When Collaborative Federated Learning Meets Blockchain to Preserve Privacy in Healthcare

Zakaria Abou El Houda, *Member, IEEE,* Abdelhakim Senhaji Hafid, *Member, IEEE,* Lyes Khoukhi, *Senior Member, IEEE,* and Bouziane Brik *Member, IEEE*

Abstract—**Data-driven Machine and Deep Learning (ML/DL) is an emerging approach that uses medical data to build robust and accurate ML/DL models that can improve clinical decisions in some critical tasks (*e.g.*, cancer diagnosis). However, ML/DL-based healthcare models still suffer from poor adoption due to the lack of realistic and recent medical data. The privacy nature of these medical datasets makes it difficult for clinicians and healthcare service providers, to share their sensitive data (*i.e.*, Patient Health Records (PHR)). Thus, privacy-aware collaboration among clinicians and healthcare service providers is expected to become essential to build robust healthcare applications supported by next-generation networking (NGN) technologies, including Beyond sixth-generation (B6G) networks. In this paper, we design a new framework, called HealthFed, that leverages Federated Learning (FL) and blockchain technologies to enable privacy-preserving and distributed learning among multiple clinician collaborators. Specifically, HealthFed enables several distributed SDN-based domains, clinician collaborators, to securely collaborate in order to build robust healthcare ML-based models, while ensuring the privacy of each clinician participant. In addition, HealthFed ensures a secure aggregation of local model updates by leveraging a secure multiparty computation scheme (*i.e.*, Secure Multiparty Computation (SMPC)). Furthermore, we design a novel blockchain-based scheme to facilitate/maintain the collaboration among clinician collaborators, in a fully decentralized, trustworthy, and flexible way. We conduct several experiments to evaluate HealthFed; in-depth experiments results using public Breast Cancer dataset show the efficiency of HealthFed, by not only ensuring the privacy of each collaborator's sensitive data, but also providing an accurate learning models, which makes HealthFed a promising framework for healthcare systems.**

Index Terms—**Healthcare; Federated Learning; B6G; SDN; Blockchain.**

I. Introduction

BREAST cancer is considered as one of the most severe and aggressive diseases with a low median survival; it has a major and severe impact on society across the world. The number of cancer cases is growing rapidly with an estimation of 29.5 million new cancer cases per year by the end of 2040 [1]. Traditional and laboratory diagnostics are facing several challenges due to the huge costs and lack of accurate and timely decisions/diagnostics. Also, some of these diagnostics are painful and have shown several side-effects on the patients. Thus, novel capabilities need to be carefully designed for an early/accurate diagnosis (*i.e.*, fewer mistakes) of these diseases, while maintaining low costs. The growing in machine and deep learning (ML/DL) field has led to a disruptive innovation in many fields, such as vision and speech [2], including healthcare. Thus, healthcare has largely adopted DL algorithms to efficiently and timely detect diseases (*e.g.*, Breast cancer) in their early stages [3]–[7]. However, ML/DL-based healthcare applications still suffer from poor adoption due to the lack of realistic medical datasets. Indeed, the privacy nature of these medical datasets makes it difficult for clinicians and healthcare service providers, to share their sensitive data, for the development of effective and robust ML/DL models, that can effectively/timely handle disease diagnosis for smart healthcare [8].

In the field of Machine Learning (ML), Federated learning (FL) has attracted a lot of attention from both industrial and academic communities, due to its ability to train a particular model, across multiple distributed healthcare stakeholders without centrally sourcing their medical/sensitive data. In a FL setting, each involved clinician builds a local learning model, leveraging its own medical data. Then, it sends only small model updates to a central node. Therefore, each clinician collaborator defines its own data governance and associated privacy requirements. FL can highly improve the communication overhead as well as ensure the data privacy. FL is expected to be highly applied in privacy-aware industry systems, due to its collaborative learning as well as privacy preserving [9]–[14]. FL opens the door for a novel research path on rare/new diseases, where the data is very limited to some institutions. Thus, FL will allow major healthcare institutions to collaboratively build an accurate and efficient model, in a fully distributed way, while preserving the privacy of each healthcare institution's sensitive data. FL has numerous privacy advantages in comparison to the centralized setting/training. However, it suffers from reverse-engineering attacks that may extract sensitive information from only the locally computed updates [15]. To alleviate this issue, HealthFed is based on a new secure aggregation approach to securely aggregated local models. Besides, Software defined networks (SDN) and blockchain have emerged as key enablers of smart networking in Beyond sixth-generation (B6G) [16]. SDN is a new concept of network programmability that is expected to improve the management of network resources [17]–[21], by deploying centralized nodes, called SDN controller. Besides, blockchain and smart contracts have shown their effectiveness in achieving

Z. Abou El Houda and is with the Department of Computer Science and Operational Research, University of Montreal, Canada, e-mail: (zakaria.abou.el.houda@umontreal.ca).

A. S. Hafid is with the Department of Computer Science and Operational Research, University of Montreal, Canada, e-mail: (ahafid@iro.umontreal.ca).

L. Khoukhi is with the GREYC CNRS, ENSICAEN, Normandie University, France: (lyes.khoukhi@ensicaen.fr).

B. bouziane is with the DRIVE EA1859, university of Bourgogne Franche-Comté, France, e-mail: (bouziane.brik@u-bourgogne.fr).

security, transparency, decentralization and trustworthiness in major industry segments (*e.g.*, Internet of Things (IoT) [22]–[24]). Blockchain addresses the challenges related to the traditional healthcare management systems (*e.g.*, single-point-of-failure). We believe that the joint of SDN, FL, and BC will change the healthcare landscape. In this paper, we exploit the privacy-preserving of FL, the programmability of SDN and the decentralized of blockchain to design a privacy-aware collaborative learning between SDN domains (*i.e.*, clinician collaborators), that is fully trustworthy, decentralized, and efficient.

First, we implemented HealthFed using pysyft [25]. Then, we deployed HealthFed on Ropsten official Ethereum test network [26]. We leverage the well-known Breast Cancer dataset [27], to validate HealthFed, that contains digitized images of a fine needle aspirate of a breast mass; it contains two classes: (1) Benign (B): it represents data samples of patients who were cancer-free; and (2) Malignant (M): it represents data samples of patients who have been diagnosed with cancer.

The main contributions of this work are as follows:

- We propose a novel secure framework (HealthFed) for healthcare system, that allows multiple SDN-based domains, referred to as clinician collaborators, to securely collaborate in order to build robust healthcare-based model, while ensuring the privacy of each clinician.
- We design a novel secure aggregation approach of local learning models, based on Secure Multiparty Computation (SMPC).
- We develop a blockchain based scheme to facilitate/maintain the collaboration among clinician collaborators in a fully decentralized, trustworthy, and flexible way.
- We evaluate and validate HealthFed in terms of several metrics (*e.g.*, Accuracy and F1 score); we also compare it with centralized approaches using a well-known Breast Cancer dataset. The in-depth experiments results confirm that HealthFed achieves privacy and high accuracy/F1 score.

The remainder of this paper is structured as follows. Section II discusses the background and related work. The design and specification of HealthFed are described in Section III. Section IV highlights the implementation and the performance evaluation of HealthFed. Finally, we conclude the paper in section V.

II. Background And Related Work

Recently, ML/DL techniques have revolutionized the healthcare industry segment; since then, ML/DL techniques are largely used to increase/boost efficiency of traditional Healthcare systems. Our previous works [28], [29] have proven to be effective at protecting organizations from network intrusions. In this paper, we extend these works to cover healthcare applications. Thus, developing privacy-aware healthcare applications supported by NGN technologies, including B6G networks. In what follow, we outline the main ML/DL-based healthcare contributions. Zheng *et al.* [30] designed a hybrid model that applies Support Vector Machine (SVM) and K-means (K-SVM), to detect breast cancer using the extracted tumor features. K-SVM has two stages. First, it uses K-means scheme to extract the hidden features of the Benign and Malignant data samples. Then, it determines the most informative features in order to use them in the elaboration of the SVM model. The authors validated their approach using Wisconsin Diagnostic Breast Cancer (WDBC) dataset. Pritom *et al.* [31] combined multiple models (*i.e.*, Naive Bayes (NB), C4.5 algorithm, Decision Tree (DT), and SVM) to detect breast cancer as well as recurrent breast cancer. The proposed solution used a feature selection method to improve the accuracy and the detection rate for each of the model. The WDBC dataset is used to demonstrate the effectiveness of the proposed solution. Hamsagayathri *et al.* [32] combined multiple models (*i.e.*, RepTree Classifier, J48, Random Forest (RF), and Random Tree (RT)) to detect breast cancer. The proposed solution used dimensionality reduction techniques as well as SEER dataset, to extract the most relevant features. Sun *et al.* [33] designed a Multi-modal Deep Neural Network (MDNNMD), which uses a Multi-dimensional data to predict breast cancer diseases. MDNNMD is composed of (1) A pre-processing bloc which processes three sub-data (*i.e.*, CNA, gene expression, and clinical data) of the multidimensional data of breast cancer; (2) A feature selection module to dynamically extract informative features to reduce the training time complexity; and (3) A Deep Neural Network (DNN) model to effectively predict breast cancer diseases. The authors evaluated the effectiveness of MDNNMD using breast cancer dataset. Sangaiah *et al.* [34] proposed a novel hybrid prediction model (RF-EGA) that combines a ReliefF ranking model with an entropy based genetic scheme to detect breast cancer; with the combination of these schemes, RF-EGA can handle a high dimension datasets. RF-EGA scheme was evaluated using Wisconsin breast cancer dataset. Kumar *et al.* [35] proposed a new voting scheme that combines multiple ML models (*i.e.*, J48, SVM, and Naïve Bayes (NB)), to effectively detect breast cancer. First, the authors used a data selection scheme to select/determine the most important tumor features. The authors validated the effectiveness of their scheme using a 10-fold cross-validation scheme on Wisconsin breast cancer dataset. Lakshmi *et al.* [36] combined two supervised ML techniques, namely, SVM and Artificial Neural Network (ANN) to automate the process of detecting breast cancer diseases. Nourelahi *et al.* [37] designed a new scheme, which uses Logistic Regression (LR) algorithm to predict breast cancer survivability. The proposed scheme can predict a 60-month survivability in patients who have been diagnosed with cancer. The authors used a realistic dataset of breast cancer patients from Shiraz University of Medical Sciences. Dutta *et al.* [38] proposed a new scheme based on a fuzzy logic/inference system, to predict breast cancer. The authors have assessed the effectiveness of their proposed scheme [38] in terms of precision, detection rate, recall, and F1 score using a clinical dataset. Huang *et al.* [39] designed a new approach, based on SVM ensembles to predict breast cancer. The authors have assessed the performance of their solution, by applying different SVM kernel functions (*e.g.*, linear kernel and RBF kernel). The authors evaluated the effectiveness of their proposed scheme [39] in terms of ROC curve, detection accuracy and F1 score.

Based on our analysis above works [30]–[39], we observe that most of them are computationally expensive. In addition, the lack of updated and realistic medical datasets, providing large data samples for an efficient training is still an ongoing challenge. The privacy nature of these medical datasets makes it difficult for clinicians and healthcare service providers to share their private. To alleviate this issue, we propose a fully decentralized framework, enabling several SDN controllers to build a robust healthcare DL-based model in a secure and collaborative way, while ensuring the privacy of such sensitive data. In addition, our framework ensures a secure aggregation of local models. Indeed, the use of the new emerging technologies, namely FL, SDN, and blockchain, will allow major healthcare institutions to collaborate with the research communities, without without disclosing the privacy of their patients' data. Thus, developing robust healthcare application supported by NGN technologies, including B6G networks.

III. HealthFed

In this section, we describe our framework (HealthFed) leveraging FL, SDN, and blockchain, to build robust healthcare-based model, while preserving the privacy of sensitive data. First, we describe the architecture of our proposed framework. Second, we briefly describe the HealthFed smart contract. Finally, we provide more details about our proposed framework.

A. HealthFed Architecture

HealthFed enables a collaborative learning among many healthcare domains ($i.e.$, SDN-based domain), by exchanging only encrypted ML model updates. Fig. 1 shows the system architecture of HealthFed. HealthFed comprises three main planes (1) A data plane which contains several OpenFlow (OF) equipment, as forwarders. SDN controllers generates OF rules that will be used by OF equipment through southbound API (Application Programming Interfaces), $e.g.$, OF protocol [40], to ensure the packets forwarding/monitoring of multiple IoT-enabled healthcare devices. Each IoT-enabled healthcare device can be used to monitor patients health or to track location of medical equipment; (2) A control plane composed of many SDN controllers; where each one will manage a particular geographical data plane, and deploys the rules issued by a application plane; and (3) An application plane which ensures a privacy-aware collaborative learning, between healthcare SDN-based domains (HDs) (see Fig. 1). First, a healthcare institution leader, referred to as organization, builds smart contracts on top of the Ethereum blockchain. It then uses functions of collaboration smart contract to add clinician collaborators. Thus, only authorized/authenticated clinician collaborators can participate in the collaboration process. Blockchain will ensure the availability, reliability, and transparency, of collaboration between healthcare SDN-based domains. Once the authentication phase is done, the healthcare institution leader will initiate the training of a shared ML model among healthcare SDN-based domains. Each healthcare SDN-based domain builds its local ML model using its own medical dataset. It then transfers the encrypted weights of its model to a central node, in order to perform a secure aggregation. The latter aggregates the encrypted updates sent from each healthcare SDN-based domain and send the aggregated value to the healthcare institution leader. Finally, the healthcare institution leader decrypts the global model, and sends its parameters to the healthcare SDN-based domains for another round of training. This process is re-executed for maximum number of iterations.

B. HealthFed's Smart Contract

We consider a healthcare institution leader who wants to manage a privacy-aware collaborative learning in healthcare systems. First, the healthcare institution leader creates and deploys the HealthFed smart contract in Ethereum blockchain. It then uses the smart contract functions of HealthFed to add clinician collaborators to the federated learning collaborative process. It includes the address of clinician collaborator, an initial credibility that can be changes over time. The credibility score gives a strongly incentive to clinician collaborators to behave correctly. The healthcare institution leader uses the HealthFed smart contract to add or remove clinician collaborators from the federated learning collaborative process. HealthFed smart contract provides the healthcare institution leader with the flexibility to manage this process in a distributed, transparent, and trustworthy way. We use solidity language to implement the HealthFed's smart contract [41]. It has several functions, including: (1) addHealthCollaborator: it can be called only by the healthcare institution leader to consider a new healthcare participant; it uses the healthcare Externally Owned Account (EOA) address as well as some public information about the healthcare organizations to add the healthcare participant to the Ethereum blockchain and timestamp when the new healthcare participant was added; and (2) removeHealthCollaborator: it can be called only by the healthcare organization leader to remove a healthcare institution from the federated system; it considers EOA of the healthcare, and removes it from the federated collaboration system.

C. HealthFed Framework

Fig. 2 shows the main interactions among HealthFed's actors. The operation of HealthFed includes the following steps. First, the healthcare institution leader initiates the process of collaboration. It initializes the parameters of the shared model with an initial weights w_0 and selects N authenticated clinician collaborators. Then, the healthcare institution leader transfers the shared model to participating clinicians n, $1 <= n <= N$. Each clinician performs local updates using its local medical training data. Afterwards, each clinician encrypts its ML model and divides it to Z_r sub-models. When receiving local models, an aggregator node z, $1 <= z <= Z_r$ aggregates the received local models and generates a new global model, before sending it to the organization. Finally, the healthcare institution leader decrypts and sends the global

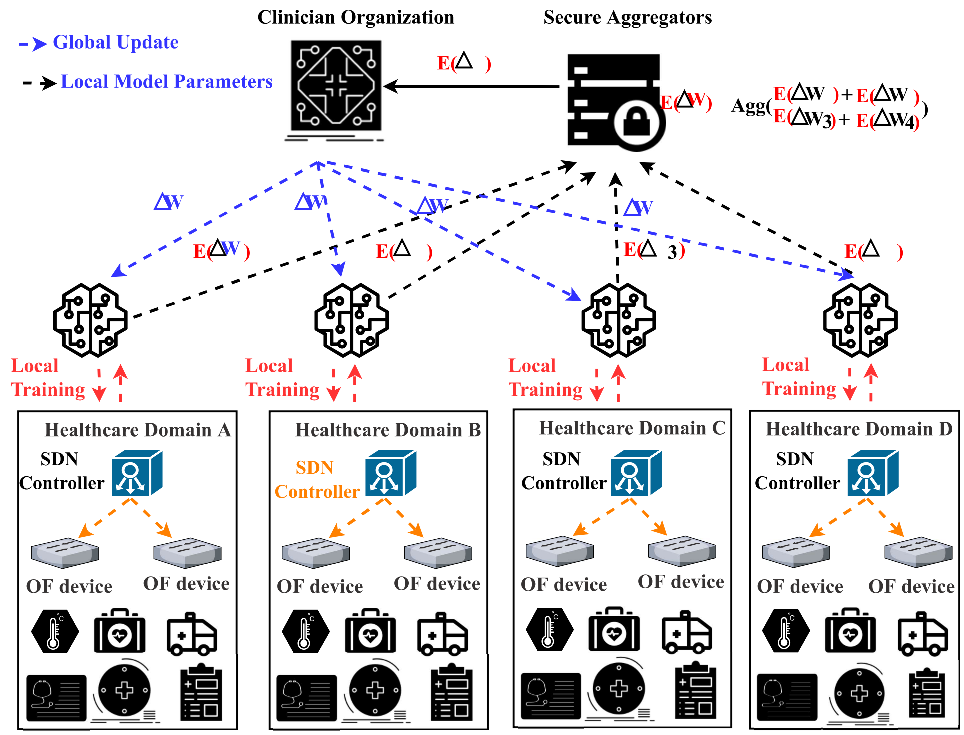


Fig. 1. HealthFed's System Architecture

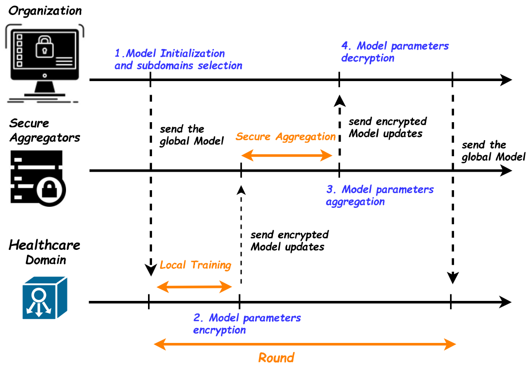

Organization
1 .Model Initialization and subdomains selection
4. Model parameters decryption
Secure Aggregators
send the global Model
Secure Aggregation
send encrypted Model updates
send the global Model
3. Model parameters aggregation
Healthcare Domain
send encrypted Model updates
Local Training
2. Model parameters encryption
Round

Fig. 2. Flowchart of HealthFed

model parameters to each healthcare SDN-based domain for another round of training. This process is re-executed during a maximum number of iterations. HealthFed is based on SDN during the collaborative learning to allow each healthcare domain to easily extract health information.

We develop a novel secure scheme, using SMPC, to secure the local models against reverse-engineering attacks. In our study, the objective function of our neural network is defined as follows:

$$\min_{w_r \in R^d} \varphi(w_r) \quad \text{where} \quad \varphi(w_r) = \frac{1}{N}\sum_{n=1}^{N} \varphi_n(w_r) \tag{1}$$

with N is the clinician participants and $\varphi_n(w_r)$ is the n^{th} clinician's objective function; φ_n. Hence, during the training phase, each clinician aims to determine the best value of parameter w_r, at each iteration r, reducing the loss function, as follows:

$$\forall n, \qquad \varphi_n(w_r) = \frac{1}{J_n}\sum_{j_n=1}^{J_n} \varphi_{j_n}(w_r; x_{j_n}, y_{j_n}) \tag{2}$$

where J_n is the number of medical data observations (x_{j_n}, y_{j_n}) of the n^{th} clinician participant.

Periodically, each clinician calculates the average gradient using its own medical data, as follows:

$$g_r^n = \nabla\varphi_n(W_r; b_r) \tag{3}$$

Then, each clinician participant performs a local gradient descent step on the shared ML Model, utilizing their local medical records as follows:

$$\forall n, \qquad W_r^n \leftarrow W_r - \eta\nabla\varphi_n(W_r; b) \tag{4}$$

To protect the locally computed models from reverse engineering attacks, we apply SMPC, which enables each clinician n to divide its learning models into multiple shares. Therefore, aggregator nodes $z = 1, \ldots, Z_r$ will receive encrypted shares $w_{r,z}^{c,n}$, and learn nothing about the secret w_r^n. SMPC provides better computation performance, as compared to other cryptography approaches, such as Homomorphic.

First, we consider a finite fixed range from $0, 1, \ldots, P$ for a prime P. Each clinician n, encodes its learning parameters w_r^n as integers. Then, it calculates the modulo of the result with the prime P. After that, each clinician n divides its encrypted values to Z_r shares. The aggregator nodes z, $1 \leq z \leq Z_r$, will receive such shares $w_{r,z}^{n,c}$ (*i.e.*, locally encrypted model

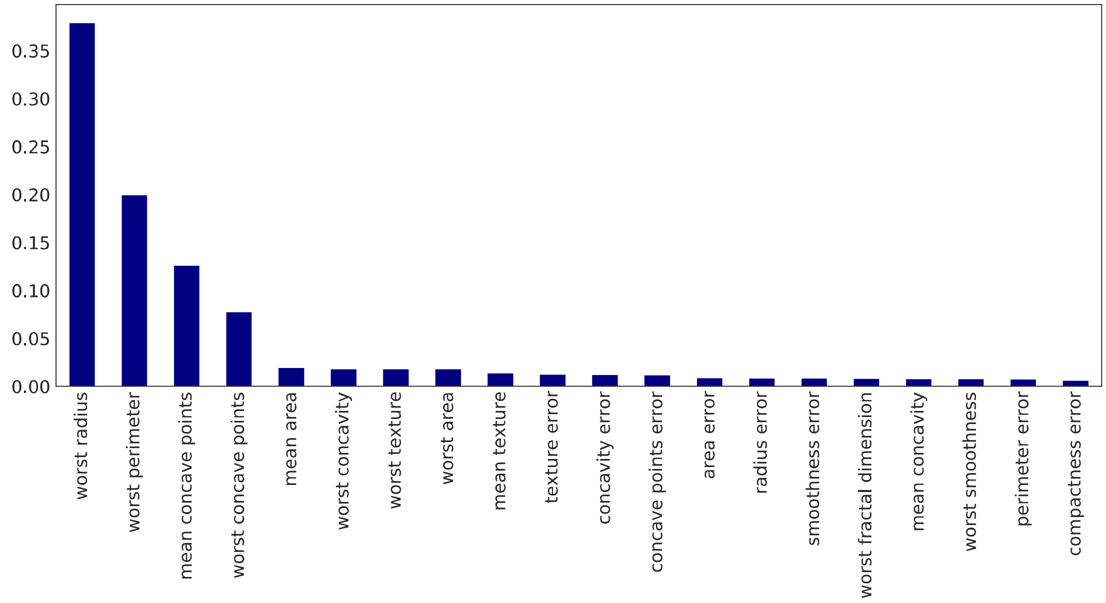


Fig. 3. TIGFS Score of features for Breast Cancer Wisconsin dataset

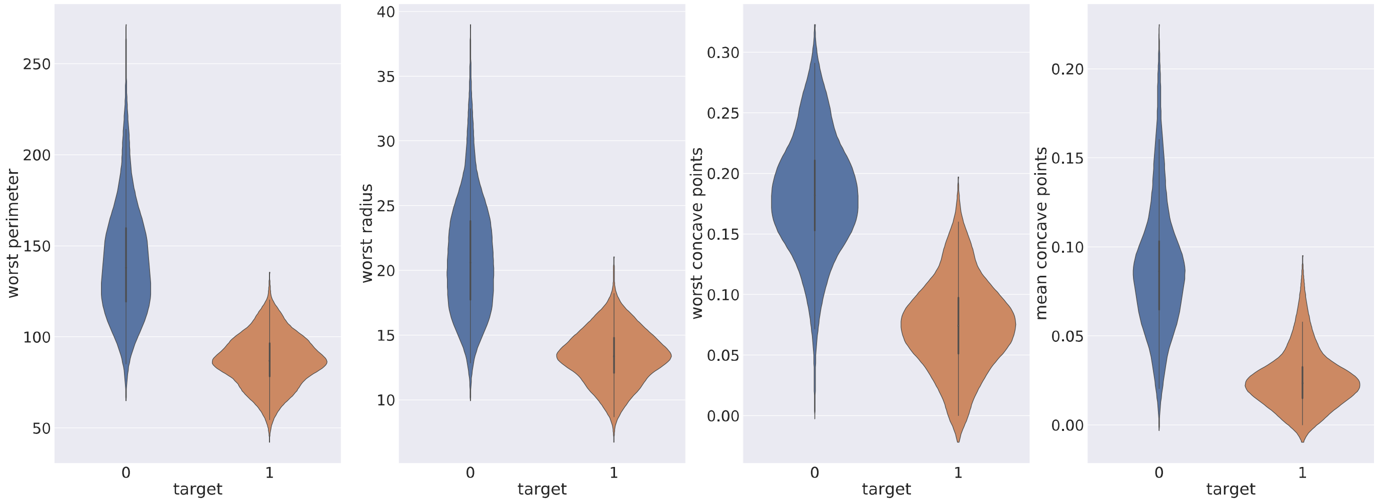


Fig. 4. Data distribution of the highest scoring features of Breast Cancer Wisconsin dataset

updates). Then, each secure aggregator z aggregates these encrypted values and updates the weights as follows:

$$\forall z, \qquad W^c_{r+1,z} \leftarrow W^c_{r,z} - \eta \frac{1}{N} \sum_{n=1}^{N} g^n_{r,z} \tag{5}$$

where η is a fixed learning rate on each clinician collaborator, and $\frac{1}{N}\sum_{n=1}^{N} g^n_{r,z}$=$\nabla\varphi(W^c_{r,z}; b^c_{r,z})$. For each collaborator n:

$$\forall n, \qquad W^{c,n}_{r+1} \leftarrow W^c_n - \eta g^n_r \tag{6}$$

Thus, each aggregator z generates a new model update, as follows:

$$\forall z, \qquad W^c_{r+1,z} \leftarrow \frac{1}{N} \sum_{n=1}^{N} W^{c,n}_{r+1,z} \tag{7}$$

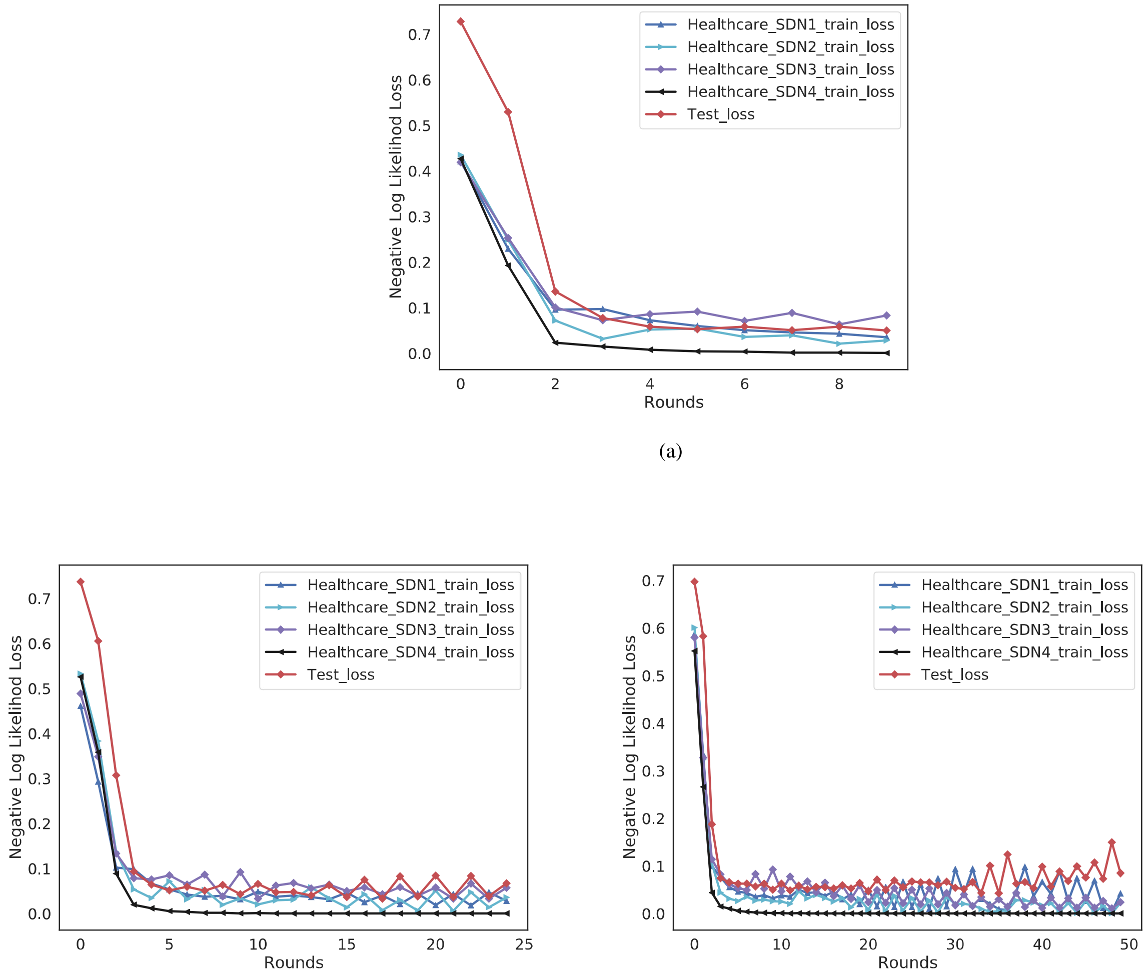
Finally, the healthcare institution leader sums the received model updates, from each aggregator z, in order to reconstruct the secret. It then decrypts the learning models and sends the aggregated model to the involved clinicians. This process is run during a maximum number of iterations r_{max}.

IV. Experimental Study

This section validates the performance of HealthFed through an experimental study. First, we present the simulation parameters. Second, we discuss experimental results. Last, we assess the effectiveness of HealthFed.

A. Simulation Parameters

To implement HealthFed, we used Pysyft library which is privacy-aware DL tool, built on top of PyTorch [42]. In



(a)

(b)

(c)

Fig. 5. Model loss for (a) 10 rounds; (b) 25 rounds; and (c) 50 rounds.

addition, we used Mininet emulator [43], to create a real SDN-based network, which enables to deploy virtual OpenFlow switches (*e.g.*, OpenVswitch [44]) and containers, including multiple SDN controllers (*i.e.*, four SDN controllers, see Fig.1), to perform local training on their local data. We also used truffle framework to deploy smart contract of our HealthFed [45]. We first used Ganache simulator to deploy HealthFed smart contract on a private blockchain [46]. We then deployed it on Ethereum official Ropsten [47]. For the test data, we leverage the well-known public Breast Cancer Wisconsin dataset; it contains real-valued features, extracted from digitized images of a fine needle aspirate (FNA) of a breast mass; it contains two classes: (1) Benign (B): it represents data samples of patients who were cancer-free; and (2) Malignant (M): it represents data samples of patients who have been diagnosed with cancer.

B. Experimental Results

We evaluate the performance of HealthFed using Breast Cancer Wisconsin dataset; the objective is to build a global model that is capable to accurately classify each data sample, as either Benign or Malignant data sample. The data distributions of features of Breast Cancer Wisconsin dataset vary widely. Thus, we applied Equation (8) to re-scale these data values.

$$X'_i = \frac{X_i - Mean(X_i)}{stdev(X_i)} \tag{8}$$

With X_i is an input feature, for example, radius and fractal dimension, $stdev(X_i)$ and $Mean(X_i)$ are the standard and mean deviation values of each input feature, respectively.

We constructed a shared DL Model comprising: one input layer of 30 dimensions, that considers the 30 input features of the Breast Cancer Wisconsin data, 2 hidden layers and an output layer of 2 dimensions. In addition, to prevent overfitting, we used Dropout technique [48]. We used cross entropy $\mathcal{L}$, as a loss function (Equation. (9)). We used Adam (Adaptive Moment Estimation) [49] as optimizer, to reduce the value of $\mathcal{L}$.

$$\mathcal{L} = -\frac{1}{M}\sum_{m=1}^{M} y_m * log(\hat{y_m}) \tag{9}$$

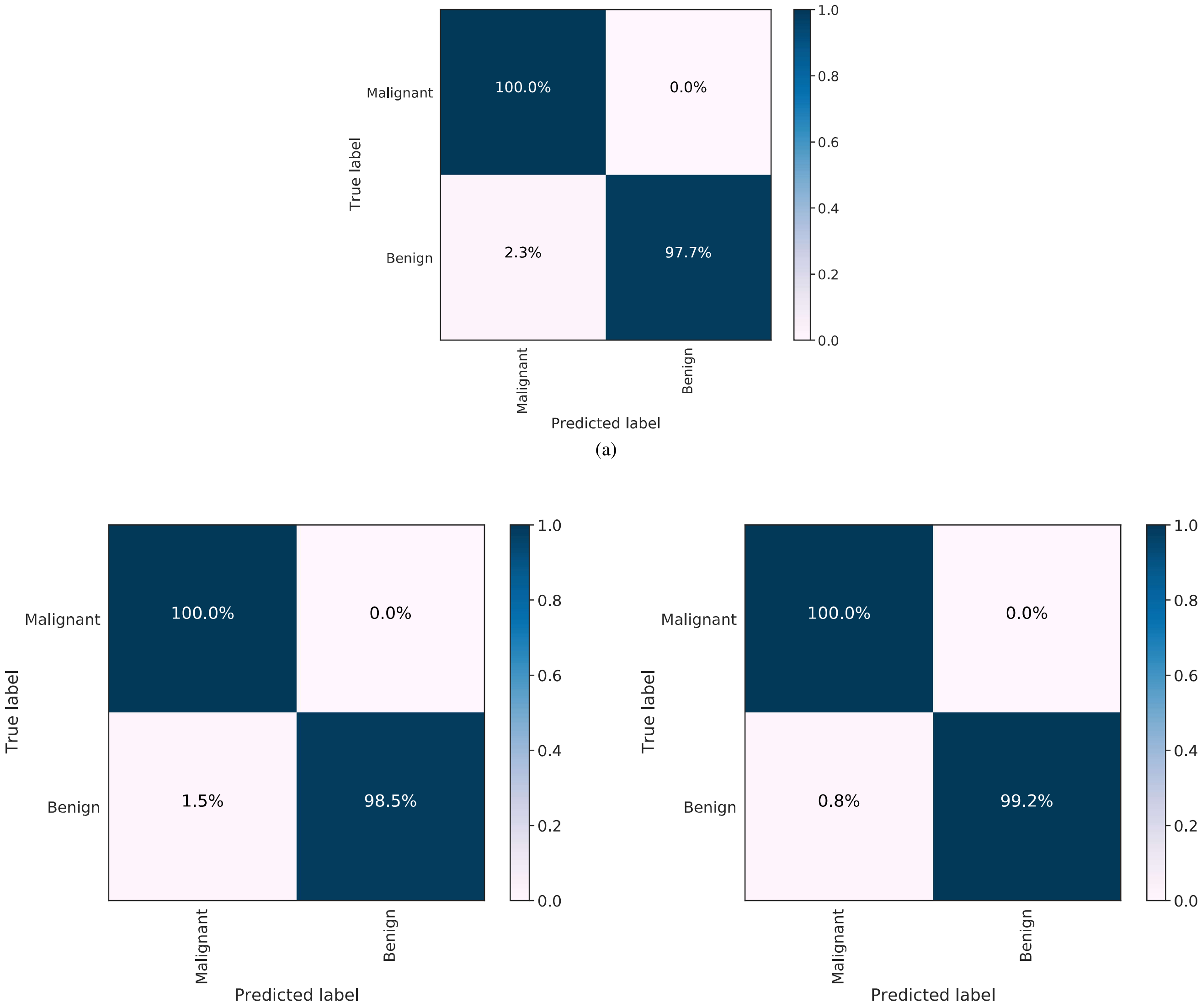


Fig. 6. Confusion matrices for (a) 10 rounds; (b) 25 rounds; and (c) 50 rounds.

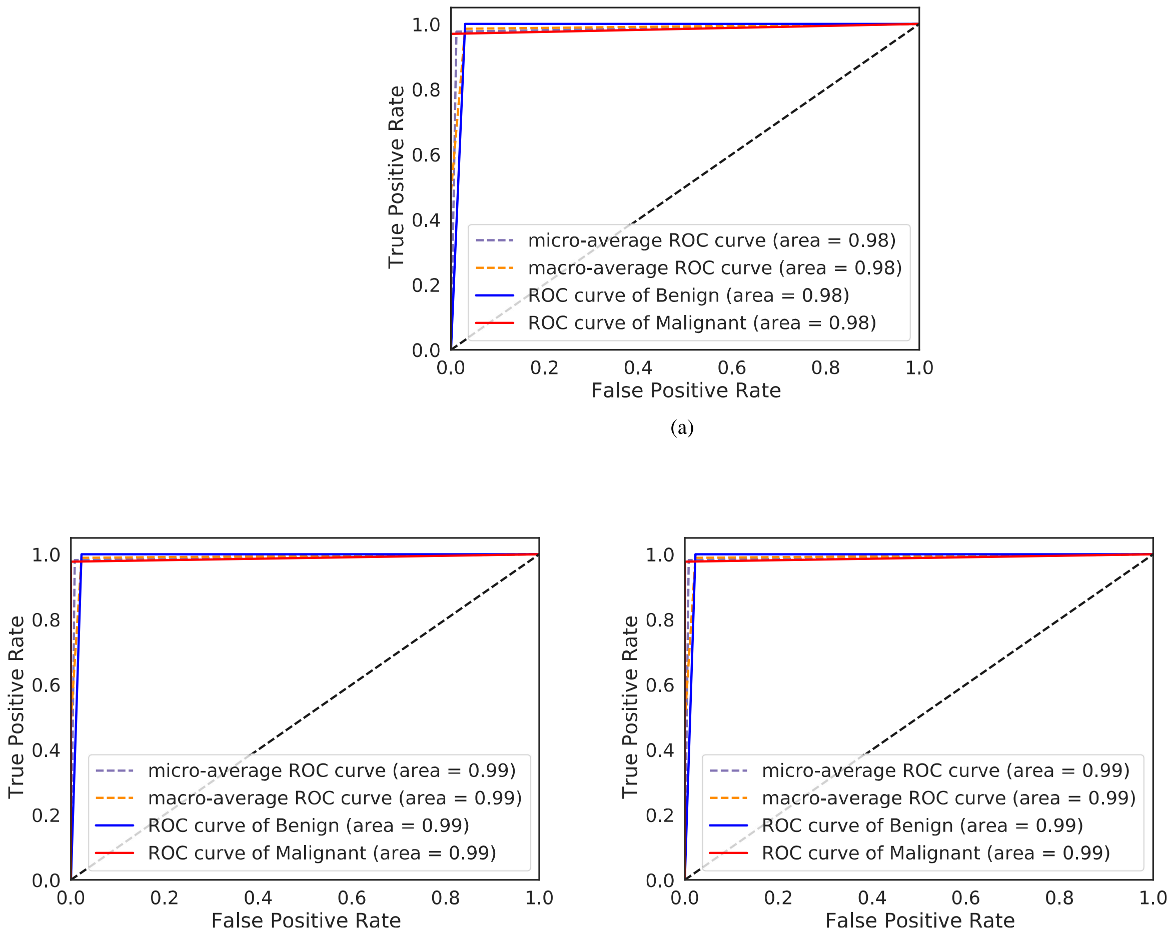
with $\hat{y_m}$ is the predicted values of the m^{th} class, while y_m is the real values of m^{th} class, and M represents the number of data observations.

Besides, we used Tree-based Information gain Feature Selection (TIGFS), to determine the most important features in the Breast Cancer dataset. Fig. 3 shows TIGFS score of the used dataset features; it depicts such scoring features in descending order. We clearly notice that more than 85% of the 30 features of Breast Cancer Wisconsin dataset do not have an important contribution. Fig. 4 shows the data distribution of the highest scoring features ($i.e.$, 'worst perimeter', 'worst radius', 'worst concave points', and 'mean concave points') of Breast Cancer Wisconsin dataset. We notice that the most informative features can effectively separate the two classes ($i.e.$, Benign (B) and Malignant (M)). Hence, the TIGFS method can highly help to deduce relevant features, that can improve directly the performance of the trained model, and hence generate an accurate ML/DL prediction model.

To test HealthFed, we divide Breast Cancer Wisconsin dataset, in order to give each clinician a subset of the data. We note also that each SDN domain trains its models during 10 to 50 rounds/iterations of federated training and for a number of epochs from 1 to 5. Figs. 5 (a), 5(b), and 5(c) illustrate the learning curves of the four SDN-based healthcare domains for 10, 25, and 50 of rounds/iterations of training, respectively. We notice that the loss decreases to a minimum value ($i.e.$, almost zero), which shows clearly that models are able to learn from each other, without being able to exchange/share their private medical data.

C. Performance Validation

We validate the HealthFed performance in terms of several metrics, including precision, accuracy, True Positive Rate (TPR), False Positive Rate (FPR), F1-score, and Area Under the ROC Curve (AUC). In addition, we also consider confusion matrix in addition to ROC curves, which reflect TPR based on FPR (see Table I). Noting that FN (False Negatives) measures the rate of malignant data observations, which are predicted



(a)

(b)

(c)

Fig. 7. ROC Curves for (a) 10 rounds; (b) 25 rounds; and (c) 50 rounds.

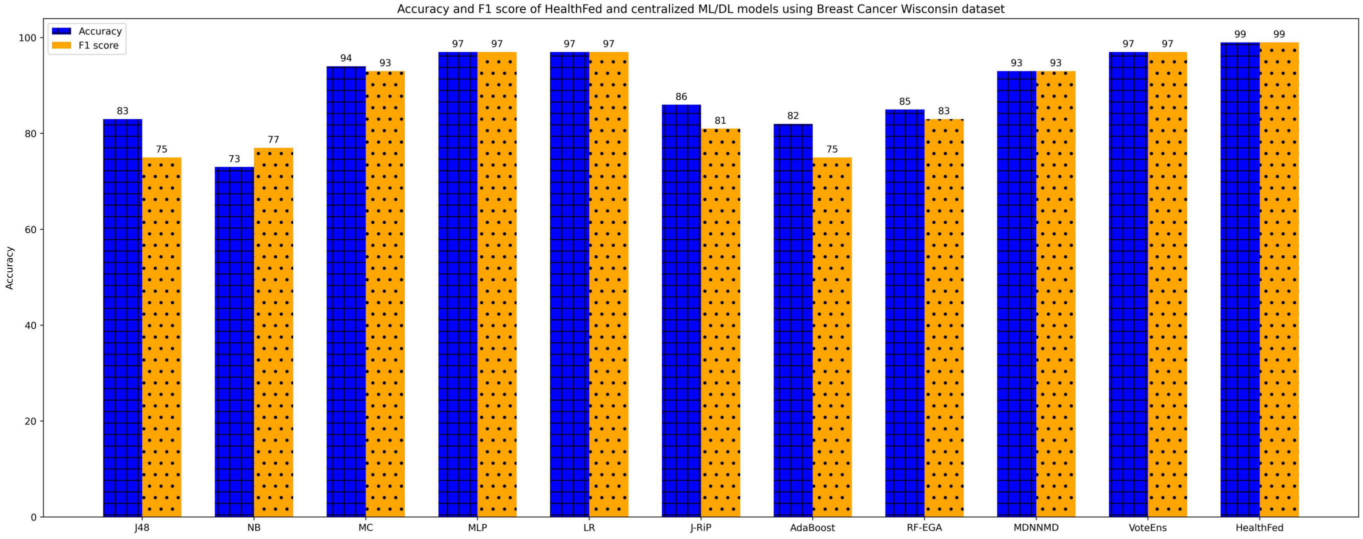


Fig. 8. Accuracy and F1 Score of HealthFed and centralized ML/DL models using Breast Cancer Wisconsin dataset

as benign data samples, TP (True Positives) measures the rate of malignant data observations, predicted as malignant data as well, FP (False Positives) measures the rate of benign data, and are predicted as malignant data, while TN (True Negatives)

TABLE I
CONFUSION MATRIX.

	Classified as Malignant	Classified as Benign
Malignant	True Positives	False Negatives
Benign	False Positives	True Negatives

reflects the rate of benign data, which are correctly predicted, as benign data.
Accuracy is the rate of corrected predictions, on the total number of predictions, calculated as follows:

$$Accuracy = \frac{TN + TP}{FP + TP + FN + FP} \tag{10}$$

Precision is the rate of corrected predictions of malignant data samples, on the total number of predictions related to malignant data samples, as follows:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

True Positive Rate (TPR), called also Detection Rate (DR) is calculated as follows:

$$TPR = \frac{TP}{TP + FN} \tag{12}$$

False positive rate (FPR) is the rate of malignant data, predicted incorrectly as benign data, on the total benign data, calculated as follows:

$$FPR = \frac{FP}{FP + TN} \tag{13}$$

F1 score considers both detection rate and precision and metrics to calculate the harmonic mean of both recall and precision, as follows:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{14}$$

TABLE II
PERFORMANCE METRICS OF HEALTHFED

Iterations	Precision	Accuracy	Recall	F1	Time(s)
10	98%	98%	98%	98%	12.04
25	99%	99%	99%	99%	32.31
50	99%	99%	99%	99%	60.48

Figs. 6(a), 6(b), and 6(c) show the confusion matrices for 10 iterations, 25 iterations, and 50 iterations of training, respectively. HealthFed reaches 98% in precision, accuracy, F1 score, and recall, respectively, for 10 iterations of learning (with only 12.04 seconds of training). During 25 iterations/rounds of training, HealthFed reaches 99% of precision, accuracy, F1 score, and recall, respectively (with only 32.31 seconds of training). For 50 iterations/rounds of training, HealthFed reaches 99% in precision, accuracy, F1 score, and recall, respectively (with only 60.48 seconds of training). Table II lists the overall performance of HealthFed. The AUC metric shows the degree of separability between the benign and malignant classes. Figs. 7(a), 7(b), and 7(c) show the ROC curves of the inference FL model on Breast Cancer Wisconsin dataset. We obtain a high AUC score of 0.99 for 10, 25, and 50 iterations of training, respectively. The experiment results demonstrate that HealthFed reaches promising performance, while ensuring the privacy of the involved clinician healthcare.

D. Comparative Analysis

In this section, we compare the performance of HealthFed with centralized existing approaches, leveraging Breast Cancer dataset. Kumar *et al.* [50] implement seven different ML/DL models: AdaBoost, J-Rip, Naive Bayes (NB), J48, Multiclass Classifier (MC), Logistics Regression (LR), and Multi-layer Perceptron (MLP). In addition, we also compare HealthFed with: Multi-modal Deep Neural Network [33], RF-EGA [34], a voting ensemble learning model (VoteEns), that combines common models (*i.e.*, J48, SVM, and Naïve Bayes (NB)) [35]. Table III gives the performance of existing approaches as well as our HealthFed scheme, when using Breast Cancer Wisconsin dataset. Fig. 8 shows both the accuracy and F1 score metrics of HealthFed and centralized existing ML/DL models. We observe that HealthFed needs only 32.31 seconds of training time, to generate better accuracy and F1 score (99%). The obtained results confirm that HealthFed not only has better F1 score and detection accuracy than centralized ML/DL contributions, but also ensures the privacy of the collaborators. Thus, HealthFed is a promising privacy-aware framework for healthcare systems

TABLE III
PERFORMANCE EVALUATION COMPARISON

Techniques	Accuracy	F1	Time (second)
J48	0.83	0.75	N/A
NB	0.73	0.77	N/A
MC	0.94	0.93	N/A
MLP	0.97	0.97	N/A
LR	0.97	0.97	N/A
J-Rip	0.86	0.81	N/A
AdaBoost	0.82	0.75	N/A
MDNNMD	0.93	N/A	N/A
RF-EGA	0.85	0.83	N/A
VoteEns	0.97	N/A	N/A
HealthFed	**0.99**	**0.99**	**32.31**

V. CONCLUSION

In this paper, we proposed a novel framework, called HealthFed, that enables multiple clinical participants (SDN domains) to collaboratively train an effective DL-based healthcare model, while ensuring their privacy. First, we introduced a novel secure aggregation approach that uses SMPC to securely aggregate local updates. Then, we built a blockchain-based approach that uses Ethereum blockchain to maintain collaboration between clinicians in a fully decentralized, reliable, and flexible way. The obtained results on the public Breast Cancer dataset showed that HealthFed achieved privacy and high accuracy. HealthFed provides better performance as compared to centralized ML/DL-based solutions, in terms of F1 score and accuracy, while ensuring the privacy of each collaborator's sensitive data. This makes HealthFed a promising framework for healthcare systems. As a future work, we plan to consider other healthcare datasets in order to cover other types of critical diseases.

References

[1] N. cancer institute, "Cancer statistics." Accessed: June. 1, 2020. [Online]. Available: https://www.cancer.gov/about-cancer/understanding/statistics

[2] Z. A. E. Houda, A. Hafid, and L. Khoukhi, "Brainchain - a machine learning approach for protecting blockchain applications using sdn," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[3] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, "Deep learning for health informatics," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4–21, 2017.

[4] F. Yang, M. Poostchi, H. Yu, Z. Zhou, K. Silamut, J. Yu, R. J. Maude, S. Jaeger, and S. Antani, "Deep learning for smartphone-based malaria parasite detection in thick blood smears," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1427–1438, 2020.

[5] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2733–2742, 2020.

[6] M. H. Sarhan, M. A. Nasseri, D. Zapp, M. Maier, C. P. Lohmann, N. Navab, and A. Eslami, "Machine learning techniques for ophthalmic data processing: A review," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 12, pp. 3338–3350, 2020.

[7] A. S. Panayides, A. Amini, N. D. Filipovic, A. Sharma, S. A. Tsaftaris, A. Young, D. Foran, N. Do, S. Golemati, T. Kurc, K. Huang, K. S. Nikita, B. P. Veasey, M. Zervakis, J. H. Saltz, and C. S. Pattichis, "Ai in medical imaging informatics: Current challenges and future directions," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1837–1857, 2020.

[8] W. Y. B. Lim, S. Garg, Z. Xiong, D. Niyato, C. Leung, C. Miao, and M. Guizani, "Dynamic contract design for federated learning in smart healthcare applications," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16 853–16 862, 2021.

[9] B. Brik, M. Messaadia, M. Sahnoun, B. Bettayeb, and M. A. Benatia, "Fog-supported low-latency monitoring of system disruptions in industry 4.0: A federated learning approach," *ACM Trans. Cyber-Phys. Syst.*, vol. 6, no. 2, may 2022.

[10] B. Brik, A. Ksentini, and M. Bouaziz, "Federated learning for uavs-enabled wireless networks: Use cases, challenges, and open problems," *IEEE Access*, vol. 8, pp. 53 841–53 849, 2020.

[11] B. Brik and A. Ksentini, "On predicting service-oriented network slices performances in 5g: A federated learning approach," in *LCN 2020, 45th IEEE Conference on Local Computer Networks, 16-19 November 2020, Sydney, Australia (Virtual Conference)*, IEEE, Ed., Sydney, 2020.

[12] Z. Chen, P. Tian, W. Liao, and W. Yu, "Zero knowledge clustering based adversarial mitigation in heterogeneous federated learning," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1070–1083, 2021.

[13] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, "Distributed federated learning for ultra-reliable low-latency vehicular communications," *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 1146–1159, 2020.

[14] H. Moudoud, Z. Mlika, L. Khoukhi, and S. Cherkaoui, "Detection and prediction of fdi attacks in iot systems via hidden markov model," *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2022.

[15] Y. Li, Y. Zhou, A. Jolfaei, D. Yu, G. Xu, and X. Zheng, "Privacy-preserving federated learning framework based on chained secure multiparty computing," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6178–6186, 2021.

[16] Z. A. El Houda, L. Khoukhi, and A. Hafid, "Chainsecure - a scalable and proactive solution for protecting blockchain applications using sdn," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[17] S. Scott-Hayward, S. Natarajan, and S. Sezer, "A survey of security in software defined networks," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 623–654, 2016.

[18] Z. A. El Houda, A. S. Hafid, and L. Khoukhi, "A novel machine learning framework for advanced attack detection using sdn," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[19] D. B. Rawat and S. R. Reddy, "Software defined networking architecture, security and energy efficiency: A survey," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 325–346, 2017.

[20] Z. Abou El Houda, L. Khoukhi, and A. Senhaji Hafid, "Bringing intelligence to software defined networks: Mitigating ddos attacks," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2523–2535, 2020.

[21] Z. Abou El Houda, A. S. Hafid, and L. Khoukhi, "Cochain-sc: An intra- and inter-domain ddos mitigation scheme based on blockchain using sdn and smart contract," *IEEE Access*, vol. 7, pp. 98 893–98 907, 2019.

[22] Z. Abou El Houda, "Renforcement de la sécurité à travers les réseaux programmables," Ph.D. dissertation, Université de Montréal, 2021.

[23] H. Moudoud, S. Cherkaoui, and L. Khoukhi, "An iot blockchain architecture using oracles and smart contracts: the use-case of a food supply chain," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1–6.

[24] Z. Abou El Houda, "Security enforcement through software defined networks (sdn)," Ph.D. dissertation, Troyes, 2021.

[25] T. Ryffel, A. Trask, M. Dahl, B. Wagner, J. Mancuso, D. Rueckert, and J. Passerat-Palmbach, "A generic framework for privacy preserving deep learning," *CoRR*, vol. abs/1811.04017, 2018. Accessed: June. 1, 2020. [Online]. Available: http://arxiv.org/abs/1811.04017

[26] Wood G., "Ethereum: A secure decentralised generalised transaction ledge." Accessed: June. 1, 2020. [Online]. Available: https://ethereum.org/

[27] W. William, S. Nick, and M. Olvi, "Breast cancer wisconsin (diagnostic) data set," 2021. Accessed: June. 1, 2020. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

[28] Z. Abou El Houda, A. S. Hafid, and L. Khoukhi, "Cofed: A privacy preserving collaborative ddos mitigation framework based on federated learning using sdn and blockchain," *IEEE Transactions on Network Science and Engineering*, 2021.

[29] Z. Abou El Houda, B. Brik, A. Ksentini, L. Khoukhi, and M. Guizani, "When federated learning meets game theory: A cooperative framework to secure iiot applications on edge computing," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2022.

[30] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of k-means and support vector machine algorithms," *Expert Systems with Applications*, vol. 41, no. 4, Part 1, pp. 1476–1482, 2014. Accessed: June. 1, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417413006659

[31] A. I. Pritom, R. Munshi, Ahadur, S. A. Sabab, and S. Shihab, "Predicting breast cancer recurrence using effective classification and feature selection technique," pp. 310–314, 2016.

[32] P. Hamsagayathri and P. Sampath, "Decision tree classifiers for classification of breast cancer," *International Journal of Current Pharmaceutical Research*, vol. 9, pp. 31–36, 2017.

[33] D. Sun, M. Wang, and A. Li, "A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 3, pp. 841–850, 2019.

[34] I. Sangaiah and A. V. A. Kumar, "Improving medical diagnosis performance using hybrid feature selection via relieff and entropy based genetic search (rf-ega) approach: application to breast cancer prediction," *Cluster Computing*, pp. 1–8, 2018.

[35] U. K. Kumar, M. S. Nikhil, and K. Sumangali, "Prediction of breast cancer using voting classifier technique," pp. 108–114, 2017.

[36] I. Lakshmi and G.Krishnaveni, "Performance assessment by using svm and ann for breast cancer mammography image classification," *International Journal of Engineering Technology Science and Research*, vol. 4, p. 620–626, 2017.

[37] M. Nourelahi, A. Zamani, A. Talei, and S. Tahmasebi, "A model to predict breast cancer survivability using logistic regression," *Middle East Journal of Cancer*, vol. 10, no. 2, pp. 132–138, 2019.

[38] D. Soumi, G. Sujata, S. Abhijit, P. Rechik, P. Rohit, and R. Rohit, "Cancer prediction based on fuzzy inference system." *Advances in Intelligent Systems and Computing*, vol. 851, 2019.

[39] H. Min-Wei, C. Chih-Wen, L. Wei-Chao, K. Shih-Wen, and T. Chih-Fong, "Svm and svm ensembles in breast cancer prediction," vol. 10, no. 2, 2017.

[40] "Openflow switch specification." Accessed: June. 1, 2020. [Online]. Available: https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-spec-v1.3.0.pdf

[41] "Solidity." Accessed: June. 1, 2020. [Online]. Available: https://solidity.readthedocs.io/en/develop/

[42] "Pytorch framework." Accessed: June. 1, 2020. [Online]. Available: https://pytorch.org/

[43] "Mininet." Accessed: June. 1, 2020. [Online]. Available: http://mininet.org

[44] "Openvswicth." Accessed: June. 1, 2020. [Online]. Available: https://www.openvswitch.org/

[45] "Truffle." Accessed: June. 1, 2020. [Online]. Available: https://truffleframework.com/

[46] "Ganache." Accessed: June. 1, 2020. [Online]. Available: https://truffleframework.com/docs/ganache/overview

[47] "Ropsten." Accessed: June. 1, 2020. [Online]. Available: https://ropsten.etherscan.io/

[48] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.

[49] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2015.

[50] V. Kumar, B. K. Mishra, M. Mazzara, D. N. H. Thanh, and A. Verma, "Prediction of malignant & benign breast cancer: A data mining approach in healthcare applications," *CoRR*, vol. abs/1902.03825, 2019. Accessed: June. 1, 2020. [Online]. Available: http://arxiv.org/abs/1902.03825