FeDeFo: A Personalized Federated Deep Forest Framework for Alzheimer's Disease Diagnosis

Chen Qian*, Hui Xiong† and Jiyun Li‡
School of Computer Science and Technology, Donghua University
Shanghai, China
Email: *chen.qian@dhu.edu.cn, †xionghui0310@foxmail.com, ‡jyli@dhu.edu.cn

***Abstract*—Alzheimer's disease (AD) is a neurodegenerative disease that severely affects cognition, memory, and behavior and is incurable. Mild cognitive impairment (MCI) is a clinical precursor to AD, and early diagnosis of AD is essential for the prevention and intervention of disease progression. The hippocampus is one of the first brain regions affected by AD, and therefore structural magnetic resonance images (sMRI) are commonly used to measure the shape and volume of the hippocampus. In this paper, we propose a federal deep forest model called FeDeFo for calculating hippocampal volume using sMRI images to achieve AD classification. Firstly, to effectively protect data privacy, we use a federated learning framework to collaboratively train a gradient boosting decision tree (GBDT) model based on the local data of each client. In addition, to address the data discrepancy between clients, we introduce a deep forest model to exploit the local data beyond local interactions further and fuse it with the federally trained GBDT to personalize the model for each client. The experiments demonstrate that our proposed approach is able to personalize the model while protecting the data privacy of each client, providing a new idea for AD classification.**

Keywords—federated deep forest; personalized federated learning; Alzheimer's disease diagnosis.

I. Introduction

Alzheimer's disease (AD) is a fatal neurodegenerative disorder with insidious onset in the presenile period [1]. It affects memory and cognitive ability in an irreversible manner and gradually causes the decline of the quality of daily life and social functions. Although developed AD has no cure, it can be delayed or even prevented at the earliest stage, known as mild cognitive impairment (MCI) [2]. Thus, distinguishing MCI from AD or normal cognition (NC) is an urgent need in AD diagnosis. So far, neuroimaging is the best non-invasive technique to look for abnormalities in the human brain [3].

Undoubtedly, deep learning has revolutionized image processing. It can solve difficult problems such as image colorization, classification, segmentation, and detection [4]. For example, a deep neural network (DNN) is a stack of multiple layers, which allows models to become more efficient at learning complex features and performing more intensive computational tasks. It outshines classic machine learning paradigms in machine perception tasks involving unstructured data. However, DNN has a number of well-known limitations, e.g., living off a considerable amount of data, requiring enormous computational resources, and lacking theoretical explanations [5]. For the above reasons, researchers are looking for an alternative paradigm. Eventually, deep ensemble learning gets attention, such as GrowNet [6], S-DNN [7], DSN [8], deep forest [9], because it combines the advantages of both the DNNs as well as ensemble learning such that the final model has better generalization performance [10].

However, there is another challenge in the healthcare area, which is the well-known privacy problem [11]. Traditional deep learning generally proceeds in two phases: collecting data from different participants for preprocessing and feeding the data to a monolithic model for training. As a result, the risk of privacy leakage is unavoidable. To tackle the problem, we present a federated learning framework. It enables distributed clients to collaboratively train a shared model without sharing their training data. To be specific, model parameters are computed locally by each client device and exchanged with a central server, which aggregates the local models for a global view [12]. It is worth noting that our framework provides personalization in order to improve the privacy-accuracy tradeoffs and balance the benefits among different parties [13].

DOI reference number: 10.18293/SEKE2023-013

II. Related Work

A. Federated Learning

The concept of federated learning was first proposed by Google AI [14], which enables mobile terminals to jointly train a global machine learning model in a decentralized manner without sharing individual data. Thereby, federated learning is an interdisciplinary domain of machine learning and privacy computing. In terms of training samples, it can be broadly divided into two categories: horizontal federated learning and vertical federated learning [12]. The former refers to datasets owned by different parties that share the same feature space but differ in samples, whereas the latter refers to datasets owned by different parties that differ not only in samples but also in feature space.

In the healthcare domain, horizontal federated learning is more frequently used as clinical task demands increase [15]. For instance, PRCL coordinates multiple medical institutions and cloud servers to develop an electronic health records (EHR) system since the hospitals run a neural network using their own records, and the cloud server aggregates the update parameters [16]. Likewise, horizontal federated learning has been used in breast cancer prediction [17], blood pressure estimation [18], skin disease detection [19], etc. Recently, experts successfully built a powerful model for COVID-19 screening by federated training using chest X-ray images from different hospitals [20]. In this paper, the proposed FeDeFo framework also adopts horizontal federated learning.

Regarding application scenarios, the types of federated learning mainly include cross-device and cross-silo federated learning [21]. The former usually requires massive mobile devices as trainers, each owning a small amount of raw data, while the latter involves few reliable organizations, each holding medium to large datasets. In this paper, we study cross-silo federated learning due to its suitability for healthcare scenarios [22]. Thereby, our work does not need to pay attention to the scheduling problem and communication bottleneck. In addition, we can assume that the central server is an honest and reliable third party in the training process.

General federated learning approaches face several fundamental challenges. One of them is that the global model can only capture the statistical characteristics of different parties rather than the unique personal styles. For example, differences in age-adjusted Alzheimer's dementia prevalence exist among regions of the world due to the combination of low average educational attainment and high vascular risk profile [23]. Another is the heterogeneous computing resources and network conditions of different federated learning devices. Fortunately, both challenges can be markedly alleviated by personalized federated learning [24]. Furthermore, since medical data are highly sensitive and valuable, personalized federated learning can even more amplify the model quality. For example, FedHome put forward an edge cloud federated learning architecture for home healthcare services, which allows a client to train a personalized model by the global model and its private data [25]. In this paper, the proposed FeDeFo framework also adopts personalized federated learning in order to heighten the patient experience.

B. Ensemble Learning

Ensemble learning is not a machine learning algorithm but a technical framework that aggregates outputs of multiple models in order to improve the overall performance and generalization. In general, there are three primary ensemble techniques, viz., bagging, boosting, and stacking [10].

Boosting is an influential ensemble methodology referring to a family of algorithms that convert a cluster of weak learners into a strong one. Unlike bagging, boosting can reduce bias by learning in a sequence that iteratively adjusts the weight of observation as per the last classification [26]. Gradient boosting decision tree (GBDT) is a classic additive model that uses a boosting ensemble of decision trees to predict a target label [27]. More specifically, it uses a forward distribution algorithm to perform greedy learning. In each iteration of learning, a classification and regression tree (CART) is used to fit the residuals of the previous one.

Random forest is an enhanced version of the decision trees, which uses the bagging strategy to build multiple decision trees and aggregate them for an accurate result with as little bias as possible [28]. As an evolution of random forest, gcForest achieved an innovation breakthrough by constructing a multi-layer neural network in which a number of random forests, instead of neurons, are embedded in each layer [9]. Overall, gcForest becomes more attractive due to the simple algorithm, fewer initial hyperparameters, and thorough extraction of feature relations. Furthermore, it can ensure higher prediction accuracy with a smaller dataset, adapt to different situations by automatically settling its complexity, and moderate the overfitting issue through its robustness. At present, gcForest and its extensions have been increasingly utilized in the real world. For example, a revised gcForest, namely BCDForest, is proposed to classify cancer subtypes based on small biological datasets [29]. Another one can successfully identify ADHD and control subjects [30].

It is worth noting that any type of classifier is applicable in gcForest, such as GBDT. Additionally, GBDT re-weights the original training sample in every boosting step, so it has an excellent generalization ability suitable for solving regression problems similar to disease diagnosis. Therefore, we tailor-make a deep ensemble algorithm by integrating GBDT into gcForest to diagnose Alzheimer's dementia.

III. FeDeFo Framework

A. Overview

This paper proposes a novel and practical framework called personalized federated deep forest framework (abbreviated as FeDeFo), which aims to classify subjects with AD, MCI, and NC while preserving privacy. In FeDeFo, multiple clients have their own data that share the same features. Thus, FeDeFo focuses on scenarios suitable for horizontal federated learning. In addition, FeDeFo supports personalization, so each client can have a slightly different local model and hence supply a better customer experience.

We are given data from N different clients, which are denoted by $\{C_1, C_2, \dots, C_N\}$, while the data they provide are denoted by $\{D_1, D_2, \dots, D_N\}$. Conventional methods train a model M_{ALL} by combining all the data $D = D_1 \cup D_2 \cup \dots \cup D_N$. All the data have different distributions. In our problem, we want to collaborate all the data to train a federated model M_{FED}, where any client C_i does not expose its data D_i to each other. If we denote the accuracy as $\mathcal{A}$, then the objective of our model is to ensure the accuracy of federated learning is close to that of conventional learning denoted by:

$$|\mathcal{A}_{FED} - \mathcal{A}_{ALL}| < \Delta,$$

where Δ is an extremely small non-negative real number.

The FeDeFo framework aims to achieve accurate AD classification tasks through federated learning and deep forests without compromising privacy security. There are two participants in FeDeFo: the central server and the client. Every client keeps its patients' data safe by preventing any other participants from accessing it. In this paper, we choose the GBDT model for local training. As shown in Figure 1, in

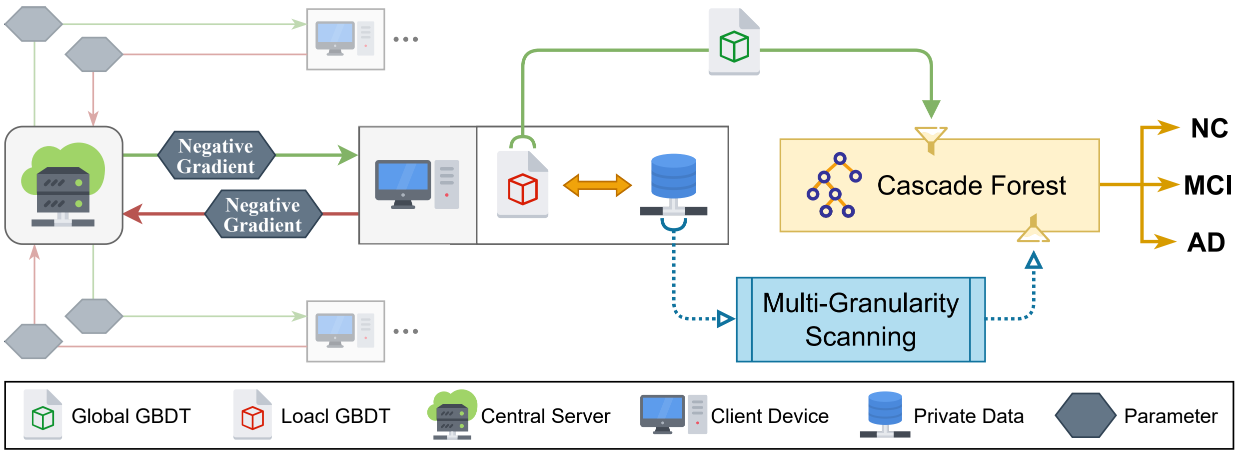

Fig. 1. An overview of FeDeFo framework.

each training iteration, the central server broadcasts the model to every client. Then, each client trains its own GBDT model using its data and uploads the gradients to the central server. Finally, the server aggregates the gradients in order to upgrade the global model. After the federated training stops, every client weaves the resulting model to the local deep forest model, i.e., gcForest, aiming to train a personalized model and further provide a more accurate AD diagnosis system. To summarize, the FeDeFo framework overcomes information segregation through higher-order information scattered across different clients without exchanging their privacy data.

B. GBDT

GBDT is an ensemble model which trains a sequence of decision trees. Formally, a dataset D with n instances and d features can be described as $D\{(x_i, y_i)\}$ ($|D| = n, x_i \in \mathbb{R}^d, y_i \in \mathbb{R}$). So, the output can be predicted via K-additive functions as follows:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \;\; f_k \in \mathcal{F},$$

where $\mathcal{F} = \{f(x) = \omega_{q(x)}\}$ ($q : \mathbb{R}^d \to T, \omega \in \mathbb{R}^T$) is the space of regression trees. Here q denotes the structure of each tree that maps an instance to the corresponding leaf index. T is the number of leaves in the tree. Each f_k corresponds to an independent tree structure q and leaf weights ω. Moreover, we use ω_i to represent the score on the i-th leaf. For a given example, we use the decision rules in the trees (given by q) to classify it into the leaves and perform the final prediction by summing up the score in the corresponding leaves (given by ω). To learn the set of functions used in the model, we minimize the following regularized objective as follows:

$$\mathcal{L} = \sum_{i} l(\hat{y}_i, y_i) + \sum_{k} \Omega(f_k),$$

where l is a differentiable convex loss function that measures the difference between the prediction $\hat{y}_i$ and the target y_i. We hereby define $\Omega(f_k) = \gamma T_l + \frac{1}{2}\lambda\|\omega\|^2$ as a regularization term to penalize the complexity of the model, in which γ and λ are hyperparameters. Hence, GBDT minimizes the following objective function at the t-th iteration as follows:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t),$$

where $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ and $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ are the first-order and second-order gradient statistics on the loss function.

C. Federated Training

Structurally speaking, GBDT is a collection of decision trees constructed in a serial manner. However, it is generally implemented and trained as an integral whole under a standalone environment. Thereby, in FeDeFo, GBDTs are revised to coincide with the decentralized environment. As Figure 2 shows, the process of updating the gradients is modified with the aim of receiving updates from the external.

For each iteration of federated training, each client starts with initializing a local GBDT model at first and feeds the model with private data subsequently. In the GBDT, each decision tree is constructed by fitting the negative gradient of the previous one. Hence, the complete GBDT is shaped as a chain of decision trees. When a client finishes the local training, it calculates the first-order and second-order gradient statistics of the loss function. Then, both gradients are uploaded to the server for aggregation. Eventually, the central server collects all clients' feedback and updates the global GBDT model. Notably, the server will continue to broadcast the aggregated gradients to every client. Therefore, in the next iteration, those gradients become the building blocks of the new local GBDT. The aforementioned procedure is repeated until the global model has converged and the federated learning result is satisfactory.

D. Personalized Training

Federated learning breakthroughs the data silos by collaborative model training with decentralized datasets. However,

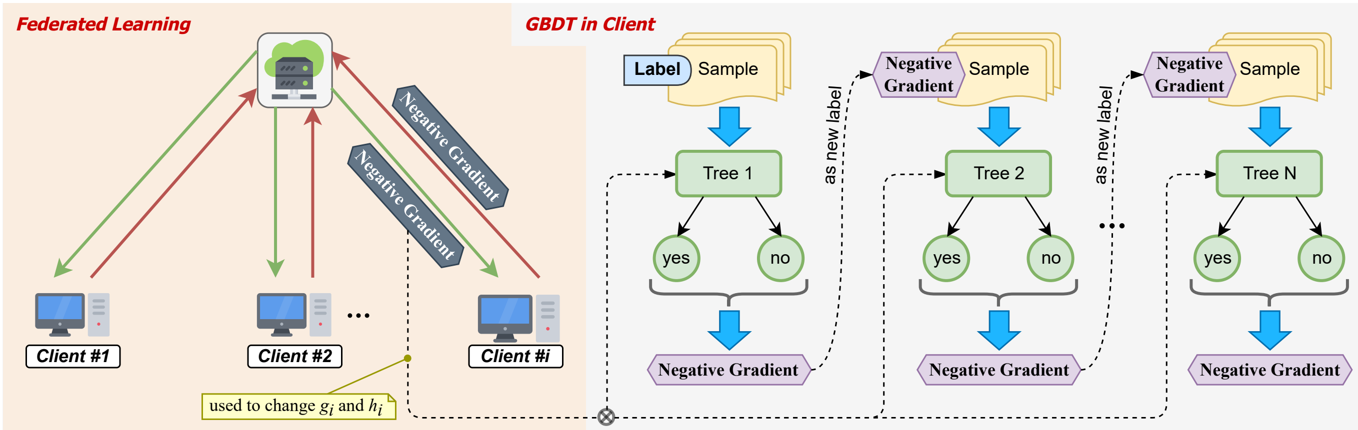

Fig. 2. Training process of federated GBDT.

its popularity is not growing in practical terms because the global model usually performs unsatisfactorily on participant-specific data. On this account, the FeDeFo framework provides a personalized training method allowing the model to learn fine-grained information from the ad-hoc client as well as the coarse-grained features from all participants. Inspired by gcForest, we hereby design a personalization model using multi-grained scanning and cascading forests with a view to achieving high performance on feature representation learning with high-dimensional data in the context of AD classification.

As Figure 3 illustrates, the personalization model takes the client's private data as input. Then, the raw features of input are extracted and processed by a multi-grained scanner to generate feature vectors, which are further sent to the cascade forest to complete the classification task. Notably, the original cascade forest in gcForest is defined by multi-level integration of decision trees, which are theoretically replaceable by any other classifiers that can output class distribution vectors. Here, we adopt GBDTs instead of decision trees to construct the personalization model. In detail, each cascade level contains four GBDTs, each of which outputs a class distribution vector. Subsequently, the four vectors at the same level and the output of the multi-grained scanner are concatenated. The result becomes the input vector of the next level. Finally, the classification result is the maximum of the average class vectors outputted from the last level.

IV. Experiment

A. Dataset

The data used in this paper are extracted from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [31] database. As demonstrated in Table I, a total of 400 subjects are involved, including 93 AD patients, 201 MCI patients, and another 106 normals as the control group. Because the hippocampal atrophy can evidently suggest the AD diagnosis [32], we select their sMRI with magnetization-prepared rapid gradient-echo (MP-RAGE). Such an image can offer a clear vision of the hippocampus, which is advantageous for volume estimation.

TABLE I
Subject composition

Type	Amount	Age	Sex (Male/Female)
AD	93	75.15±8.14	50 / 43
MCI	201	73.61±6.92	133 / 68
NC	106	76.16±7.14	54 / 52

B. Experimental Process

The hippocampal volume is crucial to diagnose AD and MCI. In this paper, we classify the patients with AD, MCI, and NC by assessing the relative hippocampal volume, which is calculated by the whole brain volume dividing the absolute hippocampal volume. The relative volume can eliminate the difference in brain volume between different populations.

The brain images are initially preprocessed by the FSL tool [33], including intensity inhomogeneity correction, skull removal, intensity normalization, and image cutting. The unified image size is $196 \times 271 \times 181$. After that, we can use the 3DUnet-CBAM model [34] to extract the hippocampal part. Finally, we use IBASPM [35] to label the voxel of images in a neuroanatomical manner and further calculate the volumes automatically. The activities in data processing are briefly expressed in Figure 4.

Subsequently, we input the relative hippocampal volume to the GBDT models deployed on both central server and client devices. During the learning process, 70 percent of the data are used for training, whereas the rest are for model evaluation. Once the federated GBDT is mature, it will be adopted as a classifier and loaded into the deep forest for personalization. Finally, every client can use its own model to diagnose AD and MCI.

C. Experimental Results

We also perform experiments on AD classification using deep learning methods (e.g., LeNet and VGGNet), and machine learning method (e.g., SVM). In comparison with those

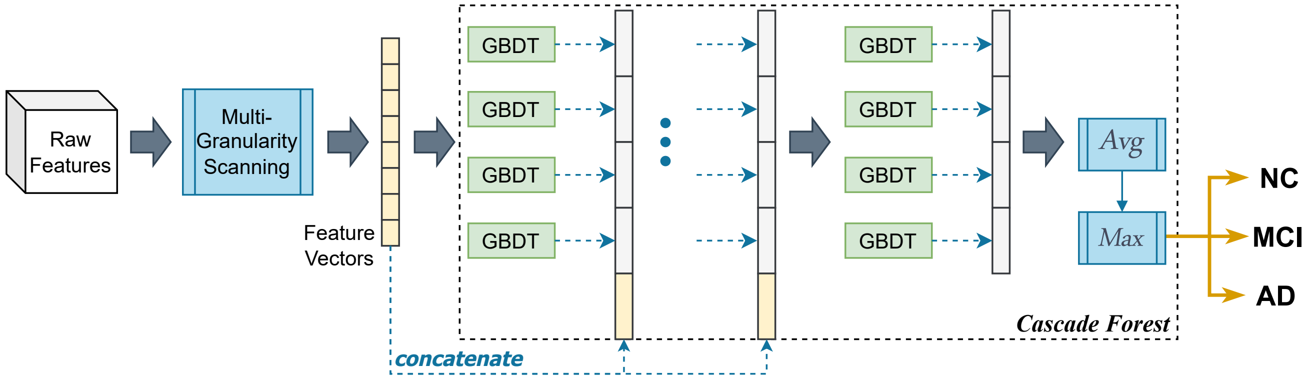

Fig. 3. Personalized model training.

experimental results, we prove the outperformance of our FeDeFo framework. Moreover, to demonstrate the effectiveness of the personalization mechanism in FeDeFo, we conduct an extra experiment using federated GBDT only, i.e., an equivalent of FeDeFo without personalized training.

To thoroughly investigate the classification performance, we use three metrics for evaluation: accuracy (ACC), sensitivity (SEN), and specificity (SPE). In clinical practice, accuracy measures how correctly a diagnostic test identifies and excludes a given condition, yet sensitivity evaluates how good the test is at detecting a positive disease, whereas specificity estimates how likely patients without the disease can be correctly ruled out. If we denote TP as true-positive samples, FP as false-positive samples, TN as true-negative samples, and FN as false-negative samples, the equations of the metrics can be defined by:

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

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{FP + TN}$$

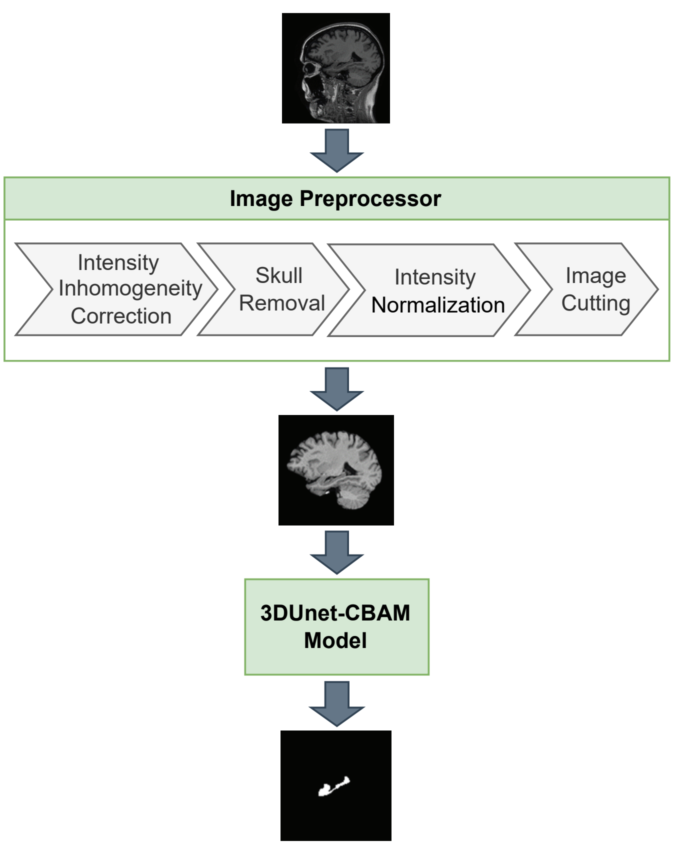

Fig. 4. Data processing.

As shown in Table II, FeDeFo outshines LeNet and VGGNet on every metric. The sensitivity and specificity of classifying MCI and NC are slightly lower than SVM, probably because the sample size is too small. Meanwhile, it improves the average results by 6.1% compared to the federated GBDT. In summary, our FeDeFo framework achieves exceptional performance in classifying subjects with AD, MCI, or NC while preserving patients' privacy.

TABLE II
COMPARISON OF CLASSIFICATION RESULTS

	AD vs. NC (%)			MCI vs. NC (%)		
	ACC	SEN	SPE	ACC	SEN	SPE
LeNet	83.8	73.2	92.4	68.8	79.3	63.4
VGGNet	84.7	77.3	90.8	70.9	81.9	65.2
SVM	83.3	71.4	92	72	**84.1**	**71.3**
Fed. GBDT	86.4	76.5	89.3	65.3	73.4	61.6
FeDeFo	**91.6**	**85.7**	**93.1**	**72.4**	83.5	68.7

V. CONCLUSION

The FeDeFo framework provides a rational combination of deep learning, ensemble learning, federated learning, and personalization. Together, they suggest a promising solution for privacy-preserving AI services. In this paper, we use FeDeFo to produce a family of AD diagnosis systems. The slight difference between them makes each system more suitable for its own scenario. Theoretically, the FeDeFo framework is extensible to other medical fields or even more. In the future, we will test our FeDeFo framework on distinct datasets using different training models and personalization techniques in more domains.

ACKNOWLEDGMENTS

This work has been supported by the National Key R&D Program of China under Grant 2019YFE0190500, the Fundamental Research Funds for the Central Universities of Ministry of Education of China (Grant No.2232021D-22), and the Initial Research Funds for Young Teachers of Donghua University.

REFERENCES

[1] H. Hippius and G. Neundörfer, "The discovery of Alzheimer's disease," *Dialogues in clinical neuroscience*, 2022.

[2] V. Lissek and B. Suchan, "Preventing dementia? Interventional approaches in mild cognitive impairment," *Neuroscience & Biobehavioral Reviews*, vol. 122, pp. 143–164, 2021.

[3] A. L. Dallora, S. Eivazzadeh, E. Mendes, J. Berglund, and P. Anderberg, "Prognosis of dementia employing machine learning and microsimulation techniques: A systematic literature review," *Procedia Computer Science*, vol. 100, pp. 480–488, 2016.

[4] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification in BioApps*, pp. 323–350, 2018.

[5] S. Sood, H. Singh, M. Malarvel, and R. Ahuja, "Significance and limitations of deep neural networks for image classification and object detection," in *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE, 2021, pp. 1453–1460.

[6] S. Badirli, X. Liu, Z. Xing, A. Bhowmik, K. Doan, and S. S. Keerthi, "Gradient boosting neural networks: GrowNet," *arXiv preprint arXiv:2002.07971*, 2020.

[7] C.-Y. Low, J. Park, and A. B.-J. Teoh, "Stacking-based deep neural network: deep analytic network for pattern classification," *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 5021–5034, 2019.

[8] C. Sun, M. Ma, Z. Zhao, and X. Chen, "Sparse deep stacking network for fault diagnosis of motor," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3261–3270, 2018.

[9] Z.-H. Zhou and J. Feng, "Deep forest," *National Science Review*, vol. 6, no. 1, pp. 74–86, 2019.

[10] M. A. Ganaie, M. Hu *et al.*, "Ensemble deep learning: A review," *arXiv preprint arXiv:2104.02395*, 2021.

[11] K. M. Chong, "Privacy-preserving healthcare informatics: A review," in *ITM Web of Conferences*, vol. 36. EDP Sciences, 2021, p. 04005.

[12] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, "Federated learning," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 13, no. 3, pp. 1–207, 2019.

[13] A. Bietti, C.-Y. Wei, M. Dudik, J. Langford, and S. Wu, "Personalization improves privacy-accuracy tradeoffs in federated learning," in *International Conference on Machine Learning*. PMLR, 2022, pp. 1945–1962.

[14] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[15] C.-R. Shyu, K. T. Putra, H.-C. Chen, Y.-Y. Tsai, K. T. Hossain, W. Jiang, and Z.-Y. Shae, "A systematic review of federated learning in the healthcare area: From the perspective of data properties and applications," *Applied Sciences*, vol. 11, no. 23, p. 11191, 2021.

[16] M. Hao, H. Li, G. Xu, Z. Liu, and Z. Chen, "Privacy-aware and resource-saving collaborative learning for healthcare in cloud computing," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.

[17] C. Beguier, J. O. d. Terrail, I. Meah, M. Andreux, and E. W. Tramel, "Differentially private federated learning for cancer prediction," *arXiv preprint arXiv:2101.02997*, 2021.

[18] E. Brophy, M. De Vos, G. Boylan, and T. Ward, "Estimation of continuous blood pressure from PPG via a federated learning approach," *Sensors*, vol. 21, no. 18, p. 6311, 2021.

[19] M. A. Hashmani, S. M. Jameel, S. S. H. Rizvi, and S. Shukla, "An adaptive federated machine learning-based intelligent system for skin disease detection: A step toward an intelligent dermoscopy device," *Applied Sciences*, vol. 11, no. 5, p. 2145, 2021.

[20] I. Feki, S. Ammar, Y. Kessentini, and K. Muhammad, "Federated learning for COVID-19 screening from chest X-ray images," *Applied Soft Computing*, vol. 106, p. 107330, 2021.

[21] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, "Advances and open problems in federated learning," *Foundations and Trends® in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[22] N. Rieke, J. Hancox, W. Li, F. Milletari, H. R. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. A. Landman, K. Maier-Hein *et al.*, "The future of digital health with federated learning," *NPJ digital medicine*, vol. 3, no. 1, pp. 1–7, 2020.

[23] L. Rizzi, I. Rosset, and M. Roriz-Cruz, "Global epidemiology of dementia: Alzheimer's and vascular types," *BioMed research international*, vol. 2014, 2014.

[24] A. Z. Tan, H. Yu, L. Cui, and Q. Yang, "Towards personalized federated learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[25] Q. Wu, X. Chen, Z. Zhou, and J. Zhang, "Fedhome: Cloud-edge based personalized federated learning for in-home health monitoring," *IEEE Transactions on Mobile Computing*, 2020.

[26] H. A. Ali, C. Mohamed, B. Abdelhamid, N. Ourdani, and T. El Alami, "A comparative evaluation use bagging and boosting ensemble classifiers," in *2022 International Conference on Intelligent Systems and Computer Vision (ISCV)*. IEEE, 2022, pp. 1–6.

[27] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, R. Tibshirani, and J. Friedman, "Boosting and additive trees," *The elements of statistical learning: data mining, inference, and prediction*, pp. 337–387, 2009.

[28] Y. Liu, Y. Wang, and J. Zhang, "New machine learning algorithm: Random forest," in *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3*. Springer, 2012, pp. 246–252.

[29] Y. Guo, S. Liu, Z. Li, and X. Shang, "BCDForest: A boosting cascade deep forest model towards the classification of cancer subtypes based on gene expression data," *BMC bioinformatics*, vol. 19, no. 5, pp. 1–13, 2018.

[30] L. Shao, D. Zhang, H. Du, and D. Fu, "Deep forest in ADHD data classification," *IEEE Access*, vol. 7, pp. 137 913–137 919, 2019.

[31] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 27, no. 4, pp. 685–691, 2008.

[32] G. Halliday, "Pathology and hippocampal atrophy in Alzheimer's disease," *The Lancet Neurology*, vol. 16, no. 11, pp. 862–864, 2017.

[33] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012.

[34] S. Li, L. Zhao, J. Li, and Q. Chen, "Segmentation of hippocampus based on 3DUnet-CBAM model," in *2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*. IEEE, 2021, pp. 595–599.

[35] S.-H. Han, M.-A. Lee, S. S. An, S.-W. Ahn, Y. C. Youn, and K.-Y. Park, "Diagnostic value of Alzheimer's disease-related individual structural volume measurements using IBASPM," *Journal of Clinical Neuroscience*, vol. 21, no. 12, pp. 2165–2169, 2014.