

A comprehensive review of federated learning for COVID-19 detection

Sadaf Naz  | Khoa T. Phan  | Yi-Ping Phoebe Chen 

Department of Computer Science and Information Technology, School of Engineering and Mathematical Sciences, La Trobe University, Bundoora, Victoria, Australia

Correspondence

Yi-Ping Phoebe Chen, Department of Computer Science and Information Technology, School of Engineering and Mathematical Sciences, La Trobe University, Bundoora, VIC 3086, Australia.
Email: phoebe.chen@latrobe.edu.au

Abstract

The coronavirus of 2019 (COVID-19) was declared a global pandemic by World Health Organization in March 2020. Effective testing is crucial to slow the spread of the pandemic. Artificial intelligence and machine learning techniques can help COVID-19 detection using various clinical symptom data. While deep learning (DL) approach requiring centralized data is susceptible to a high risk of data privacy breaches, federated learning (FL) approach resting on decentralized data can preserve data privacy, a critical factor in the health domain. This paper reviews recent advances in applying DL and FL techniques for COVID-19 detection with a focus on the latter. A model FL implementation use case in health systems with a COVID-19 detection using chest X-ray image data sets is studied. We have also reviewed applications of previously published FL experiments for COVID-19 research to demonstrate the applicability of FL in tackling health research issues. Last, several challenges in FL implementation in the healthcare domain are discussed in terms of potential future work.

KEYWORDS

COVID-19 detection, deep learning, federated learning, machine learning, privacy preservation

1 | INTRODUCTION

The coronavirus disease of 2019 (COVID-19) instigated a global pandemic of viral pneumonia which commenced in late 2019. In the span of a year, there have been more than 123 million cases and more than 2.7 million deaths worldwide. While different parts of the world are at different levels of outbreak, despite early precautionary actions, quality clinical measures and mandatory implementations of public health practices, coronavirus cases are still soaring globally. There is a universal growing urgency to slow the COVID-19 spread by efficient testing and isolation. The research community can help by applying the most advanced artificial intelligence (AI) techniques to generate new insights and methods for COVID-19 detection, which is possible as the significant increase in the number of COVID-19 cases enables a huge amount of relevant data to be collected daily.

With advancements in computer technologies, access to big data, and significant algorithmic developments, machine learning (ML) helps address COVID-19 challenges by refining diagnosis capacity, modelling techniques, and predicting likely epidemics.¹ Traditional ML as shown in Figures 1 and 2A uses manually extracted features that are not only prone to errors but also time-consuming and tedious to develop, particularly in the case of COVID-19-like situations when data is highly sensitive and massively scattered. Instead of manual extraction, deep learning (DL) as shown in Figures 1 and 2B learns hierarchical representations from the data itself and scales better with more data. However, individual COVID-19 data may be scarce for DL analyses. To overcome data scarcity, data integration across scattered locations in a centralized repository is both expensive and complex. Time, resources, and privacy constraints to pool and train such enormously scattered data is a major challenge. While DL addresses ML challenges to learn hidden pattern from COVID-19 data and to build much more efficient decision rules, DL is impractical where data sharing invades the company's privacy in settings such as those where personal data directly affects the owners' privacy, that is, when hospitals would like to protect the privacy of their patients or where competitors are competing for patient populaces, grants, and researchers.²

Recently, there has been an explosion of intelligent devices that are able to collect and process a substantial amount of data, especially in health systems such as personal wearable devices. These devices typically gather data in a private environment, often without the consent and knowledge of the users. Thus, it is crucial to develop a learning technique which trains a model for decentralized data while maintaining privacy. Federated learning (FL), also known as collaborative learning, is

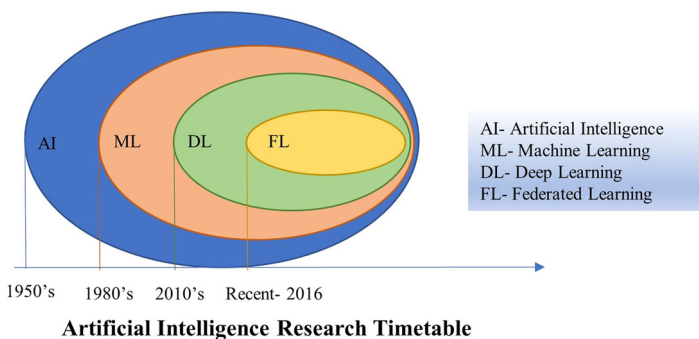


FIGURE 1 The relationship between the subsets of artificial intelligence is shown in the Venn diagram [Color figure can be viewed at wileyonlinelibrary.com]

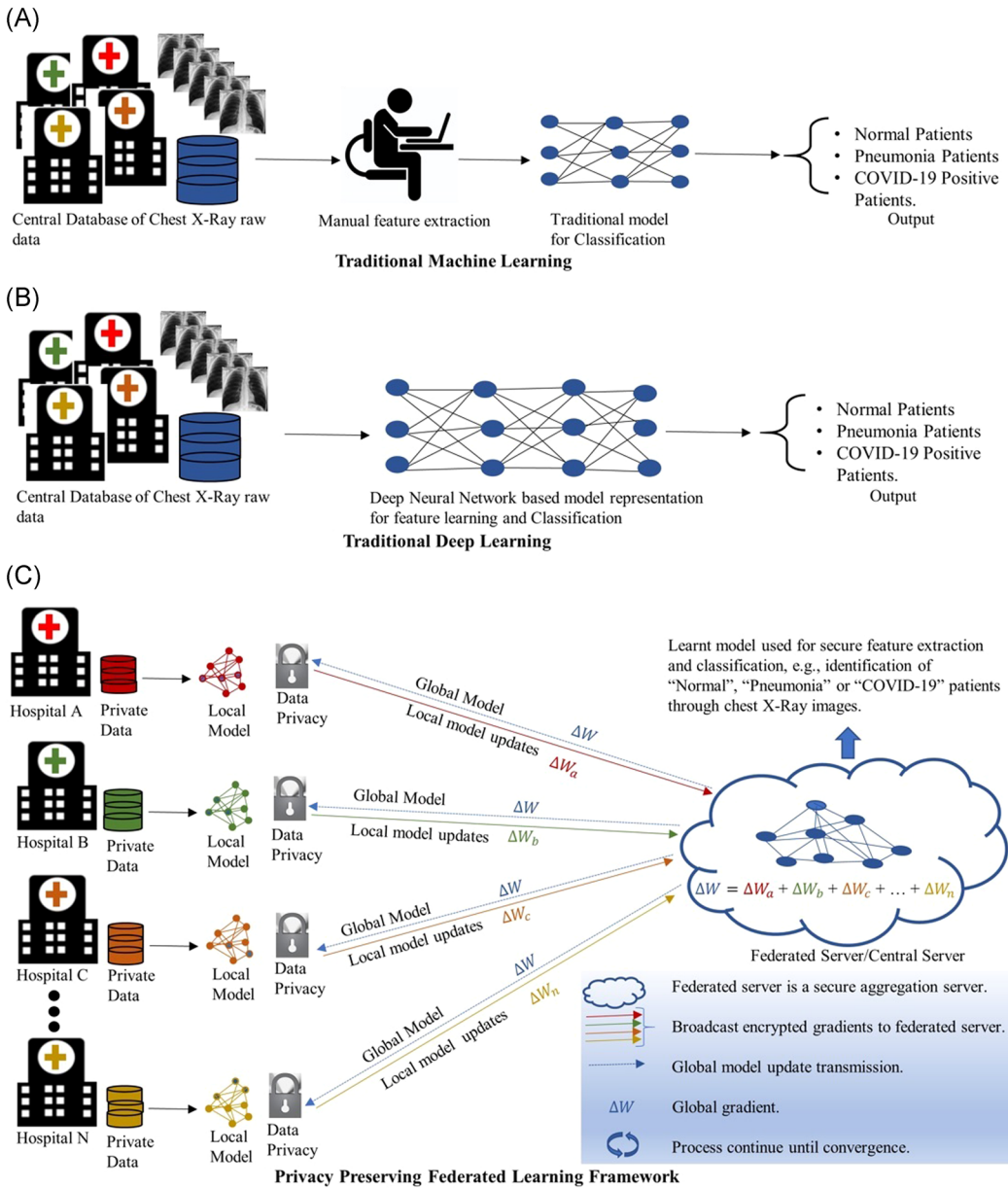


FIGURE 2 Basic structure of (A) traditional ML, (B) traditional DL, and (C) privacy preserving FL framework. DL, deep learning; FL, federated learning; ML, machine learning [Color figure can be viewed at wileyonlinelibrary.com]

such a technique. FL was initially developed for mobiles, the Internet of Things (IoT), and edge devices,² and recently attained popularity in the health domain for data privacy preservation.³⁻⁵ FL allows users to train an algorithm across multiple decentralized databases without sharing their data samples as shown in Figures 1 and 2C. FL is broadly used in the last few years in various fields; however, FL implementation is still a challenging task. Table 1 highlight some of the FL potential risks and benefits. In Yang et al.,⁶ author extensively discussed possible solutions to potential FL

TABLE 1 Potential risks and benefits of federated learning

	Types	Major consequence
Risk of Information leakage	<ul style="list-style-type: none"> • During model training^{7,8} • Through gradient^{7,9} • To central server or third party^{10,11} 	Compromise privacy ⁷⁻⁹
Risk of data poisoning	Data poisoning attacks ¹² can be random or targeted <ul style="list-style-type: none"> • By data¹⁴ • By model^{13,15,16} • By backdoor¹⁷ 	Miss-classification with high confidence ¹³
Risk of model attacks	<ul style="list-style-type: none"> • Insider¹⁸—by the FL server/participants in the FL system • Outsider¹⁹—by the FL server/participants during communication, and by users of the final FL model • At training¹⁵ or inference phase⁷⁻⁹ 	Compromise the integrity of the learning process ¹⁸⁻²¹
Benefits of FL	<ol style="list-style-type: none"> 1. Allow fruitful collaboration with private data. 2. Unlock data sets analysis rarely available to public. 3. Collaboratively analyse sensitive data sets.²²⁻²⁴ 4. Generalizable outputs. 	

challenges. However, still there exist many open issues presented by Google in Kairouz et al.,² which may help set future directions for the researchers.

Currently, there are specifically designed platforms to support FL implementation such as PySyft- a python library for secure DL, TFF (<https://www.tensorflow.org/federated>), FATE (<https://www.fedai.org/cn/>), and Tensor/IO (<https://github.com/doc-ai/tensorio>), which are developed by OpenMined, Google, Webank, and Dow et al. respectively.²⁵

The rest of the manuscript is organized as follows. In Section 2, we briefly review the existing DL results on COVID-19 detection. In Section 3, we present a simple FL implementation with an exemplary COVID-19 detection use case using chest X-ray (CXR) image data sets. Existing works on FL for COVID-19 detection are reviewed in Section 4. In Section 5, we discuss scope of FL in medical research. We discuss the implementation challenges of FL in medical research as a conclusion in Section 6 followed by the future work in Section 7.

2 | DL FOR COVID-19

The clinical symptoms of COVID-19 are mostly a dry cough, fever, chills, and systemic pain although some patients have abdominal manifestations.²² Current COVID-19 detection and classification DL approaches are mainly based on, including but not limited to, pre-scan, laboratory testing and medical image (CXR and computed tomography [CT]) analysis, that is, see Table 2, for more details.

Prescanning approaches^{26,27} analyse coughing and breathing data, which could be a first step in the diagnosis and detection of COVID-19. However, these approaches are not robust and cannot replace clinical testing. The timely infection detection by additional screening and

TABLE 2 Review of various dl approaches for covid-19

	Data	What to predict	Methods	Advantages	Disadvantages
Pre-scan approaches	<ul style="list-style-type: none"> Coughing samples and sound wave data analysed via mobile app Breathing data and thermal videos 	<ul style="list-style-type: none"> Detect and identify COVID-19 	<ul style="list-style-type: none"> Diagnose COVID-19 in respiratory system. Detection of COVID-19 	<ul style="list-style-type: none"> Deep transfer learning based multiclassifier model approach Bidirectional gate recurrent unit (GRU) with attention layer mechanism 	<ul style="list-style-type: none"> Data privacy is not addressed. It cannot not be an alternative to clinical testing. Data privacy is not addressed. Not robust at all as multiple other factors may affect the results, for example, mask variety, temperature, and so on.
Laboratory testing approaches	<ul style="list-style-type: none"> Pathology lab data 	<ul style="list-style-type: none"> Detection of COVID-19 	<ul style="list-style-type: none"> Deep sequencing or polymerase chain reaction-based techniques 	<ul style="list-style-type: none"> Identify virus and diagnose COVID-19 by genomic techniques²⁸ 	<ul style="list-style-type: none"> Data privacy is not addressed. Incorrect sample collection. Diagnosis of false negative.
CXR and/or CT images analysis	<ul style="list-style-type: none"> Chest X-ray images 	<ul style="list-style-type: none"> Detection of COVID-19 	<ul style="list-style-type: none"> CNN 	<ul style="list-style-type: none"> Deep transfer learning based multiclassifier model approach 	<ul style="list-style-type: none"> Data privacy is not addressed. Prescreening app²⁶ could be a first step and possibly fill the gap in the lack of testing. Help distinguish abnormal breathing in many scenarios²⁷
					<ul style="list-style-type: none"> Detect and identify COVID-19 positive and negative states^{29,30} Showed higher accuracy in presence of pneumonia in Brunese et al.²⁹ and identification of COVID-19 positive and negative states in Fátima and Inigo³⁰

(Continues)

TABLE 2 (Continued)

Data	What to predict	Methods	Advantages	Disadvantages
• CXR and CT images	• Detection of COVID-19	• GAN with CNN and ConvLSTM-based DL models	<ul style="list-style-type: none"> • Detected COVID-19 infection. • Tried to address the data deficiency issue by data augmentation³¹ 	<ul style="list-style-type: none"> • Data privacy is not addressed. • Data deficiency issue. • This approach needs refinement to enhance accuracy.
• CXR and CT images	• Detection and classification of COVID-19	• Several traditional DL architectures	<ul style="list-style-type: none"> • Detection and classification of CXR and CT images with bacterial pneumonia, coronavirus, and normal³² 	<ul style="list-style-type: none"> • Data privacy is not addressed. • Data deficiency issue. • Requires more data sets for better performance.
• Chest X-ray images	• Identification of COVID-19	• Commercialized DL based CAD system	<ul style="list-style-type: none"> • Identify patients suspected of having COVID-19³³ 	<ul style="list-style-type: none"> • Data privacy is not addressed. • The centralized processing with a limited number of COVID-19 patients is not feasible for generalization.
• X-ray images	• Detection of COVID-19	• An ANN approach based on Convolutional CapsNet used for detection of COVID-19.	<ul style="list-style-type: none"> • Capsule network effectively performs binary and multiclass feature classification³⁴ 	<ul style="list-style-type: none"> • Data privacy is not addressed. • A huge number of images with an inconsistent scale require hardware resources and enough time for processing.

TABLE 2 (Continued)

Data	What to predict	Methods	Advantages	Disadvantages
<ul style="list-style-type: none"> CXR and CT images 	<ul style="list-style-type: none"> Identification, classification, and quantification of COVID-19 	<ul style="list-style-type: none"> Iteratively pruned DL models. 	<ul style="list-style-type: none"> Built ensembles for identification, classification, and quantification of COVID-19 like disease patterns^{3,5} 	<ul style="list-style-type: none"> Data privacy is not addressed. Improvement required in pruning strategy. Visualization and interpretation of learned model is still unaddressed.

Abbreviations: ANN, artificial neural network; CNN, convolutional neural network; CT, computed tomography; CXR, chest X-ray.

combining the antibody testing with quantitative-polymerase chain reaction (qPCR) can significantly improve detection sensitivity and accuracy,³⁶ however, incorrect sample collection in qPCR or false-negative diagnosis can result in grave consequences by allowing diseased patients to spread the virus. Medical imaging such as CXR and CT scan analysis is one of the most auspicious research fields which facilitates the diagnosis of viral infections like COVID-19.³² In comparison, CT images are more powerful in detecting viral infections however less accessible and a costly test to the public, although CXR images perform the same task with greater accessibility and relatively at a lower cost.³⁷ However, it should be noted that these approaches are unable to efficiently address data privacy concerns,^{29,30} the infeasibility of model generalization due to small data sets,^{31–33} centralized processing,^{32,33} a lack of set criteria for the selection of the most suitable algorithm for a precise problem,

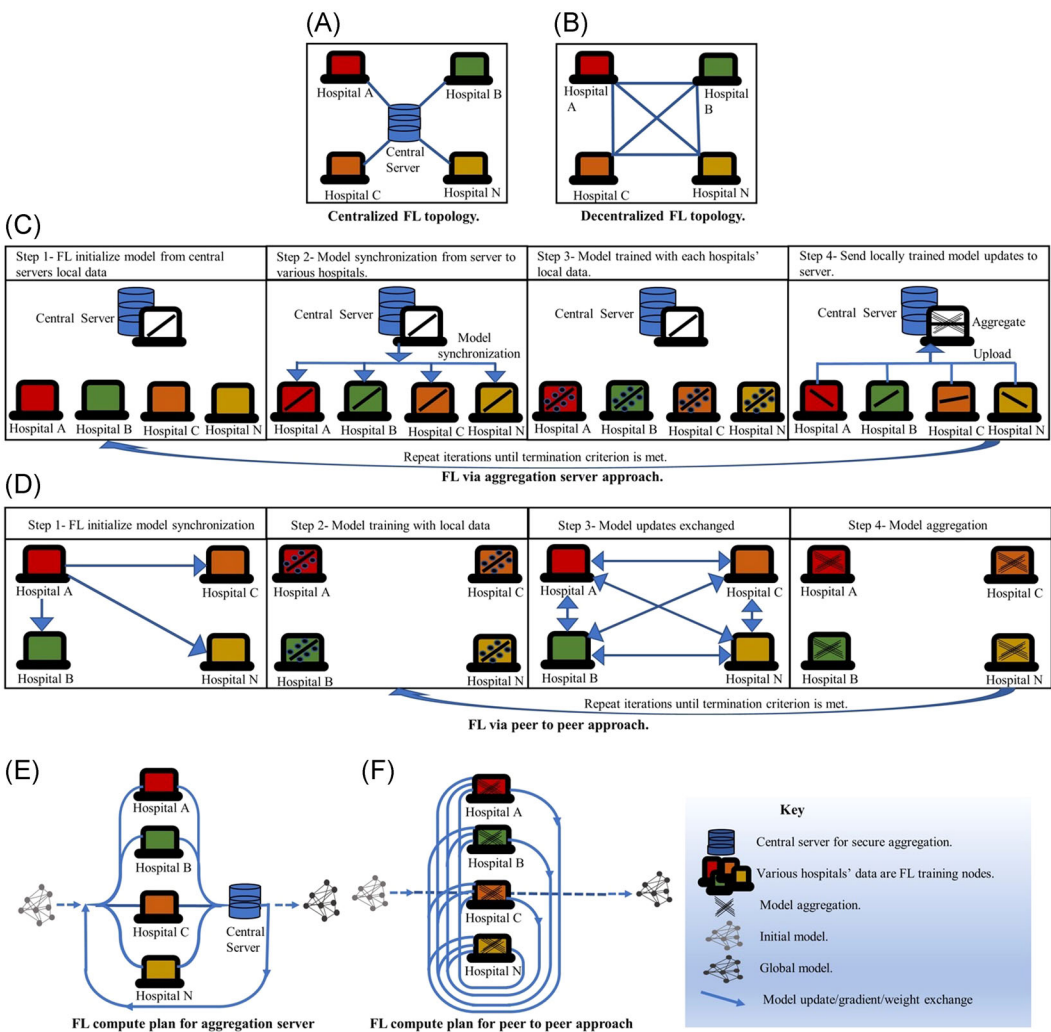


FIGURE 3 FL workflow. (A) Centralized FL topology, (B) decentralized FL topology, (C) FL via aggregation server approach, (D) FL via peer-to-peer approach, (E) FL computation plan for aggregation server, and (F) FL computation plan for peer-to-peer approach. FL, federated learning. [Color figure can be viewed at wileyonlinelibrary.com]

expensive model training, and communications and implementations requirement.³⁴ Here FL is useful in addressing these issues to some extent.

3 | FL SYSTEM

FL is an ML architecture to address the data privacy issue by collaborative training approaches that do not require a single pool of centralised data.

3.1 | A model FL implementation in the healthcare setting

FL is an iterative process as shown in Figures 2C and 3C.

- Step 1: Central server initializes the training model from its local data.
- Step 2: Central server synchronises (or transmits) the model to participating hospitals/clients.
- Step 3: Upon receiving the model from the server, each hospital trains the model locally with their own data samples.
- Step 4: Each hospital returns the locally trained incremental model updates to the central server as shown in Figures 2C and 3E. Then, the central server aggregates the model results and generates a global model without knowing the individual data samples of the hospitals.

The process from Steps 1 to 4 is termed as one FL round. In Step 4, the central server pools all the updated models from the clients and generates the new global model. The activated nodes' generated data are stored and treated locally, and the central server received model updates only. Multiple FL rounds are executed. The central server ends the iteration process when a prespecified termination criterion is met.

Healthcare applications commonly use FL via either aggregation server approaches (Figure 3C) or peer-to-peer approaches (Figure 3D).³⁸ The basic topology and computation plans of FL via the aggregation server and peer-to-peer approaches are presented in Figure 3A-F, respectively. Although FL mainly serves for privacy preservation, where aggregation server approaches ensure participants remain unknown from each other, models subject to conditions retain some information.³⁹ To overcome privacy leakage in the FL framework, differential privacy^{40,41} or encrypted data learning approaches⁴² have been suggested. Peer-to-peer workflow creates connections between all or a subsection of directly linked nodes.⁴³ Overall, FL benefits its stakeholders, such as clinicians, patients, hospitals, AI researchers, pharmaceutical companies, health care providers and software developers.⁴⁴ Overall, FL is an emerging approach to break down barriers to share data between industries while the local data is protected.⁶

3.2 | FL case study for Covid-19 detection

Input: Chest X-ray images data set labelled as

- Normal
- COVID-19

Output: Classification model for the identification of CXR images with COVID-19.

Notations: Let

- N = Total number of hospitals participating in FL model building.
- K = Number of hospitals that participate in every communication round where $K \leq N$.

k th hospital holds n_k training data samples: $x_{k,1}, x_{k,2}, x_{k,3}, \dots, x_{k,n_k}$

- T = Total number of rounds.
- E = Number of local (epochs) iterations performed in a hospital between two communications. Hence, T/E = Number of communications.
- w_0 = Initial model weight.
- w_k = k th hospital's model update.
- b = Local minibatch size.
- η = Learning rate or step size.
- l = Loss function.

Problem formulation: Let l be a user specified global loss function obtained from a weighted combination of K local losses $\{l_k\}_{k=1}^N$, calculated from private data which is stored in the individual hospitals' repository and is never shared between them:

$$\min_{\varphi} l(X; \varphi) \text{ with } l(X; \varphi) = \sum_{k=1}^N w_k l_k(X_k; \varphi)$$

where $w_k \geq 0$ denote the weights of the k th hospital and $\sum_{k=1}^N w_k = 1$.⁴⁴

Pseudo code of FL: The pseudo-code of FL via the aggregation approach (FedAvg with centralized topology) is presented in Table 3, targeting updates from K clients per round.

Communication cost: The centralized aggregation (FedAvg) approach incurs costs in two ways per iteration:

1. The central server transmits the latest model update to all the participating hospitals and then performs local updates.
2. The central server aggregates the outputs from all the participating hospitals.

Learning parameters: There are three key parameters:

1. K , the partial contribution of the hospitals that perform computation on each round.
2. E , the number of local training iterations each hospital makes over its local data set on each round.
3. b , the local minibatch size intended for the client updates.

Client's participation: When there are many participants, the partial participation of the collaborators is more realistic and cost efficient. Users set the threshold K ($1 \leq K < N$). For one iteration, the central server aggregates the output of the first K responded hospitals and stops waiting for the rest.

4 | FL FOR COVID-19

In this section, we review the three most recent results^{22–24} to apply FL for COVID-19 detection, which study medical diagnostic images, for example, CT scan and/or X-ray. The information and comparisons are summarized in Table 4. Overall, the insufficiency of data and privacy concerns are the two main motivations for applying the FL approach in these works.

TABLE 3 Algorithm of FL via aggregation approach (FedAvg with centralized topology)⁴⁵**Server-side execution:**

//Start procedure

Initialize w_0 //Initialize global model**for** each round $t = 1, 2, 3, \dots$ **do** Select K collaborating hospitals to calculate model updates Wait for updates from K hospitals. $(\Delta^k, n^k) =$ Client's update (w) from hospital $k \in [k]$. $\bar{w}_t = \sum_k \Delta^k$ //Sum of weighted updates $\bar{n}_t = \sum_k n^k$ //Sum of weights $\Delta_t = \bar{\Delta}_t / \bar{n}_t$ //Average update $w_{t+1} \leftarrow w_t + \Delta_t$ **Client-side updates (w):** $n \leftarrow |\beta|$ //Update weight $w_{init} \leftarrow w$ **for** batch $b \in \beta$ **do** $w \leftarrow w - \eta \nabla l(w; b)$ $\Delta \leftarrow n \cdot (w - w_{init})$ //Weighted update.

//End procedure

//Procedure stops when some user prespecified criteria is met.

// Δ can be compressed more than w return (Δ, n) to server.

//In a real-world situation, the assumption of independent and identical distribution (i.i.d.) data does not meet. The aggregation step varies in this case.

//The aggregation step also varies in the case of the full or partial participation of the hospitals.

To stress the importance and to motivate further research on FL, experiments are performed on open-source pneumonia CXR and/or CT Images data sets to detect COVID-19. Research data access information, availability and sources are also detailed in Table 4.

Mostly up-to-date COVID-19 data are provided by government organizations organizations.⁴⁶ Open source COVID-19 data sets are in raw text format and are often unstructured. Raw data in the form of comma separated value (CSV) files permit a quick and easy data download yet require substantial data pre-processing is required to prepare it for further analysis. CXR images resize model augmentation technique is adopted in Yan et al.²² for model training, whereas the authors used scaling in Kumar et al.²³ for data preprocessing. Most COVID-19-related research work deals with the binary class (positive or negative) which may incur vagueness for the detection of a disease. For example, Kumar et al.²³ concerns the binary class for recognition of COVID-19, which is unable to detect other viral pneumonia, whereas the multiclass approach in Yan et al.²² provides a better and deeper understanding of data and helps achieve better screening.

TABLE 4 Detailed review of three recent studies of federated learning for COVID-19

Problem identification/ Motivations	<ul style="list-style-type: none"> • Lack of enough data samples. • Data privacy concerns. • Inspire researchers in relation to the novel FL approach. • Data privacy concerns. • Lack of enough training data. • Find patterns via lung screening. 	<ul style="list-style-type: none"> • Insufficient data for training. • Massive communication overhead in the case of heterogeneity of the clients. • Model performance and data privacy at risk.
What want to predict	In Yan et al., ²² the author wants to detect COVID-19 using Pneumonia CXR images	<ul style="list-style-type: none"> • The author of Zhang et al.²⁴ wants to use medical diagnostic image analysis, such as a CT scan or X-ray, to detect COVID-19.
Data source	<ul style="list-style-type: none"> • Open source • GitHub (https://github.com/ieee8023/COVID-chestxray-dataset), • Kaggle and RSNA, Pneumonia Detection Challenges data set. 	<ul style="list-style-type: none"> • Open source • Github (https://github.com/UCSD-A14H/COVID-CT), (https://www.kaggle.com/tawsifurrahman/covid19-radiography-database), (https://github.com/agchung/Figure1-COVID-chestxray-dataset).
Data/Sample size/ Sample details	<ul style="list-style-type: none"> • COVIDx data set (CXR images) labelled as: <ul style="list-style-type: none"> Normal images = 8851 Pneumonia images = 6045 COVID19 images = 386 Total images = 15,282 	<ul style="list-style-type: none"> • COVID-19 CT images data set: <ul style="list-style-type: none"> COVID-19 positive images = 349 COVID-19 negative images = 397 Viral Pneumonia = 0 Total = 746 • COVID-19 CXR images from data set 1 and 2: <ul style="list-style-type: none"> COVID-19 positive images = 219 + 55 COVID-19 negative images = 1341 + 0 Viral pneumonia = 1345 + 0 Total = 2905 + 55 = 2965

TABLE 4 (Continued)

Platform requirement	<ul style="list-style-type: none"> • GPU acceleration of NVIDIA Tesla V100 (32 GB) on Ubuntu 18.04 system. • PyTorch used for model implementation. Grad-CAM++ performed model visual explanations. 	<ul style="list-style-type: none"> • DL models pretrained either on ImageNet or Scratch. • ImageNet is an ongoing project for an easily accessible image database. • Scratch is a visual programming language. 	<ul style="list-style-type: none"> • GPU architecture with different configurations.
Data pre-processing	<ul style="list-style-type: none"> • Chest X-ray images resized to (224,224) model augmentation technique is adopted. 	<ul style="list-style-type: none"> • Discarded the CT scans with discrepancies. 	<ul style="list-style-type: none"> • 3326 images were chosen from the collected data sets and their ratios were adjusted for both types of images.
Specific learning problem	<ul style="list-style-type: none"> • Learn the best image classification DL model by comparing the performance of the models, trained with and without the FL framework. 	<ul style="list-style-type: none"> • Learn the appropriate class on the extracted features with high performance. 	<ul style="list-style-type: none"> • FL's default setting incurs a high communication cost when exchanging model parameters and can barely guarantee model efficiency with heterogeneous data.
Best solution/FL models/methodology	<ul style="list-style-type: none"> • FL approach to address the data shortage issue when performing DL approaches to detect COVID-19. 	<ul style="list-style-type: none"> • Secure data sharing with the integration of FL and blockchain and proposed a modified Capsule Network for feature extraction. 	<ul style="list-style-type: none"> • To optimize communication efficiency and model performance, a dynamic fusion-based technique was suggested.
Input	<ul style="list-style-type: none"> • Training images = 13,703 • Testing images = 1579 	<ul style="list-style-type: none"> • Training images = 22,556 • Testing images = 11,450 	<ul style="list-style-type: none"> • Training images = 2800 • Testing images = 526
Main features	<ul style="list-style-type: none"> • Main task was image classification. • Applied 4 DL models and compared the performance of all 4 models by training with and without FL. • On performance, ResNeXt and ResNet18 are chosen for COVID-19 identification. 	<ul style="list-style-type: none"> • Designed a new FL model by modifying Capsule Network for secured data sharing. • Applied various DL models for training and testing and compared model outputs. • Relative importance of data sample and privacy is well taken care of. • Blockchain technology advance the recognition accuracy. 	<ul style="list-style-type: none"> • Once the network is weak and the model comprises of a lot of parameters, the suggested DF-based solution will greatly reduce training time. • As compared to the default setting of FL (D_FL) for GhostNet, ResNet50, and ResNet101, the figure of uploads for DF decreased 1/3, 1/10, and 1/16 of D_FL time.

(Continues)

TABLE 4 (Continued)

Parameters	<ul style="list-style-type: none"> Number of agents = 5, The proportion of agents participating in the central model update for each round = 0.4, Epochs update per round = 3, Batch size update per round = 10, Learning rate = 2.00E-05, Weight decay = 1.00E-07 	<ul style="list-style-type: none"> The parameters $(x; y)$ of H hospitals are defined as: $VAR[\Delta w_{x,y}] = \frac{1}{H} \sum_{H=0}^H (\Delta w_{x,y}^H - \mu_{x,y})^2$ 	<ul style="list-style-type: none"> Client training epochs are adjusted to 90, and the extreme network speed for model upload/download is set to 10 MB/s. The upload figure and time were composed of the total figure of three clients, which in their case was 30 and 90 times for each client and in total, respectively.
Output	<ul style="list-style-type: none"> ResNet18 showed the best performance both in training with and without FL. ResNeXt has the best performance in images with COVID-19 labels. ResNeXt and ResNet18 are better chosen for COVID-19 identification. 	<ul style="list-style-type: none"> Designed a smart capsule network (FL model) to securely share the medical data with privacy and achieved high performance (0.967 Sensitivity) in detecting diseases in the medical images. A significant performance increase was noticed with increased data providers. 	<ul style="list-style-type: none"> The proposed method is viable and outperforms the default FL setting regarding model accuracy, robustness, communication efficiency and fault tolerance.

In these works, FL models are implemented on PyTorch²² and pretrained on ImageNet and Scratch,²³ where the authors used different configurations of GPU.^{22,24} The data size for training and testing is specified in Table 4. The works [22,23] highlight a trade-off between model accuracy and privacy-preserving but do not consider communication efficiency. In Yan et al.,²² the authors provide visual explanations on the models and highlight the critical regions on the patient's CXR images in addition to generating maps for classification. In contrast, in Kumar et al.,²³ the decentralized blockchain technology is an obvious development in the recognition of DL models. In a blockchain, the integration of differential privacy and FL is complicated as the design lacks clarity because of several opposing features. To mitigate the complexity, the authors propose a theoretical framework to enable differential privacy to COVID-19 CT imaging data using FL. Blockchain technology ensures the traceability of data which helps identify the social connections between people which is a key risk factor in the spread of COVID-19. In Kumar et al.,²³ the authors provide all the technical details of the DL model implementation and achieved enhanced sensitivity for COVID-19 detection from lung CT scans. In contrast to [22,23], the author proposed a novel dynamic fusion-based FL approach to achieve communication efficiency and improved model performance while securing data privacy for COVID-19 identification in Zhang et al.,²⁴ however, the performance is not significant for the models with a simple structure and few parameters.

5 | SCOPE OF FL IN MEDICAL RESEARCH

Machine DL techniques have shown efficiency in tackling a huge amount of curated data to feature millions of parameters to gain precise, unbiased, secure, and generalizable medical grade outputs.⁴⁷⁻⁴⁹ However, high quality full spectrum curated medical data are often hard to obtain such as sensitive and well controlled data.⁵⁰ The collection of such data is challenging and may have substantial business value as it requires significant time, cost, and energy, thus making it improbable to access publicly. Data privacy could be preserved by removing meta data but not anymore,⁵¹ as CT or MRI data can possibly restructure the patient's face.⁵² In such situations, the FL approach comes in handy.

The databases, for example, pathology,⁵³ radiology,⁵⁴ and so on, store a huge amount of medical data, however, such data collaborations are prone to scalability issues, in addition to technical and privacy concerns.⁵⁵ FL is currently gaining popularity in medical research where each institute can hold its data and executes decentralized computing which not only preserves privacy but also captures greater data variability. For example, FL helps to discover patients with similar symptoms,⁵⁶ predicting hospitalizations due to heart diseases,⁴ brain cancer segmentation,⁵ and whole brain segmentation through MRIs,⁵⁷ and smartwatches/smartphones classify human activity using a huge amount of sensor data,⁵⁸ multisite fMRI analysis to classify biomarkers related to disease disease,⁵⁹ breast density classification based on breast imaging,⁶⁰ and so on. Recent FL-based model approaches comparatively perform better than ML traditional approaches which either require centralized or single-sited data.⁵ There is a huge scope in this field as limited research has been conducted on FL so far in healthcare settings.

6 | CONCLUSION

COVID-19 has brought unprecedented challenges. However, FL has been promising in solving the issues related to COVID-19 detection and classification as aforementioned reviewed.

However, as a conclusion of our review paper, we are listing several challenges must be addressed before FL can be applied more broadly.

1. *Shortage of data:* Analysing multisite data without pooling is an inherent ability of FL which helps solve the data shortage issue to some extent. However, better model training largely depends on data quality, bias, and scalability.⁴⁷ Some problems are general, such as, the shortage of quality data, data cluttering and a lack of efficiency issues within the healthcare system. In such cases, sample results cannot be generalized. There are a few specific data-related problems in the COVID-19 situation. Only lab-confirmed COVID-19 infections are agreed to be confirmed cases. A limited diagnostic capacity and a shortage of testing kits are a major problem mostly in low-income countries. To produce generalizable results, the availability and access to biased data which share similar demographics, device brands and environment is a challenging part of health care research.
2. *Data heterogeneity issues:* Multi-institutional collaboration causes data standardization problems. The harmonization of heterogeneous COVID-19 data requires preprocessing such as data scaling,^{22,23} resizing of images,²³ resizing of model augmentation,²² and so on, to make it compatible for FL analysis. Intrinsically, traditional FL frameworks are designed for balanced data, that is, each institution consists of the same amount of data, which is typically not feasible in the COVID-19 situation. The FL algorithm, FedAvg, is likely to fail under an extremely skewed data distribution.² Due to data imbalance, the FL model experiences accuracy degradation as observed in several studies.^{61,62} Although a few novel FL frameworks which cater for such imbalanced data⁶³ are favourable, more researchers are encouraged to explore FL further.
3. *Communication overhead:* Naively, the synchronization procedure of FL model training from distributed data entails uplink (user to server) and downlink (server to user) communication.²⁵ In general, model performance is directly proportional to the number of users who participated in training, and the computation and communication overhead.⁶² Communication efficiency is discussed in very few COVID-19 research studies²⁴ and is not considered in most.^{22,23} A huge communication overhead is reported in other areas of research^{62,64} and an effort to reduce communication overhead while preserving data privacy is also reported in Xia et al.⁶⁵
4. *Trade-off between privacy and performance:* A trade-off between FL model accuracy and privacy has been observed not only in COVID-19 research^{22,23} but also in other fields.^{64,66} Better quality data is fundamental to achieve the optimized performance of the model. Ensuring secure access to several organizations to find relevant data for FL model training is a challenging task and may greatly affect model performance.
5. *Privacy leakage issue:* FL naturally promises secure collaboration; however, it does not tick all the boxes to provide guaranteed privacy. Healthcare data collection is directly linked with the augmented risk of privacy leakage. Moreover, the FL training process based on shared information is largely at risk of leakage by model gradients, reverse engineering of model updates, model manipulation, and so on. Data leakage issues have been reported in multiple studies.^{9,52} Patient's information can be back-tracked from the shared gradient.⁶⁷ Research addressing this problem was reported in Kumar et al.,²³ however, more secure FL frameworks⁶² are encouraged for COVID-19-like sensitive research areas. Some untapped counter steps are required to secure data privacy, which makes it an active research area.²

6. *Mutual trust issues*: FL systems collaborate with decentralized parties either in trusted or nontrusted relationships. Trusted collaboration is a kind of standard collaboration with enforceable agreements and set principles, and vice-versa in nontrusted collaboration. Nontrusted collaboration provides a broad spectrum of information, however, it introduces risks, such as privacy concerns, integrity execution, model encryption, malicious attacks, and so on. A strong trusted collaboration is vital in the health care setting, particularly in light of the COVID-19 situation, where each party is not only concerned about the privacy of their patients but they also want to keep information from their business rivals or from the general public to avoid panic. The FL collaborative mechanism requires either a trustworthy third party to play the part of overall controller or stricter mutually agreed protocols, both of which involve extra cost and effort.
7. *Low participation issues*: COVID-19 data is either stored in data centres or sits in data silos, where all users are almost always available. The low participation issue is mostly reported in cross-device FL,² such as wireless communication and IoT settings. The federated averaging procedure by default takes into equal consideration the likely contribution of each user to complete one round, which is sometimes not feasible in practice. Users may not participate sufficiently in the FL process for several reasons, such as low battery power, poor connection, and so on. The low participation issue during FL model training has been highlighted in several studies.^{2,68}
8. *Reliability issues*: A user's reliability depends on its availability to participate in a round of computation for FL model training. Data-centre distributed learning and cross-silo FL results are relatively reliable as both face few dropouts, whereas cross-device FL may produce highly unreliable output as more than 5% of dropouts are likely in a round of computation.² Healthcare collaborators equipped with strong computational resources and advanced systems for better model training are considered relatively more reliable.⁴⁴
9. *Traceability and accountability issues*: The traceability of resources is mandatory in FL systems which includes data access history, training structure, hyperparameter selection, and modifications, and so on. Once the optimality of the model is achieved, traceability and accountability determine the level of contribution of the participants to give them relevant compensation and build a revenue model.⁶⁹ Traceability and accountability may help researchers in explaining and interpreting a global model by investigating the data source from which the models are being trained, where each user can view its own raw data with intra-node security imposed. Issues related to the traceability of FL training data records are discussed and addressed in a few research studies.⁷⁰
10. *Implementation/System architecture issues*: FL system implementation is a significant task in the healthcare setting which faces challenges, but somehow all are manageable. The continuous efforts of researchers make it certainly more surmountable. The healthcare setting holds high-throughput relatively reliable data whose model training requires more communication rounds and more local training steps, and carries with it certain challenges, such as data integrity, communication with redundant nodes, data leakage prevention, reduction in model training time, and so on.⁴⁴ Fortunately, to stay ahead of the curve when implementing FL, we use readily available resources, for example, TensorFlow, a free and open-source platform or PyTorch, a free and open-source ML library. There are still significant nontechnical challenges linked with the healthcare setting, such as health protocols, intellectual property, legal and agreement issues.

7 | SUMMARY AND FUTURE WORK

The medical sector produces an enormous amount of data which is not being fully exploited by MLs yet.⁴⁴ Privacy concerns demand that medical data be stored in data silos where the sedentary behavior of data prevents ML approaches from unleashing their full potential. FL, a promising approach in ML, is a true definition of global collaboration. FL efficient and robust models exploit sensitive medical data stored across different health care institutions without accessing or decentralizing the actual data and help to improve diagnosis and drug discovery which eventually improves patient care worldwide. Since the beginning of the COVID-19 situation, FL has been used by researchers and industry not only for the detection and identification of coronavirus patients, but also for timely and cost-efficient drug discovery, data privacy, data fairness, optimization, statistical solutions, and cryptography. FL is an interesting and growing research topic in recent times² and a revolutionizing collaborative learning approach for training ML models.

Few FL reviews have been published recently. Current reviews covers diverse fields, for example, potential general privacy preservation techniques which could be implemented in an FL setting are reviewed in Yang et al.,⁶ a detailed discussion of recent advances and open problems are surveyed in Kairouz et al.,² FL system heterogeneity is reviewed in Kairouz et al.,² personalization techniques for the FL setting are surveyed in Kulkarni et al.,⁷¹ potential threats to FL models are reviewed in Lyu et al.,⁷² applications in FL are reviewed in Li et al.⁷³ and Rehouma et al.,⁷⁴ FL blockchain with a particular focus on the in vitro fertilization (IVF) field is reviewed in Hickman et al.⁷⁵ and a comprehensive survey is conducted on mobile edge networks in Lim et al.⁷⁶ We reviewed FL as a crucial AI framework, envisioned the scope of FL research in healthcare including but not limited to COVID-19, and highlighted the main FL challenges in the health sector particularly in COVID-19-like situations. Our work aims to motivate researchers to help build a more secure FL setup which is compliant with ethical data handling.

Although, FL has a potential impact on health care at a global level, not all the technicalities of this approach have been efficiently addressed yet, but it is safe to assume that FL will be a dynamic research area in the following years.² In the future, we will continue our significant interest in exploring FL capabilities in healthcare settings over the wireless network.

Here we summarize several unaddressed problems in the FL setting to give future directions to researchers:

- Availability and accessibility of quality data that share similar demographics and environments to produce more generalizable results.
- Legal, regulatory, or ethical issues that may encourage or coerce the use of FL.
- Business issues that will possibly inspire or constrain the use of FL.
- Electronic Health Records to help build a prediction model for patients' readmission risk while keeping patients' information secure.

Regardless of a few technical restrictions, we strongly believe that FL has a promising impact on improving health care. We hope this review motivates and helps to scope FL research, including but not limited to the COVID-19 situation.

ORCID

Sadaf Naz  <https://orcid.org/0000-0002-0267-8371>

Khoa T. Phan  <https://orcid.org/0000-0003-0471-9402>

Yi-Ping Phoebe Chen  <https://orcid.org/0000-0002-4122-3767>

REFERENCES

1. van der Schaar M, Alaa AM, Floto A, et al. How artificial intelligence and machine learning can help healthcare systems respond to COVID-19. *Int J Mach Learn Cybern*. 2020;110:1-14.
2. Kairouz P, McMahan H, Avent B, et al. Advances and open problems in federated learning. *Found Trends Mach Learn*. 2021;14(1-2):1-210.
3. Eysenbach G, Luo Y, Noman M, et al. Privacy-preserving patient similarity learning in a federated environment: development and analysis. *JMIR Med Inf*. 2018;6(2):e20. doi:10.2196/medinform.7744
4. Brisimi TS, Chen R, Mela T, Olshevsky A, Paschalidis IC, Shi W. Federated learning of predictive models from federated Electronic Health Records. *Int J Bio-Med Comput*. 2018;112:59-67.
5. Li W, Milletari F, Xu D, et al. Privacy-preserving federated brain tumour segmentation. In: *Conference Proceedings 10th International workshop on MLMI*, Shenzhen, China, Vol. 11861; 2019:133-141.
6. Yang Q, Liu Y, Chen TJ, Tong YX. Federated machine learning: concept and applications. *ACM Trans Intell Syst Technol*. 2019;10(2):1-19.
7. Zhu L, Han S. Deep leakage from gradients. Springer; 2020:17-31.
8. Melis L, Song C, De Cristofaro E, Shmatikov V. Exploiting unintended feature leakage in collaborative learning. In: *Conference Proceedings of 2019 IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA; 2019:691-706.
9. Phong LT, Aono Y, Hayashi T, Wang L, Moriai S. Privacy-preserving deep learning via additively homomorphic encryption. *IEEE Trans Inf Foren Sec*. 2018;13(5):1333-1345.
10. Agarwal N, Suresh AT, Felix XY, Kumar S, McMahan B. cpSGD: Communication-efficient and differentially-private distributed SGD. In: *Conference Proceedings of 32nd Conference on NeurIPS*, Montréal, Canada; 2018:7575-7586.
11. McMahan HB, Ramage D, Talwar K, Zhang L. Learning differentially private recurrent language models. In: *Conference Proceedings of 6th International Conference on Learning Representations*, Vancouver, BC, Canada; 2018.
12. Mahloujifar S, Mahmoody M, Mohammed A. Data poisoning attacks in multi-party learning. In: Kamalika C, Ruslan S, eds. *Proceedings of the 36th International Conference on Machine Learning*, California, 2019:4274-4283.
13. Bhagoji AN, Chakraborty S, Mittal P, Calo S. Analyzing federated learning through an adversarial lens. In: *Proceedings of the 36th International Conference on Machine Learning*, California; 2019:634-643.
14. Liu X, Li H, Xu G, Chen Z, Huang X, Lu R. Privacy-enhanced federated learning against poisoning adversaries. *IEEE Trans Inform Forensics Security*. 2021;16:4574-4588.
15. Zhou X, Xu M, Wu Y, Zheng N. Deep model poisoning attack on federated learning. *Future Internet*. 2021; 13(3):73.
16. Chen Z, Tian P, Liao W, Yu W. Towards multi-party targeted model poisoning attacks against federated learning systems. *High-Confidence Computing*. 2021;1(1):100002.
17. Zhang J, Chen B, Cheng X, Binh HTT, Yu S. PoisonGAN: generative poisoning attacks against federated learning in edge computing systems. *IEEE Internet of Things J*. 2021;8(5):3310-3322.
18. Pan X, Zhang M, Wu D, Xiao Q, Ji S, Yang Z, Justinian's GA. Averno: robust distributed learning with gradient aggregation agent. In: *Conference Proceedings of 29th USENIX Security Symposium*, Boston, USA; 2020:1641-1658.
19. Bagdasaryan E, Veit A, Hua Y, Estrin D, Shmatikov V. How to backdoor federated learning. In: *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. Palermo, Sicily, Italy; 2020:2938-2948.
20. Biggio B, Nelson B, Laskov P. Support vector machines under adversarial label noise. In: *Conference Proceedings of The 3rd Asian Conference on Machine Learning*, Taoyuan, Taiwan; 2011:97-112.
21. Barreno M, Nelson B, Sears R, Joseph AD, Tygar JD. Can machine learning be secure? In: *Conference Proceedings of the 2006 ACM Symposium on Information, computer and communications security*, Taipei, Taiwan, 2006:16-25.
22. Yan B, Wang J, Cheng J, et al. Experiments of federated learning for COVID-19 chest X-ray images. In: *International Conference on Artificial Intelligence and Security (ICAIS)*, Dublin, Ireland; 2021:41-53.
23. Kumar R, Khan AA, Zhang S, Wang W, Abuidris Y, Amin W, Kumar J. Blockchain-federated-learning and deep Learning models for COVID-19 detection using CT Imaging. *IEEE Sens J*. 2021;21(14):16301-16314.

24. Zhang W, Zhou T, Lu Q, et al. Dynamic fusion-based federated learning for COVID-19 detection. *IEEE Internet Things J.* 2021;21(14):16301-16314.
25. Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated Learning for Healthcare Informatics. *J Healthc Inform Res.* 2020;5:1-19.
26. Imran A, Posokhova I, Qureshi HN, et al. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Inform Med.Unlocked.* 2020;20:100378.
27. Jiang Z, Hu M, Gao Z, et al. Detection of respiratory infections using RGB-infrared sensors on portable device. *IEEE Sens J.* 2020;20(22):13674-13681.
28. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med.* 2020;382(13):1199-1207.
29. Brunese L, Mercaldo F, Reginelli A, Santone A. Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Comput Meth Prog Biomed.* 2020;196:105608.
30. Fátima S, Iñigo B. COVID-19 detection in chest X-ray images using a deep learning approach. *Int J Interact.* 2020;6(2):1-4.
31. Sedik A, Iliyasa AM, Abd El-Rahiem B, et al. Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections. *Viruses.* 2020;12(7):769.
32. El Asnaoui K, Chawki Y. Using X-ray images and deep learning for automated detection of coronavirus disease. *J Biomol Struct Dyn.* 2020:1-12.
33. Hwang EJ, Kim H, Yoon SH, Goo JM, Park CM. Implementation of a deep learning-based computer-aided detection system for the interpretation of chest radiographs in patients suspected for COVID-19. *Korean J Radiol.* 2020;21(10):1150-1160.
34. Toraman S, Alakus TB, Turkoglu I. Convolutional capsnet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos, Soliton. Fract.* 2020;140:110122.
35. Rajaraman S, Siegelman J, Alderson PO, Folio LS, Folio LR, Antani SK. Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays. *IEEE Access.* 2020;8:115041-115050.
36. Guo L, Ren L, Yang S, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). *Clin Infect Dis.* 2020;71(15):778-785.
37. Varela-Santos S, Melin P. A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays using texture features and neural networks. *Inf Sci.* 2021;545:403-414.
38. Chang K, Balachandar N, Lam C, et al. Distributed deep learning networks among institutions for medical imaging. *J Am Med Inform Assn.* 2018;25(8):945-954.
39. Anees A, Chen YPP. Discriminative binary feature learning and quantization in biometric key generation. *Pattern Recogn.* 2018;77:289-305.
40. Abadi M, Chu A, Goodfellow I, et al. Deep learning with differential privacy. In: *Conference Proceedings 2016 ACM SIGSAC Conference on Computer and Communications Security*, Vienna, Austria; 2016:308-318.
41. Shokri R, Shmatikov V. Privacy-preserving deep learning. In: *Conference Proceedings 22nd ACM SIGSAC Conference on Computer and Communications Security*, Denver, Colorado, 2015 **Oct.**215:1310-1321.
42. Hao M, Li H, Luo X, Xu G, Yang H, Liu S. Efficient and privacy-enhanced federated learning for industrial. *Artif Intell IEEE Trans Ind Inform.* 2020;16(10):6532-6542.
43. Pappas C, Chatzopoulos D, Lalis S, Valalis M. IPLS: a framework for decentralized federated learning. In: *2021 IFIP Networking Conference*, Espoo and Helsinki, Finland; 2021:1-6.
44. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. *NPJ Digit. Med.* 2020; 3(1):1-7.
45. Bonawitz K, Eichner H, Grieskamp W, et al. Towards federated learning at scale: system design. In: *Conference Proceedings of 2nd SysML Conference*, California, 2019:Online.
46. Agapito G, Zucco C, Cannataro M. COVID-WAREHOUSE: a data warehouse of Italian COVID-19, pollution, and climate data. *Int J Environ Res Public Health.* 2020;17(15):1-22.
47. Wang F, Casalino LP, Khullar D. Deep learning in medicine—promise, progress, and challenges. *JAMA Intern Med.* 2019;179(3):293-294.
48. Amin N, McGrath A, Chen YPP. Evaluation of deep learning in non-coding RNA classification. *Nature Mach Intell.* 2019;1:246-256.
49. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med.* 2018;24(9):1342-1350.

50. van Panhuis W, Paul P, Emerson C, et al. A systematic review of barriers to data sharing in public health. *BMC Public Health*. 2014;14(1):1-9.
51. Luc R, Julien MH, Yves-Alexandre de M. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun*. 2019;10(1):1-9.
52. Schwarz CG, Kremers WK, Therneau TM, et al. Identification of anonymous MRI research participants with face-recognition software. *N Engl J Med*. 2019;381(17):1684-1686.
53. Borovec J, Kybic J, Arganda-Carreras I, et al. ANHIR: automatic non-rigid histological image registration challenge. *IEEE (T-MI)*. 2020;39(10):3042-3052.
54. Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE (T-MI)*. 2015;34(10):1993-2024.
55. Chen M, Qian Y, Chen J, Hwang K, Mao S, Hu L. Privacy protection and intrusion avoidance for cloudlet-based medical data sharing. *IEEE Trans Cloud Comput*. 2020;8(4):1274-1283.
56. Eysenbach G, Luo Y, Noman M, et al. Privacy-preserving patient similarity learning in a federated environment: development and analysis. *JMIR Med Inf*. 2018;6(2):e7744.
57. Knolle M, Kaissis G, Jungmann F, et al. Efficient, high-performance semantic segmentation using multi-scale feature extraction. *PLOS One*. 2021;16(8):e0255397.
58. Sozinov K, Vlassov V, Girdzijauskas S Human activity recognition using federated learning. In *Conference Proceedings of IEEE Intl Conf on SPA, UCC, BDCloud, SocialCom, SustainCom, Melbourne, Australia*; 2018:1103-1111.
59. Li X, Gu Y, Dvornek N, Staib LH, Ventola P, Duncan JS. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Med Image Anal*. 2020;65:101765.
60. Roth H, Chang K, Singh P, et al. Federated learning for breast density classification: a real-world implementation. In: *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. Springer; 2020:181-191.
61. Amiri MM, Gunduz D. Federated learning over wireless fading channels. *IEEE Trans Wirel Commun*. 2020;19(5):3546-3557.
62. Xu G, Li H, Liu S, Yang K, Lin X. VerifyNet: secure and verifiable federated learning. *IEEE Trans Inf Foren Sec*. 2019;15(99):911-926.
63. Huang X, Ding Y, Jiang Z, Shuhan Q, Wang X. DP-FL: a novel differentially private federated learning framework for the unbalanced data. *W3J*. 2020;23(4):2529-2545.
64. Sattler F, Wiedemann S, Muller K-R, Samek W. Robust and communication-efficient federated learning from non-i.i.d. data. *IEEE Trans Neural Netw Learn Syst*. 2020;31(9):3400-3413.
65. Xia W, Quek TQS, Guo K, Wen W, Yang HH, Zhu H. Multi-armed bandit based client scheduling for federated learning. *IEEE Trans Wirel Commun*. 2020;19(11):7108-7123.
66. Liu Y, Yu JJQ, Kang J, Niyato D, Zhang S. Privacy-preserving traffic flow prediction: a federated learning approach. *IEEE Internet Things J*. 2020;7(8):7751-7763.
67. Zhu L, Liu Z, Han S. Deep leakage from gradients. In: *Conference Proceedings of 33rd Conference on Neural Information Processing Systems*, Vancouver, Canada, 2019:14774-14784.
68. Li T, Sahu AK, Zaheer M, Sanjabi M, Talwalkar A, Smithy V. FedDANE: a federated newton-type method. In: *Conference Proceedings of 53rd Asilomar Conference on Signal, Systems, and Computers*, Pacific Grove, CA; 2019:1227-1231.
69. Ghorbani A, Zou J. Data shapley: equitable valuation of data for machine learning. In: *Conference Proceedings of 36th International Conference on Machine Learning*, Long Beach, CA; 2019:2242-2251.
70. Nasr M, Shokri R, Houmansadr A. Comprehensive privacy analysis of deep learning: passive and active white-box inference attacks against centralized and federated learning. In: *Conference Proceedings IEEE Symposium on SP*, San Francisco, CA; 2019:739-753.
71. Kulkarni V, Kulkarni M, Pant A. Survey of personalization techniques for federated learning. In: *Conference Proceedings of Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, London, UK; 2020:794-797.
72. Lyu L, Yu H, Zhao J, Yang Q. *Threats to federated learning*. Springer; 2020:3-16.
73. Li L, Fan Y, Tse M, Lin K-Y. A review of applications in federated learning. *Comput Ind Eng*. 2020;149:106854.

74. Rehouma R, Buchert M, Chen YPP. Machine learning for medical imaging based COVID19 detection and diagnosis. *Int J Intell Syst.* 2021;36:5085-5115.
75. Hickman CFL, Alshubbar H, Chambost J, et al. Data sharing: using blockchain and decentralized data technologies to unlock the potential of artificial intelligence: what can assisted reproduction learn from other areas of medicine? *Fertil Steril.* 2020;114(5):927-933.
76. Lim WYB, Luong NC, Hoang DT, et al. Federated learning in mobile edge networks: a comprehensive survey. *IEEE Commun Surv Tutor.* 2020;22(3):2031-2063.

How to cite this article: Naz S, Phan KT, Chen Y-PP. A comprehensive review of federated learning for COVID-19 detection. *Int J Intell Syst.* 2022;37:2371-2392.

doi:[10.1002/int.22777](https://doi.org/10.1002/int.22777)