Highly expressive heart disease prediction based on Federated Learning

Tongxuan Bie[1] and Longjie Cao[1]

[1]Department of Science, China University of Petroleum (East China)

1 Abstract

Over the past 20 years, heart disease has remained the leading cause of death worldwide. The growing connection between deep learning and the medical field has led us to imagine more medical intelligence. However, this requires a certain amount of data to support the training of models, and medical institutions do not share data with other institutions to protect the privacy of patients. To address this problem, we propose a federated learning framework that uses deep learning to predict heart disease without disclosing patient data from individual healthcare institutions. We up-dimension features after data pre-processing to improve the expressiveness of the data and investigate the impact of uneven data volumes on the model. Experiments show that our model achieves similar accuracy to centralized training and is insensitive to data heterogeneity. In addition, the high recall rate of our framework on the test set allows patients with heart disease to be treated at an early stage, avoiding the serious consequences of worsening heart disease.

2 Introduction

With the development of deep learning, people are beginning to expect its impact on medical diagnosis. However, much of the stagnation in deep learning in healthcare is due to the lack of quality data, which is often hampered by the lack of privacy protection mechanisms[1][2]. Meanwhile, most of the current deep learning-based medical diagnostics are focused on image data, such as skin cancer classification or lung disease picture classification. But the fact is that digital data is also accumulating in large quantities, and although it cannot contain the same amount of information as image data, the utility of digital data in everyday prevention and in portability is enormous. Our article thus targets the two pain points of privacy protection and the expressiveness of digital data for medical deep learning diagnostics. The aim of our research is to use the

IEEE's Joint Heart Disease Dataset to simulate realistic situations to implement a privacy-preserving distributed highly expressive learning framework for further use in the prediction of heart disease.

Deep learning has sprouted many applications in the medical field. In deep learning for image medical diagnosis, there are good results in brain tumor segmentation[3], whole brain segmentation for MRI[4], and benign and malignant classification for Skin Lesion[5]. It is also considered promising for the detection of COVID-19[6][7][8]. But many medical deep learning-related studies do not consider privacy protection Federation learning, the hot privacy-preserving training framework, has begun to be experimented with in various areas of healthcare since it was first proposed in 2016[9]. It is used in a limited number of medical topics. For example, in the EHR (electronic health record) problem, federated learning is used to match patients with similar symptoms[10] and also to predict mortality and ICU length of stay[11].

In the diagnosis of cardiac disease, Akis Linardos et al. used MRI images for diagnosis and invoked federated learning for privacy[12]. A great deal of research has focused on image diagnosis, somewhat ignoring the value of digital data on heart disease. In digital data, traditional heart disease prediction tends to use traditional machine learning models such as Random Forest[13], Decision Trees[14] and Naive Bayes[15].

The main contributions of our articles are as follows. 1. Elevating the dimensionality of the most common 1D data for heart disease to improve expressiveness. 2. Federated learning LeNet training framework based on FedAvg for heart disease classification questions.

3 Proposed Framework

3.1 Cross Section

With this approach, the non-linear representation of the features is increased and the expressiveness of the features is greatly enhanced. In the field of cardiac diagnosis, in particular, this correlation is very prominent. Based on the holistic nature of medicine and organisms, disease does not arise from a mere pile-up of indicators like age, sex, but rather from a disorder of complex connections between systems, so that the combination of features becomes crucial.

Most data from hospital physicals and wearable devices is one-dimensional, but one-dimensional data naturally ignores associations between features. So we intend to multiply between features in an attempt to add in feature correlations

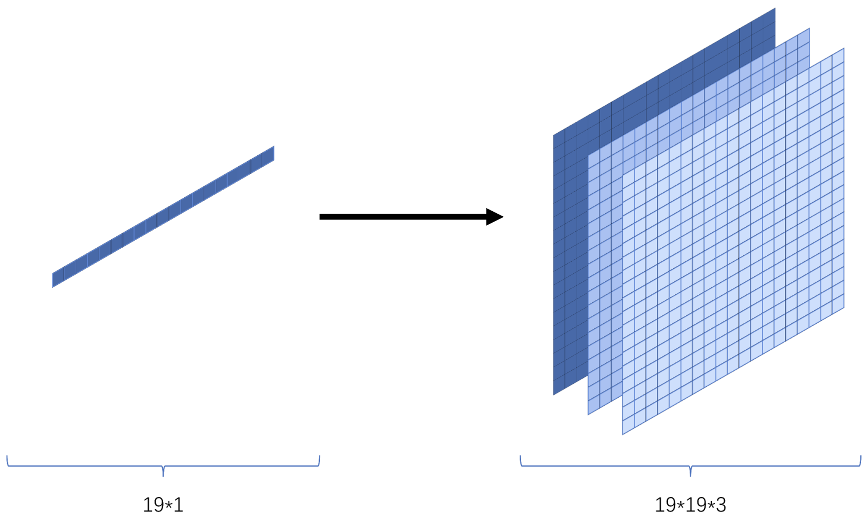

Figure 1: Cross

Algorithm 1: Transform one-dimensional features into eigenmatrices

---

Input: primitive features S,number of features n
Output: Feature combination matrix T
for $i \leftarrow 0$ **to** n **do**
    for $i \leftarrow 0$ **to** n **do**
        $T_{ij0} \leftarrow S_i$;
        $T_{ij1} \leftarrow S_j$;
        $T_{ij0} \leftarrow S_i{*}S_j$;
    end
end
return T

3.2 Federated Learning Framework

3.2.1 Loss

The loss function is often described by the following equation.

$$f(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w) \tag{1}$$

In this formulation we have K Clients, $F_k(w)$ represents the loss function of the Client k, and n_k is measured by the size of the dataset of Client k. So clients with a greater amount of data will have a stronger weighting, which means they will have a higher impact on the global loss function. We train shared models in client-sever interactions using FedAvg's algorithm, which does this in a way that controls global loss. This global loss is a weighted average of the losses of each individual customer.

Algorithm 2: FederatedAveraging.The K clients areindexed by k; B is the local minibatch size, E is the numberof local epochs, and η is the learning rate[16]

Server executes :
initialize w_0;
foreach round $t = 1, 2, \dots$ **do**
  $m \leftarrow \max(C \cdot K, 1)$;
  $S_t \leftarrow$ (random set of m clients);
  foreach *client* $k \in S_t$ **inparallel do**
    $w_{t+1}^k \leftarrow$ ClientUpdate (k, w_t)
  end
  $w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k$
end

ClientUpdate(k, w) : // Run on client k
$\mathcal{B} \leftarrow$ (split $\mathcal{P}_k$ into batches of size B)
for each local epoch i from 1 to E **do**
  for batch $b \in \mathcal{B}$ **do**
    $w \leftarrow w - \eta \nabla \ell(w; b)$
  end
end
return w to server

3.2.2 The Framework

The whole framework of federated learning can be represented by this schematic Figure 2. In traditional cardiac deep learning research, researchers tend to ignore privacy issues and put the model and data in the same device, i.e. centralized learning. But federated learning can integrate data to train models while keeping the data local. Instead of uploading data to the cloud for centralized training, we transfer models from the Center Server to each Client. Each Client can then use its own data to train these models locally and then upload the updated weights back to the Center Server. The Center Server then aggregates these incoming parameters and updates the global model, then repeats the process over multiple rounds of training. In fact, we will not train the model on every Client, but will select some at random for training.

We build a deep convolutional neural network (LeNet) to perform feature extraction and classification of two-dimensional data from heart disease. Its input is a feature matrix that has been dimensioned upwards and its output is the probability of having heart disease. Using convolution and the fit of two-dimensional data, we can extract as much information as possible from the features.

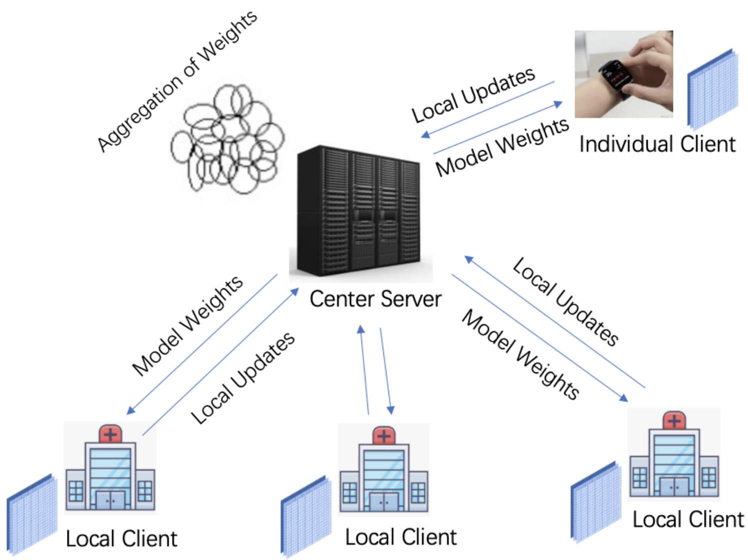

Figure 2: Federated Learning

3.2.3 Client and Server

Training is performed on the client side, with each client configuring the LeNet network for training, and updates to the weights obtained from training are recorded and uploaded to the Center Server. During this process, the data local to the Client never leaves the local area. When Center Server finishes integration, it passes the integrated model parameters back to the Client, which repeats this step and trains again.

Algorithm 3: Federated learning. Client-side training at federated round$i$[17].

Require : α : Stepsize
Require : $\beta_1, \beta_2 \in [0,1)$: Exponential decay rates for the moment estimates
Require : $f(w)$: Stochastic objective function with parameters w
Require : w_0 : Initial parameter vector
$m_0 \leftarrow 0$ (Initialize 1^{st} moment vector)
$v_0 \leftarrow 0$ (Initialize 2^{nd} moment vector)
$t \leftarrow 0$ (Initialize timestep)
while w_t not converged **do**
  $t \leftarrow t+1$
  $g_t \leftarrow \nabla_w f_t(w_{t-1})$
  $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1-\beta_1) \cdot g_t$
  $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g_t^2$
  $\widehat{m}_t \leftarrow m_t / (1-\beta_1^t)$
  $\widehat{v}_t \leftarrow v_t / (1-\beta_2^t)$
  $w_t \leftarrow w_{t-1} - \alpha \cdot \widehat{m}_t / \left(\sqrt{\widehat{v}_t} + \epsilon\right)$
end
return w_t

The main role of the Center Server is to initialize and integrate. After initializing the global models, it sends them to the selected Clients. When it receives the weights back from the Client, it aggregates them using methods such as weighted averaging.

Algorithm 4: Federated learning. Server-side aggregation procedure[17].

Require T : num_rounds
procedure AGGREGATING(C, K)
Initialize global model w^0
foreach *round* $t = 1, 2, \ldots T$ **do**
  $m \leftarrow \max(C \times K, 1)$
  $S_t \leftarrow$(random set of m clients) ▷ Selected Clients for round t
  foreach *client* $k \in S_t$ **do**
    `// Run in parallel`
    Send w^{t-1} to client k
    $w_k^t \leftarrow$CLIENTUPDATE(k, w^{t-1})
  end
  $w^t \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_k^t$ ▷ Aggregating clients updates
end
return w^T
end procedure

4 Experiment

4.1 Datasets

We used the IEEE dataset on heart disease integrated by R. Alizadehsani et al. This dataset was collated from five separate datasets, including the UCI, and contains 1190 instances and 11 features. Previous research into heart disease prediction has often been based on small, outdated data, yet the huge accumulation of real-world data and the development of deep learning calls for new research.

In the field of deep learning diagnostics for cardiac disease, image-based diagnoses tend to be more accurate, yet digital data that can be easily detected by wearable devices, for example, is often easily accessible. If a privacy-preserving deep learning framework can be implemented for diagnosis of specific diseases, patient self-testing and self-prevention will be facilitated.

With regard to data pre-processing, we first coded the dataset with One-Hot Encoding and then normalized it. We then split the data into four groups containing three Clients and a test set for the purposes of federated learning. Because there is a correlation between the data, we multiply between features and combine two by two into two-dimensional data.

4.2 Data Segmentation

We randomly divided the dataset into a training set and a test set, with the training set accounting for 75% (473 patient data and 427 healthy patient data) and the test set accounting for 25% (140 patient data and 150 healthy patient data). Depending on the distribution of the data volume, we divided the training set into multiple subsets in different proportions to be distributed to different clients. All our experiments were done using 3 clients. In order to simulate the real data volume distribution, we divided the training set into 1:1:1, 2:3:4, and 1:3:5 ratios so that we could observe the results of the model with different data volume distributions and test how well the model adapts to the different ratios of data.

4.3 Model Construction

We chose a convolutional neural network as the classification model in our experiments, which are binary classification problems, and we chose cross-entropy loss as the loss function and used Adam as the optimiser to train the model. Our goal is to demonstrate that federated learning of deep convolutional neural network models can benefit from rich private data sharing while preserving privacy. Therefore, the network structure of the model was not the direction we explored in depth. We used AlexNet and ResNet18 as classification models before employing LeNet. Different models can slightly improve or degrade the performance of classification, and there was no significant performance gap between the different models on the same test set, so we favoured simpler models to solve the problem.

The convolutional neural network consists mainly of two convolutional blocks and three fully connected layers. The training set is pre-processed and each patient's data is transformed into $19 \star 19 \star 3$ 2D data, which is passed through two convolutional blocks to obtain $3 \star 4 \star 4$ 2D data, which is expanded into the fully-connected layer, resulting in two outputs to represent the probability of being healthy and diseased.

Diagnostic models for heart disease need to be highly accurate to assist doctors in guiding patient treatment while trying to avoid unnecessary panic for healthy patients. Therefore, in order to better assess the performance of our proposed model, all our experiments were repeated 3 times and the values such as accuracy that appear below are the average results of 3 experiments. We also used various evaluation metrics to evaluate our model.

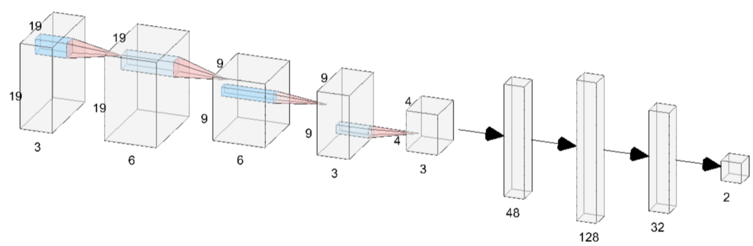

Figure 3: LeNet Framework

4.4 Results and Model Evaluation

4.4.1 Federated Framework Measurement

Again using the LeNet framework, both federated and centralised learning converge to an accuracy of around 0.9, with federated learning being only a limited value less accurate than centralised learning. However, the individual Clients only converge to 0.83 without federation as shown in Figure 4. Our federated learning framework is a significant improvement over approaches that train separately on individual clients. We can conclude that federated learning after raising the dimensional can achieve similar results to centralised learning while at the same time ensuring privacy.

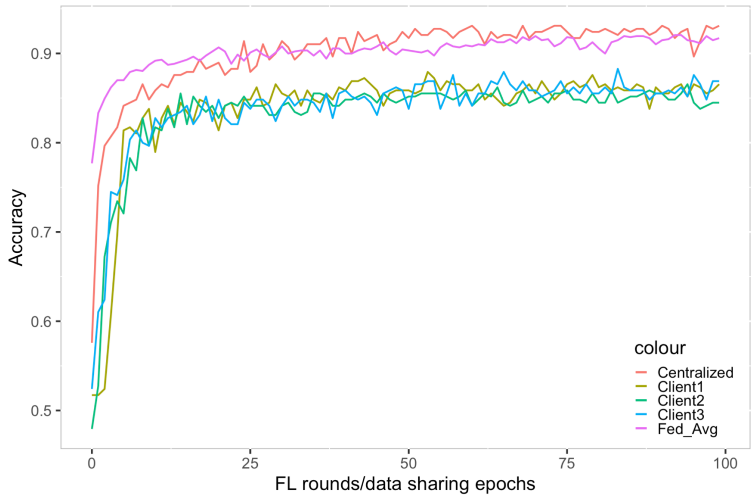

Figure 4: Training Process

4.4.2 Data Imbalance Adaptability

Variations in the amount of data available to the Client are common in reality, and imbalances in data size can affect the effectiveness of the model to some extent. It is therefore important to measure how well the model performs with different data variances. We set up three sets of Client data with different degrees of imbalance, consisting of 300:300:300, 200:300:400, and 100:300:500 respectively, and after training, we found that their convergence points only differed slightly at around 0.9. As shown in Figure 5, the model performs satisfactorily with different data ratios.

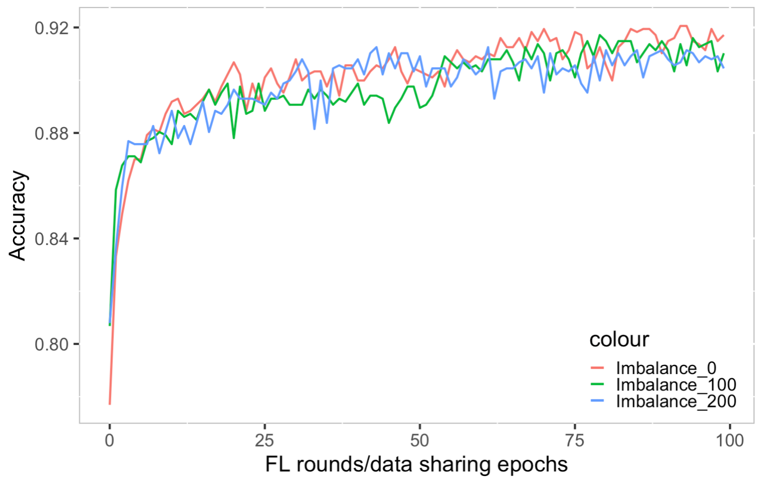

Figure 5: Data Imbalance

4.4.3 Metrics and Evaluation

Based on the training on the evenly split dataset, as seen in Figure 6, the basic region was stable after 75 rounds of communication, and we randomly selected the model once after the parameters had converged to evaluate the accuracy of the model on our test set from various aspects. The model achieved an accuracy of 92.41% on the test set, with model sensitivity and specificity of 94% and 91% respectively. The centralised learning model achieved an accuracy of 93.7%, with sensitivity and specificity of 90% and 97.1% respectively.

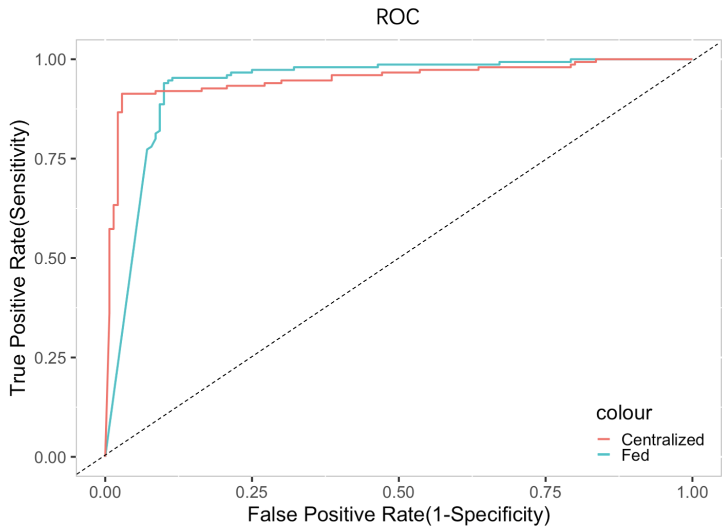

Figure 6: ROC

The following Table 1 shows a number of evaluation metrics for the models generated under both frameworks. This result highlights the effectiveness of our proposed FL-based approaches, as they provide comparable performance to centralised approaches while maintaining data privacy, indicating their suitability for privacy-constrained applications. Moreover, the models obtained from federated learning have significantly better recall rates in patients than centralised learning, which means that federated learning is better at identifying people with the disease and can provide treatment to patients earlier. Research has shown that some heart diseases can be completely cured or remitted if scientific and correct intervention and treatment are given early in the course of the disease.

Table 1: Evaluation Result

	Fed_Avg	Centralized
Accuracy	92.41	93.79
Sensitivity	94.02	90.66
Specificity	91.02	97.14
auc	0.936	0.938
precision	0.91	0.97
recall	0.94	0.90
f1-score	0.92	0.97

5 Future and Conclusion

Our results provide a good framework for the use of federated learning for deep learning medical diagnostics. However, constrained by the total number of available data, our study cannot implement more complex and practically meaningful models. If more data becomes available in the future, privacy-preserving cardiac diagnostics and federated learning will enter a new phase.

References

[1] Li Li et al. "A review of applications in federated learning". In: *Computers & Industrial Engineering* 149 (2020), p. 106854.

[2] Nicola Rieke et al. "The future of digital health with federated learning". In: *NPJ digital medicine* 3.1 (2020), pp. 1–7.

[3] Wenqi Li et al. "Privacy-preserving federated brain tumour segmentation". In: (2019), pp. 133–141.

[4] Abhijit Guha Roy et al. "Braintorrent: A peer-to-peer environment for decentralized federated learning". In: *arXiv preprint arXiv:1905.06731* (2019).

[5] Tariq Bdair, Nassir Navab, and Shadi Albarqouni. "Fedperl: Semi-supervised peer learning for skin lesion classification". In: (2021), pp. 336–346.

[6] Weishan Zhang et al. "Dynamic-fusion-based federated learning for COVID-19 detection". In: *IEEE Internet of Things Journal* 8.21 (2021), pp. 15884–15891.

[7] Ittai Dayan et al. "Federated learning for predicting clinical outcomes in patients with COVID-19". In: *Nature medicine* 27.10 (2021), pp. 1735–1743.

[8] Rajesh Kumar et al. "Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging". In: *IEEE Sensors Journal* 21.14 (2021), pp. 16301–16314.

[9] Peter Kairouz et al. "Advances and open problems in federated learning". In: *Foundations and Trends® in Machine Learning* 14.1–2 (2021), pp. 1–210.

[10] Yejin Kim et al. "Federated tensor factorization for computational phenotyping". In: (2017), pp. 887–895.

[11] Theodora S Brisimi et al. "Federated learning of predictive models from federated electronic health records". In: *International journal of medical informatics* 112 (2018), pp. 59–67.

[12] Akis Linardos et al. "Federated Learning for Multi-Center Imaging Diagnostics: A Study in Cardiovascular Disease". In: *arXiv preprint* (2021).

[13] M Akhil Jabbar, Bulusu Lakshmana Deekshatulu, and Priti Chandra. "Prediction of heart disease using random forest and feature subset selection". In: (2016), pp. 187–196.

[14] Mai Shouman, Tim Turner, and Rob Stocker. "Using decision tree for diagnosing heart disease patients". In: (2011), pp. 23–30.

[15] K Vembandasamy, R Sasipriya, and E Deepa. "Heart diseases detection using Naive Bayes algorithm". In: *International Journal of Innovative Science, Engineering & Technology* 2.9 (2015), pp. 441–444.

[16] Brendan McMahan et al. "Communication-efficient learning of deep networks from decentralized data". In: (2017), pp. 1273–1282.

[17] Qiang Yang et al. "Federated learning". In: *Synthesis Lectures on Artificial Intelligence and Machine Learning* 13.3 (2019), pp. 1–207.