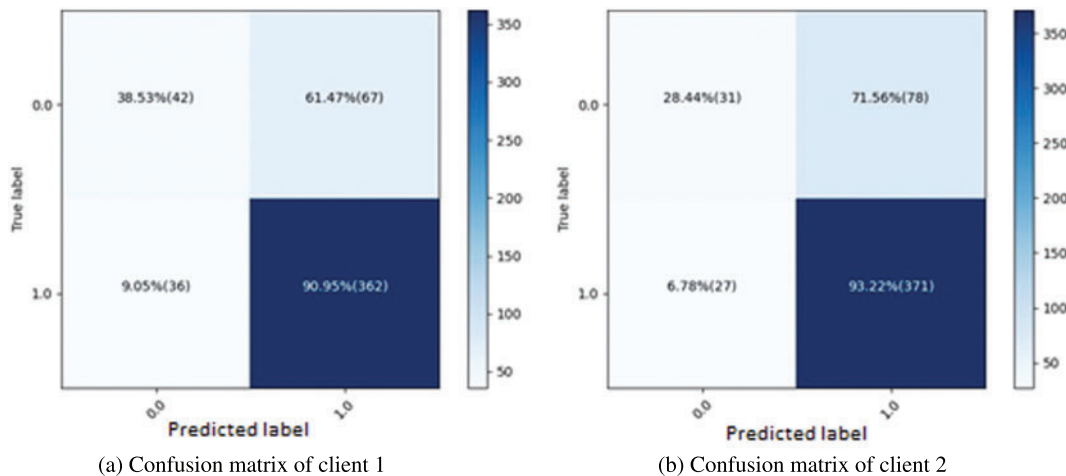
Figure 5: Confusion matrices for both clients using proposed methods

5 Conclusion

This paper proposed a federated learning based Xception neural network model to detect and diagnose pneumothorax disease. Two clients are used to train the Xception neural network model, and the server end is combined with an accuracy of 79.28%. To lessen the over-fitting aspect, every client reviewed the outcomes three times. The best result of round 2 of client one was obtained with a 79.0% accuracy, and client 2 of round 2 obtained the best result with a 77.0% accuracy. The Xception neural network model's ROC curve average of 81.0% demonstrates the suggested method's excellent performance on the tested dataset. The findings showed that federated learning accurately secures the privacy of client data, but it is limited in accuracy. In further studies, we will apply different deep learning models to improve the performance of the federated learning framework. According to a system efficiency study, side training times and storage costs favor medical devices with constrained resources. In the future, we plan to investigate this trend more thoroughly by training more systems using various smartphone device combinations and expanding our research by applying new deep learning algorithms with multiple datasets.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] D. Kim, J. H. Lee, S. W. Kim, J. M. Hong, S. J. Kim *et al.,* "Quantitative measurement of pneumothorax using artificial intelligence management model and clinical application," *Diagnostics*, vol. 12, no. 8, pp. 1823, 2022.

[2] P. Malhotra, S. Gupta, D. Koundal, A. Zaguia, M. Kaur *et al.,* "Deep learning-based computeraided pneumothorax detection using chest X-ray images," *Sensors*, vol. 22, no. 6, pp. 2278, 2022.

[3] X. Wang, S. Yang, J. Lan, Y. Fang, J. He *et al.,* "Automatic segmentation of pneumothorax in chest radiographs based on a two-stage deep learning method," *IEEE Transactions on Cognitive and Developmental Systems*, 2020.

[4] A. Gooßen, H. Deshpande, T. Harder, E. Schwab, I. Baltruschat *et al.,* "Deep learning for pneumothorax detection and localization in chest radiographs," arXiv preprint arXiv:1907.07324, 2019.

[5] K. Williams, T. A. Oyetunji, G. Hsuing, R. J. Hendrickson and T. B. Lautz, "Spontaneous pneumothorax in children: National management strategies and outcomes," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 28, no. 2, pp. 218–222, 2018.

[6] M. Annarumma, S. J. Withey, R. J. Bakewell, E. Pesce, V. Goh *et al.,* "Automated triaging of adult chest radiographs with deep artificial neural networks," *Radiology*, vol. 291, no. 1, pp. 196, 2019.

[7] K. Rami, P. Damor, G. P. Upadhyay and N. Thakor, "Profile of patients of spontaneous pneumothorax of North Gujarat region, India: A prospective study at gmers medical college, information and knowledge for health, Dharpur-Patan," 2015.

[8] A. P. Wakai, "Spontaneous pneumothorax," *BMJ Clinical Evidence*, vol. 2011, 2011.

[9] A. W. Martinelli, T. Ingle, J. Newman, I. Nadeem, K. Jackson *et al.,* "Covid-19 and pneumothorax: A multicentre retrospective case series," *European Respiratory Journal*, vol. 56, no. 5, 2020.

[10] C. Mehanian, S. Kulhare, R. Millin, X. Zheng, C. Gregory *et al.,* "Deep learning-based pneumothorax detection in ultrasound videos," in *Smart Ultrasound Imaging and Perinatal, Preterm and Paediatric Image Analysis*, Shenzhen, China, Springer, pp. 74–82, 2019.

[11] J. M. Hillis, B. C. Bizzo, S. Mercaldo, J. K. Chin, I. Newbury-Chaet *et al.,* "Evaluation of an artificial intelligence model for detection of pneumothorax and tension pneumothorax on chest radiograph," *medRxiv*, 2022.

[12] A. R. Javed, F. Shahzad, S. ur Rehman, Y. B. Zikria, I. Razzak *et al.,* "Future smart cities requirements, emerging technologies, applications, challenges, and future aspects," *Cities*, vol. 129, pp. 103794, 2022.

[13] P. Malhotra, S. Gupta and D. Koundal, "Computer aided diagnosis of pneumonia from chest radiographs," *Journal of Computational and Theoretical Nanoscience*, vol. 16, no. 10, pp. 4202–4213, 2019.

[14] C. Wang, S. Dong, X. Zhao, G. Papanastasiou, H. Zhang *et al.,* "Saliencygan: Deep learning semisupervised salient object detection in the fog of IoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2667–2676, 2019.

[15] D. Zhang, J. Han, Y. Zhang and D. Xu, "Synthesizing supervision for learning deep saliency network without human annotation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 7, pp. 1755–1769, 2019.

[16] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang *et al.,* "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[17] H. Lee, S. Yune, M. Mansouri, M. Kim, S. H. Tajmir *et al.,* "An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets," *Nature Biomedical Engineering*, vol. 3, no. 3, pp. 173–182, 2019.

[18] K. Li, G. Wan, G. Cheng, L. Meng and J. Han, "Object detection in optical remote sensing images: A survey and a new benchmark," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 296–307, 2020.

[19] M. Christ, T. Bertsch, S. Popp, P. Bahrmann, H. J. Heppner *et al.,* "High-sensitivity troponin assays in the evaluation of patients with acute chest pain in the emergency department," *Clinical Chemistry and Laboratory Medicine (CCLM)*, vol. 49, no. 12, pp. 1955–1963, 2011.

[20] S. Pandya, T. R. Gadekallu, P. K. Reddy, W. Wang and M. Alazab, "Infusedheart: A novel knowledge-infused learning framework for diagnosis of cardiovascular events," *IEEE Transactions on Computational Social Systems*, 2022.

[21] A. Amanat, M. Rizwan, M. Abdelhaq, R. Alsaqour, S. Pandya *et al.,* "Deep learning for depression detection from textual data," *Electronics*, vol. 11, no. 5, pp. 676, 2022.

[22] R. A. Haraty, B. Boukhari and S. Kaddoura, "An effective hash-based assessment and recovery algorithm for healthcare systems," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 1523–1536, 2022.

[23] B. Xue, M. Warkentin, L. A. Mutchler and P. Balozian, "Self-efficacy in information security: A replication study," *Journal of Computer Information Systems*, pp. 1–10, 2021.

[24] A. Rehman, I. Razzak and G. Xu, "Federated learning for privacy preservation of healthcare data from smartphone-based side-channel attacks," *IEEE Journal of Biomedical and Health Informatics*, 2022.

[25] Z. Lian, W. Wang, H. Huang and C. Su, "Layer-based communication-efficient federated learning with privacy preservation," *IEICE Transactions on Information and Systems*, vol. 105, no. 2, pp. 256–263, 2022.

[26] Z. Lian, Q. Yang, W. Wang, Q. Zeng, M. Alazab *et al.,* "Deep-fel: Decentralized, efficient and privacy-enhanced federated edge learning for healthcare cyber physical systems," *IEEE Transactions on Network Science and Engineering*, 2022.

[27] J. Song, W. Wang, T. R. Gadekallu, J. Cao and Y. Liu, "Eppda: An efficient privacy-preserving data aggregation federated learning scheme," *IEEE Transactions on Network Science and Engineering*, 2022.

[28] A. Khamparia, D. Gupta, V. H. C. de Albuquerque, A. K. Sangaiah and R. H. Jhaveri, "Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning," *The Journal of Supercomputing*, vol. 76, no. 11, pp. 8590–8608, 2020.

[29] M. Rizwan, A. Shabbir, A. R. Javed, M. Shabbir, T. Baker *et al.,* "Brain tumor and glioma grade classification using gaussian convolutional neural network," *IEEE Access*, vol. 10, pp. 29731–29740, 2022.

[30] A. R. Javed, M. U. Sarwar, H. U. Khan, Y. D. Al-Otaibi, W. S. Alnumay *et al.,* "Pp-spa: Privacy preserved smartphone-based personal assistant to improve routine life functioning of cognitive impaired individuals," *Neural Processing Letters*, pp. 1–18, 2021.

[31] A. Namburu, D. Sumathi, R. Raut, R. H. Jhaveri, R. K. Dhanaraj *et al.,* "Fpga-based deep learning models for analysing corona using chest X-ray images," *Mobile Information Systems*, vol. 2022, 2022.

[32] G. Salloum and J. Tekli, "Automated and personalized nutrition health assessment, recommendation, and progress evaluation using fuzzy reasoning," *International Journal of Human-Computer Studies*, vol. 151, pp. 102610, 2021.

[33] M. Usman Sarwar, A. Rehman Javed, F. Kulsoom, S. Khan, U. Tariq *et al.,* "Parciv: Recognizing physical activities having complex interclass variations using semantic data of smartphone," *Software: Practice and Experience*, vol. 51, no. 3, pp. 532–549, 2021.

[34] G. Kitamura and C. Deible, "Retraining an open-source pneumothorax detecting machine learning algorithm for improved performance to medical images," *Clinical Imaging*, vol. 61, pp. 15–19, 2020.

[35] Y. H. Chan, Y. Z. Zeng, H. C. Wu, M. C. Wu and H. M. Sun, "Effective pneumothorax detection for chest x-ray images using local binary pattern and support vector machine," *Journal of Healthcare Engineering*, vol. 2018, 2018.

[36] R. W. Filice, A. Stein, C. C. Wu, V. A. Arteaga, S. Borstelmann *et al.,* "Crowdsourcing pneumothorax annotations using machine learning annotations on the nih chest x-ray dataset," *Journal of Digital Imaging*, vol. 33, pp. 490–496, 2020.

[37] H. Maziad, J. A. Rammouz, B. E. Asmar and J. Tekli, "Preprocessing techniques for end-to-end trainable rnn-based conversational system," in *Int. Conf. Web Engineering (ICWE) 2021*, Biarritz, France, Springer, pp. 255–270, 2021.

[38] E. Hariri, M. Al Hammoud, E. Donovan, K. Shah and M. M. Kittleson, "The role of informed consent in clinical and research settings," *Medical Clinics*, vol. 106, no. 4, pp. 663–674, 2022.

[39] A. Dey, "Cov-xdcnn: Deep learning model with external filter for detecting covid-19 on chest x-rays," in *Int. Conf. Computer, Communication, and Signal Processing (ICCCSP) 2022*, Chennai, India, Springer, pp. 174–189, 2022.

[40] S. Tiwari and A. Jain, "Convolutional capsule network for covid-19 detection using radiography images," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 525–539, 2021.

[41] M. Mahdianpari, B. Salehi, M. Rezaee, F. Mohammadimanesh and Y. Zhang, "Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery," *Remote Sensing*, vol. 10, no. 7, pp. 1119, 2018.

[42] A. Sze-To and Z. Wang, "Tchexnet: Detecting pneumothorax on chest x-ray images using deep transfer learning," in *Int. Conf. Image Analysis and Recognition (ICIAR) 2019*, Waterloo, ON, Canada, Springer, pp. 325–332, 2019.

[43] J. X. Luo, W. F. Liu and L. Yu, "Pneumothorax recognition neural network based on feature fusion of frontal and lateral chest x-ray images," *IEEE Access*, 2022.

[44] X. Ouyang, Z. Xue, Y. Zhan, X. S. Zhou, Q. Wang *et al.,* "Weakly supervised segmentation framework with uncertainty: A study on pneumothorax segmentation in chest X-ray," in *Int. Conf. in Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Shenzhen, China, Springer, pp. 613–621, 2019.

[45] M. U. Rehman, A. Shafique, K. H. Khan, S. Khalid, A. A. Alotaibi *et al.,* "Novel privacy preserving non-invasive sensing-based diagnoses of pneumonia disease leveraging deep network model," *Sensors*, vol. 22, no. 2, pp. 461, 2022.

[46] M. Akter, N. Moustafa and T. Lynar, "Edge intelligence-based privacy protection framework for IoTbased smart healthcare systems," in *IEEE INFOCOM 2022-IEEE Conf. on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, pp. 1–8, 2022.

[47] S. Hakak, S. Ray, W. Z. Khan and E. Scheme, "A framework for edge-assisted healthcare data analytics using federated learning," in *2020 IEEE Int. Conf. on Big Data (Big Data)*, Atlanta, GA, USA, IEEE, pp. 3423–3427, 2020.

[48] H. Mughal, A. R. Javed, M. Rizwan, A. S. Almadhor and N. Kryvinska, "Parkinsonâs disease management via wearable sensors: A systematic review," *IEEE Access*, 2022.

[49] H. Mukhtar, S. Rubaiee, M. Krichen and R. Alroobaea, "An IoT framework for screening of covid19 using real-time data from wearable sensors," *International Journal of Environmental Research and Public Health*, vol. 18, no. 8, pp. 4022, 2021.

[50] S. U. Rehman, A. R. Javed, M. U. Khan, M. Nazar Awan, A. Farukh *et al.,* "Personalised comfort: A personalised thermal comfort model to predict thermal sensation votes for smart building residents," *Enterprise Information Systems*, vol. 16, no. 7, pp. 1852316, 2022.

[51] A. R. Javed, R. Faheem, M. Asim, T. Baker and M. O. Beg, "A smartphone sensors-based personalized human activity recognition system for sustainable smart cities," *Sustainable Cities and Society*, vol. 71, pp. 102970, 2021.

[52] S. Batra, H. Sharma, W. Boulila, V. Arya, P. Srivastava *et al.,* "An intelligent sensor based decision support system for diagnosing pulmonary ailment through standardized chest X-ray scans," *Sensors*, vol. 22, no. 19, pp. 7474, 2022.

[53] A. R. Javed, L. G. Fahad, A. A. Farhan, S. Abbas, G. Srivastava *et al.,* "Automated cognitive health assessment in smart homes using machine learning," *Sustainable Cities and Society*, vol. 65, pp. 102572, 2021.

[54] K. Saleem, M. Saleem, R. Zeeshan, A. R. Javed, M. Alazab *et al.,* "Situation-aware bdi reasoning to detect early symptoms of covid 19 using smartwatch," *IEEE Sensors Journal*, 2022.

[55] Y. Chen, X. Qin, J. Wang, C. Yu and W. Gao, "Fedhealth: A federated transfer learning framework for wearable healthcare," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 83–93, 2020.

[56] B. C. Tedeschini, S. Savazzi, R. Stoklasa, L. Barbieri, L. Stathopoulos *et al.,* "Decentralized federated learning for healthcare networks: A case study on tumor segmentation," *IEEE Access*, vol. 10, pp. 8693–8708, 2022.

[57] P. Tam, S. Math and S. Kim, "Optimized multi-service tasks offloading for federated learning in edge virtualization," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 6, pp. 4363–4378, 2022.

[58] H. Shamseddine, J. Nizam, A. Hammoud, A. Mourad, H. Otrok *et al.,* "A novel federated fog architecture embedding intelligent formation," *IEEE Network*, vol. 35, no. 3, pp. 198–204, 2020.

[59] A. Hammoud, A. Mourad, H. Otrok, O. A. Wahab and H. Harmanani, "Cloud federation formation using genetic and evolutionary game theoretical models," *Future Generation Computer Systems*, vol. 104, pp. 92–104, 2020.

[60] S. AbdulRahman, H. Tout, A. Mourad and C. Talhi, "Fedmccs: Multicriteria client selection model for optimal IoT federated learning," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4723–4735, 2020.

[61] S. AbdulRahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi *et al.,* "A survey on federated learning: The journey from centralized to distributed on-site learning and beyond," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5476–5497, 2020.

[62] O. A. Wahab, A. Mourad, H. Otrok and T. Taleb, "Federated machine learning: Survey, multi-level classification, desirable criteria and future directions in communication and networking systems," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1342–1397, 2021.

[63] B. Petrovska, E. Zdravevski, P. Lameski, R. Corizzo, I. Štajduhar *et al.,* "Deep learning for feature extraction in remote sensing: A case-study of aerial scene classification," *Sensors*, vol. 20, no. 14, pp. 3906, 2020.