mathematics

MDPI

Article

A Robust Distributed Deep Learning Approach to Detect Alzheimer's Disease from MRI Images

Tapotosh Ghosh [1,†], Md Istakiak Adnan Palash [2,†], Mohammad Abu Yousuf [3,*], Md. Abdul Hamid [4], Muhammad Mostafa Monowar [4] and Madini O. Alassafi [4]

[1] Department of Computer Science and Engineering, United International University, Dhaka 1212, Bangladesh; tapotosh@cse.uiu.ac.bd
[2] Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka 1216, Bangladesh; 19511057@student.bup.edu.bd
[3] Institute of Information Technology, Jahangirnagar University, Savar 1342, Bangladesh
[4] Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia; mabdulhamid1@kau.edu.sa (M.A.H.); mmonowar@kau.edu.sa (M.M.M.); malasafi@kau.edu.sa (M.O.A.)
* Correspondence: yousuf@juniv.edu
† These authors contributed equally to this work.

Abstract: Alzheimer's disease has become a major concern in the healthcare domain as it is growing rapidly. Much research has been conducted to detect it from MRI images through various deep learning approaches.However, the problems of the availability of medical data and preserving the privacy of patients still exists. To mitigate this issue in Alzheimer's disease detection, we implement the federated approach, which is found to be more efficient, robust, and consistent compared with the conventional approach. For this, we need deep excavation on various orientations of MRI images and transfer learning architectures. Then, we utilize two publicly available datasets (OASIS and ADNI) and design various cases to evaluate the performance of the federated approach. The federated approach achieves better accuracy and sensitivity compared with the conventional approaches in most of the cases. Moreover, the robustness of the proposed approach is also found to be better than the conventional approach. In our federated approach, MobileNet, a low-cost transfer learning architecture, achieves the highest 95.24%, 81.94%, and 83.97% accuracy in the OASIS, ADNI, and merged (ADNI + OASIS) test sets, which is much higher than the achieved performance in the conventional approach. Furthermore, in the proposed approach, only the weights of the model are shared, which keeps the original MRI images in their respective hospital or institutions, preserving privacy in the healthcare domain.

Keywords: federated learning; Alzheimer's disease; medical imaging; MRI image

MSC: 68T07; 92C55

Citation: Ghosh, T.; Palash, M.I.A.; Yousuf, M.A.; Hamid, M.A.; Monowar, M.M.; Alassafi, M.O. A Robust Distributed Deep Learning Approach to Detect Alzheimer's Disease from MRI Images. *Mathematics* **2023**, *11*, 2633. https://doi.org/10.3390/math11122633

Academic Editor: Jakub Nalepa

Received: 27 April 2023
Revised: 29 May 2023
Accepted: 7 June 2023
Published: 9 June 2023

1. Introduction

Alzheimer's disease (AD) is among the most prevalent neurodegenerative diseases in the world. The risk of developing this condition increases in tandem with the patient's chronological age. Alzheimer's disease may be the underlying cause of dementia in about 6 million people in the US, where most of them are in the more than 65 years of age group. This disease is now the sixth biggest cause of death in the US [1]. The other causes of AD include genetic, behavioral, and environmental factors that gradually impact the brain [2]. As Alzheimer's disease is incurable, and early detection of this disease can be extremely beneficial. A variety of cognitive and behavioral tests are used to diagnose this disease. However, MRI images of the brain can provide important information for diagnosing this condition [3], since Alzheimer's disease affects a wide variety of cells and different parts

of our brain, and MRI images reveal changes in the region's structure or pattern. For this reason, researchers are now trying to develop a robust and efficient approach to analyze MRI images to detect AD through machines. They have proposed a wide range of models based on deep learning (DL), machine learning (ML), and hybrid methods. These ML and DL models need a lot of image data to train and also have a high computation cost [4]. If the dataset is not large enough, then the model may be biased towards a single class, and the robustness of the models cannot be provided. However, collecting data from the healthcare sector is usually challenging due to the sensitive nature of the data. Patients also remain hesitant to share their medical information [5]. Therefore, a secure method to train the deep learning models without collecting data has become a necessity to ensure the confidentiality and integrity of medical data.

McMahan et al. [6] proposed the federated learning (FL) method, which ensures the confidentiality and security of user data. This method involves training the model on the local server and then sending the weights of the trained model to the global server. This allows the global server to update its weight without knowing anything about the user data. In this case, the user does not need to provide any personal details to the server. As a result, the user's privacy is secured. Using this strategy additionally guarantees that user data will not be compromised under any circumstances. Moreover, this approach shows more robustness compared with the conventional approach [7]. This approach was designed for Object Detection and Handwritten Character Recognition and tested on the MNIST and CIFAR datasets. However, in our study, a federated-learning-based AD detection approach employing a pretrained MobileNet transfer learning architecture is proposed, which is robust, efficient, and privacy-preserving.

MobileNet [8] is a compact and simple model that takes less time to train and is capable of providing significant performance. In this research, MobileNet is trained in a federated way to diagnose Alzheimer's disease from brain MRI images. In this study, we used two distinct datasets, namely ADNI and OASIS. There are different planes of brain MRI images available in both of these datasets. The proposed model uses coronal plane images after an orientation selection procedure. To ensure the robustness of the model, different experiments are conducted by utilizing ADNI and OASIS datasets. They are trained individually and then tested on all of the test sets of ADNI and OASIS in each of the cases. Moreover, these datasets are combined and named as merged set. This merged set is used to explore the robustness of the model further. In the federated learning approach, each of these training processes on different datasets is performed solely on different local servers using two different datasets. After that, the weights of each of these trained models are forwarded to the main or central server. In the central server, a deep learning model is updated. This deep learning model is same in architecture and parameters, just like the client side, and it is called global model. The global model is then updated by using the average weights of both of the local trained models. Here, the user data need not to be shared to the researchers, and thus, the user's privacy and data confidentiality are preserved. In addition, the proposed FL-based model achieves higher levels of accuracy and sensitivity than the usual approach. The proposed method also shows a greater level of robustness compared with the conventional approach. This further enhances the acceptability of the proposed model.

The following contributions are made in this paper.

- An appropriate model and MRI image orientation selection are performed.
- A federated approach is proposed to train the model, and it is found to be better performing and more privacy-preserving.
- To ensure robustness and consistency, the proposed approach is validated against multiple datasets and multiple test cases.

The rest of this work is organized as follows. In Section 2, the existing works on Alzheimers' disease detection is discussed. Section 3 explains the methodology, and the results are discussed in Section 4. Finally, in Section 5, we provide concluding remarks and future goals.

2. Literature Review

There have been many Alzheimer's disease detection methods developed so far. Most of the techniques are implemented with ML- or DL-based approaches. These approaches are extremely effective in other domain as well, such as COVID-19 chest X-ray classification [9], brain tumor classification [10], image quality identification from ultrasound [11], recognition of human activities [12], and so on. By combining shearlet-based descriptors with deep features, Alinsaif et al. [13] suggested a technique for the representation of characteristics that might be used for the categorization of AD. Their model can be broken down into two distinct stages. In the first place, they preprocessed the MRI images and then extracted features from them. They utilized the SVM and DTB algorithms for classification purposes. Puente-Castro et al. [14] developed a hybrid model for the classification of AD. They used the sagittal plane of the brain MRI images. After collecting the dataset, they applied the ResNet model, extracting the features and then integrating age and sex features with them. Then, these extracted features were classified using the SVM classifier. Chui et al. [15] proposed a three-layered model called GAN-CNN-TL. In this particular investigation, they used all three available variants of the OASIS dataset. They employed GAN to generate synthetic data. Then, feature extraction from these images was performed using a CNN. After that, they applied transfer learning to the classification of the images.

Folego et al. [16] developed two different models, called ADNet and ADNet-DA, using a 3D CNN. Here, they first preprocessed the MRI images by extracting the skull and also normalizing the images. After that, they employed 3D-CNN in these processed images for extracting features and classifying them. They used four different pre-trained architectures here. Furthermore, they employed the domain adaptation method to evaluate the robustness of their model, which was evaluated against various test sets. Liu et al. [17] suggested using a depthwise separable convolutional-based model as an alternative to the conventional CNN model in order to make the model and its parameters more simple. They implemented CNN and DSC for the purpose of AD detection. In the DSC's convolution layer, filtering and feature extraction were kept independent from one another. Because of the potential for overfitting or underfitting with a smaller dataset, they additionally relied on AlexNet and GoogleNet pretrained models for categorization. An et al. [18] suggested a deep ensemble learning approach for classifying AD. Their dataset includes seven distinct feature categories. Their proposed approach performed better considering the other ensemble approach.

By combining a 3D CNN with 3D CLSTM, Xia et al. [19] were able to develop a unique model. Their model could only perform simple categorization tasks. They employed 12 layers in total (6 convolution layers and 6 max polling layers) in their CNN model. On the other hand, their CLSTM model was made up of three separate gates. This CLSTM model helped improve feature extraction from images. The model's overfitting problems were fixed by employing data augmentation techniques to make the images larger. Using Grad-cam, they also annotated the MRI images to show the area impacted by AD. Wei et al. [20] applied an adaptive histogram technique for increasing the contrast of the images. In this study, features were selected from these images using the t-test. They employed SVM and RF for the classification of AD images.

Lin et al. [21] proposed an voting ensemble-learning-based classification model consisting of the discriminators GAN, VGG16, and ResNet50. Later, they also performed a domain adaptation task to check the robustness of the model. Kaplan et al. [22] presented a new feed-forward method called LPQNet to reduce computational complexity with higher performance. They used a newly collected dataset in this work which contains 1070 subjects to train the model. Bringas et al. [23] used mobility data for the detection of AD. For this purpose, they collected data from 35 patients and identified them using the CNN model. Murugan et al. [24] developed a framework called DEMNET. This framework performed a multiclass classifier with an accuracy of 95.23%. Furthermore, they applied SMOTE to reduce the class imbalance problem. Lodha et al. [25] proposed an ML-based model for AD classification. They illustrated a comparison among different algorithms in that study.

Li et al. [26] also proposed a hybrid model using the CNN and RNN architectures. They used DenseNet as the CNN and BGRU as the RNN model. In this hybrid model, they evaluated the hippocampus structure and classified AD.

Basheera et al. [27] suggested a skull-stripping technique in the preprocessing stages for eliminating unnecessary tissues. They also applied the HEICA method for collecting segmented gray matter from MRI images. Bi et al. [28] developed a CAD system for AD detection using an unsupervised method. This method first extracted features from the images using a model called PCANet, which was built using the CNN model. After that, K-means clustering was performed for grouping these features as AD or normal images. Jabason et al. [29] implemented an ensemble model for identifying different stages of AD. They first trained the hybrid CNN model separately for learning features from the images. After that, a voting classifier was employed for classification purposes. Helaly et al. [30] implemented a 2D- and 3D-based model for AD classification. They used a CNN model for this purpose. Image preprocessing, resampling, data augmentation, etc., were also performed in their study. Venugopalan et al. [31] proposed a multimodal DL method to detect AD stages. They used a 3D CNN for extracting features from different modalities of clinical and image data. They showed that shallow ML algorithms are less accurate than their proposed model. They also found that integrating multimodal data enhance the performance of the proposed approach.

In all of the above-mentioned research works, the goals are to find an architecture that may perform well in Alzheimer's disease detection. These research works did not focus on preserving the privacy of the patients. Moreover, collecting data in the medical domain is costly and requires permission from various authorities. Therefore, there is a gap in this domain which deals with patient privacy, as well as developing an efficient approach to detect Alzheimer's disease without needing data collection from hospitals. In our research, an approach has been proposed to detect Alzheimer's disease which does not require medical data to be collected. Rather, deep learning models may be trained in the respective hospitals/institutions. This mitigates the gap and provides higher performance than the conventional approach. Nevertheless, it is also essential to develop a robust model in this domain. The robustness of the architectures was not proved in the above-mentioned research. In our work, we put an effort to prove the robustness of our approach by performing several tests.

3. Methodology

In this paper, our goal is to find out the efficacy of the federated approach in Alzheimer's disease detection. To achieve this, we collect two datasets (ADNI and OASIS), which are publicly available. MRI images may be explored in different orientations, as they are in 3D form. Hence, we explore to find an appropriate orientation from which our state-of-the-art transfer learning architectures may learn the patterns and produce significant accuracy. In what follows, we evaluate three different orientations of MRI images of the OASIS dataset (with augmentation and without augmentation). We train and test them with six different transfer learning architectures, from which we selected the optimal orientation and model. Figure 1 illustrates the overall workflow diagram of this study.

After the orientation and model selection procedure, a federated approach is implemented by performing multithreading. Another dataset, ADNI, is used in this part to verify the robustness of the federated approach. Moreover, the performance of this approach is compared with the conventional approach. We describe the whole process in detail in this section.

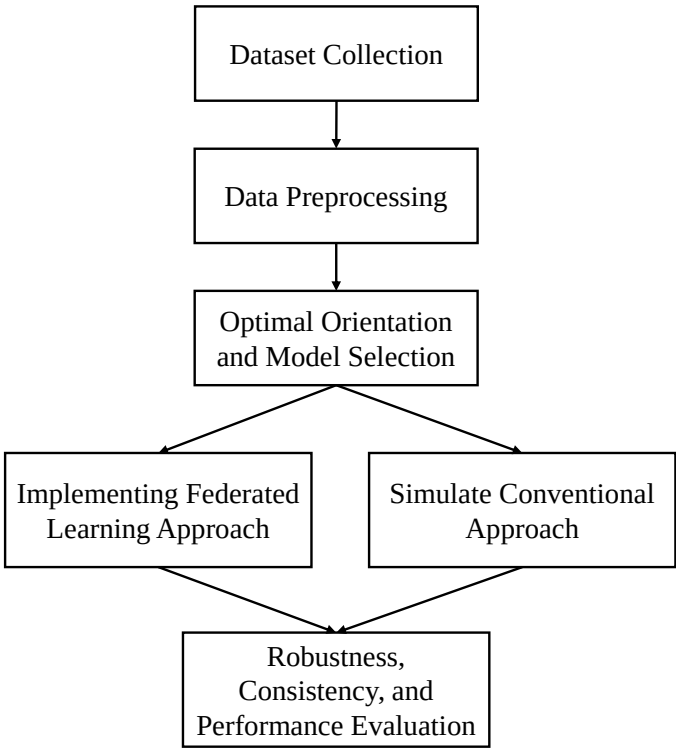

Figure 1. Overall workflow diagram of our study.

3.1. Dataset

Alzheimer's disease can be detected in many ways, such as brain MRI, PET scan images, neurological evaluations, and so on. In this study, we use two different brain MRI images datasets: ADNI and OASIS. Furthermore, these datasets are divided into many subsets, including augmented sets, merged sets, etc. Figure 2 describes the overall structure of the datasets.

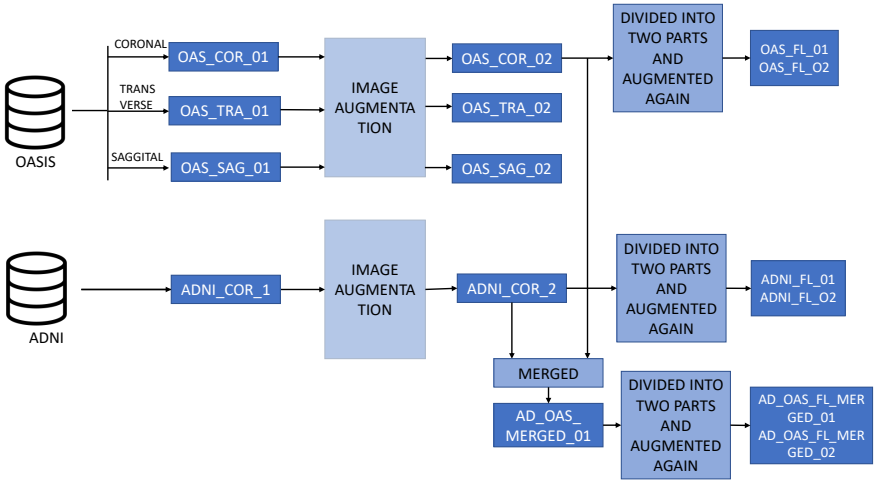

Figure 2. Descriptions of the datasets used in this study: Three different planes of the OASIS dataset are used here, which are coronal, sagittal, and transverse. Only the coronal plane of the ADNI dataset is used here. These images were further augmented and then merged as well. Then, these images were divided into two parts and constructed into datasets for the federated learning approach.

3.1.1. OASIS Dataset

The purpose of the Open Access Series of Imaging Studies (OASIS) [32] dataset is to make high-quality brain neuroimaging datasets available to the scientific community (available on https://www.oasis-brains.org/ (accessed on 2 February 2023)). By putting together neuroimaging datasets and making them available for free, they hope to speed up future scientific and clinical advances in neuroscience. For neurological, clinical, and cognitive investigations on ordinary aging and cognitive impairment, these data are widely available and assessed across a broad range of people, including neurological and genetic ranges. In this study, a total of 436 images are used, of which 98 image are AD-positive and 338 are normal MRI images, which are available in 3 orientations: coronal, sagittal, and transverse. These images are classified into four classes: Mild demented, Very mild demented, Moderate Demented, and Cognitive Normal. Because the number of images in each category was insufficient to make any of them statistically significant, we approach it as a binary classification problem. Therefore, the OASIS-1 dataset's Mild Demented, Very Mild Demented, and Moderately Demented classes are assessed as AD-positive (labeled as 1), and the Non-Demented class is determined to be AD-negative (labeled as 0). This helps the models learn more about the AD-positive images, as there are different classes of AD images that are now combined together.

3.1.2. ADNI Dataset

We collect the data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [33] database (adni.loni.usc.edu, accessed on 2 February 2023) for checking the performance and robustness of our proposed approach. The primary goal of the ADNI is to observe whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can all be used together to track the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). More information is available at www.adni-info.org.

The findings of studies conducted by the ADNI may be of significant use in clinical research pertaining to the diagnosis, prevention, and management of Alzheimer's disease (AD). It seeks biomarkers and enables precise diagnosis and tracking of AD by using its open-source datasets. The ADNI has proven to be very beneficial for long-term MRI and PET scans of elderly patients suffering from Alzheimer's disease, mild cognitive impairment, and other illnesses. We collected 321 Cognitive Normal (CN) MRI images and 136 Alzheimer's disease (AD) MRI images from the ADNI repository. Therefore, there is no need to convert it to a binary classification problem as it is already in this form. We consider the Cognitive Normal (CN) class as AD-negative (labeled as 0) and the Alzheimer's disease (AD) class as AD-positive (labeled as 1). So, a total of 457 images are collected from this dataset.

3.1.3. Augmentation Procedure

Data augmentation is used to increase the number of images. Deep-learning-based models perform well on a huge amount of data. If the amount of data is very small, then the model might be biased and may not be able to find patterns from the features of the images properly. However, collecting medical images is often a very difficult task, as most hospitals do not want to share their data. To overcome these challenges, data augmentation might be useful. Thus, we use conventional data augmentation methods in this analysis for both the ADNI and OASIS datasets. Images may be flipped, rotated, and zoomed, and their contrast can be increased using these methods. For each dataset, we only increase the images used in the training set. The testing set, however, is left unchanged from the primary dataset. As DL-based models perform well with a huge amount of data, it helps our models achieve better performance.

3.1.4. Merging Procedure

To evaluate the performance of the models, we merge the ADNI and the OASIS datasets in the following way. In the training set, all the training images from ADNI and OASIS are merged. On the other hand, the testing set consists of the merged images of the ADNI and OASIS test sets.

3.1.5. Dataset Splitting

Table 1 describes the training and testing sets used in Conventional training procedures, while Table 2 provides the training sets used in the training of the proposed model in the federated way. Some of these datasets contain augmented images, and some of the training sets were built by combining both ADNI and OASIS datasets.

OAS_COR_01, OAS_SAG_01, OAS_TRA_01: These datasets contain a total of 436 images, all of which belong to the OASIS dataset. The training set contains a total of 352 images, and the testing set contains 84 images. In this dataset, the images are not merged or augmented. The highest number of images are from the normal class. Here, COR stands for coronal, TRA stands for transverse, and SAG denotes the sagittal plane of the MRI image.

OAS_COR_02, OAS_SAG_02, OAS_TRA_02: Each of these datasets contains 2500 images. These images are augmented from the OASIS dataset. The training set contains a total of 2416 images, and the testing set remains the same as the nonaugmented dataset, which has 84 images. Here, COR stands for coronal, TRA stands for transverse, and SAG denotes the sagittal plane of the MRI image.

ADNI_COR_01: The ADNI_COR_01 dataset contains a total of 457 images. All the images in this dataset were collected from the ADNI dataset. Here, the training set contains 385 images and the testing set contains 72 images. The majority of images represent the normal classes.

ADNI_COR_02: The ADNI_COR_02 dataset consists of 2756 images. These images are augmented from the ADNI dataset. After augmentation, the training set contains a total of 2648 images, of which 1453 represent the normal class and 1195 represent the AD class. On the other hand, the testing set contains 72 images.

AD_OAS_MERGED_01: This is a merged dataset combining the ADNI and OASIS datasets together. All the training and testing images of the ADNI and OASIS datasets are merged for training and testing purposes. Here, the training set contains 5064 images and the testing set contains 156 images.

ADNI_FL_01, ADNI_FL_02: These datasets contain 1962 and 1961 images in each local client training set. These images are divided from the augmented ADNI dataset, and the number of images is increased by performing augmentation again. In the test set, both local clients have 35 test images.

OAS_FL_01, OAS_FL_02: These datasets are derived from the augmented OASIS dataset. Each local client contains 2153 images in the training set and 32 images in the local testing set.

AD_OAS_FL_MERGED_01, AD_OAS_FL_MERGED_02: This dataset is found after merging the ADNI and OASIS datasets. After merging and splitting the dataset, the two local clients contain 4115 and 4114 images in the training set after performing another layer of augmentation. On the other hand, in the local test set, each dataset contains 67 images.

LOCAL_AD_OAS_FL_01: This dataset consists of two different datasets for two local clients. Local client 1 consists of the ADNI dataset in the training and testing set, and local client 2 consists of the OASIS dataset. There are a total of 3923 images in the first local client training set and 4306 images in the second local client training set.

Table 1. Training and testing sets used for training in the conventional approach.

Dataset Code	Dataset	Merged (Yes/No)	Augmented (Yes/No)	Train Image			Test Image		
				Total	1	0	Total	1	0
OAS_COR_01	OASIS	No	No	352	80	272	84	18	66
OASIS_TRA_01	OASIS	No	No	352	80	272	84	18	66
OASIS_SAG_01	OASIS	No	No	352	80	272	84	18	66
OAS_COR_02	OASIS	No	Yes	2416	1187	1229	84	18	66
OASIS_TRA_02	OASIS	No	Yes	2416	1187	1229	84	18	66
OASIS_SAG_02	OASIS	No	Yes	2416	1187	1229	84	18	66
ADNI_COR_01	ADNI	No	No	385	112	273	72	24	48
ADNI_COR_02	ADNI	No	Yes	2648	1195	1453	72	24	48
AD_OAS_MERGED_01	ADNI + OASIS	Yes	Yes	5064	2382	2682	156	42	114

Table 2. Training and testing sets used for training in the federated approach.

Dataset Code	Dataset	Merged (Yes/No)	Augmented (Yes/No)	Train Image Local Machine			Test Image Local Machine		
				Total	1	0	Total	1	0
ADNI_FL_01	ADNI	No	Yes	1962	948	1014	35	10	25
ADNI_FL_02	ADNI	No	Yes	1961	948	1013	35	10	25
OAS_FL_01	OASIS	No	Yes	2153	1035	1118	32	8	24
OAS_FL_02	OASIS	No	Yes	2153	1035	1118	32	8	24
AD_OAS_FL_MERGED_01	ADNI + OASIS	Yes	Yes	4115	1983	2132	67	18	49
AD_OAS_FL_MERGED_02	ADNI + OASIS	Yes	Yes	4114	1983	2131	67	18	49

3.2. Orientation and Model Section

We used three different orientations (sagittal, coronal, and transverse) of MRI images and six different transfer learning architectures (DenseNet121 [34], DenseNet201, InceptionResNetV2 [35], MobileNet, MobileNetV2, and ResNet50V2). We use both the nonaugmented OASIS_COR_01, OASIS_SAG_01, and OASIS_TRA_01) and augmented (OASIS_COR_02, OASIS_SAG_02, and OASIS_TRA_02) datasets for finding the optimal model and orientation. First, we train all of our models and then evaluate the performances of all the models using different parameters, which are accuracy, precision, sensitivity, specificity, etc. The detailed description of these metrics is described in Section 4.1. In almost all of the cases, the selected models perform better in the coronal plane. Again, among these transfer learning models, MobileNet is found to be better performing in most cases. A detailed performance analysis is provided in Section 4.2. Based on the findings of this analysis, we selected MobileNet as the optimal model and coronal as an optimal plane. Therefore, MobileNet and the coronal plane were used for further analysis in this study. Figure 3 describes the procedure for selecting the best orientation and transfer learning model.

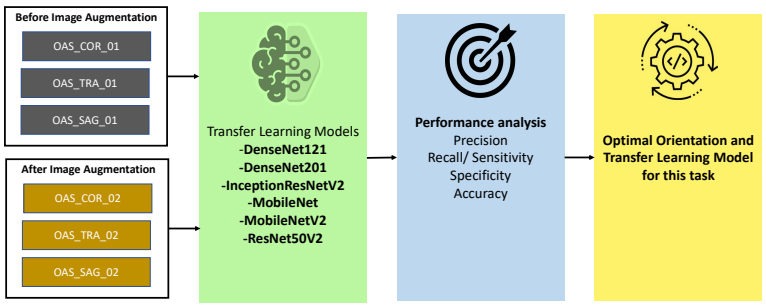

Figure 3. Pipeline for finding the best orientation and model for our study. Both augmented and nonaugmented datasets are used in this procedure. Different performance evaluation matrices are considered for analyzing the results of the transfer learning models.

3.3. Proposed Federated Learning Framework

In our study, we propose to train deep learning models in a federated way. Therefore, MobileNet is trained in a federated way to evaluate performance. The MobileNet architecture is very fast and efficient for image-processing-related tasks. We used the MobileNet's initial weights gained from training on the ImageNet database. We only changed the last layer (dense layer) of the MobileNet architecture. This dense layer has two neurons and a sigmoid as an activation function. These two neurons indicate AD or normal images.

In this proposed FL-based model, two local clients are involved in its development. These two clients have different datasets for training and validation purposes. These datasets can be varied, or two local clients can use the same dataset. Both clients use the same MobileNet model. While training the models, these two local clients run in parallel, and they send their model's weights for each epoch to the central server. Before transferring the weights to the global server, the local clients validate their model using the local validation dataset that evaluates the local client's performance. In the central server, an average of both of the local client's weights are taken, and it is considered as the weight of the global MobileNet model. Then, the global MobileNet containing this weight is used to test the global test set, which provides the global validation performance. After that, the global MobileNet weights are sent to each of the local clients for validation using their validation set. All these actions are performed for 50 epochs. After that, the found global MobileNet model is evaluated using different testing sets. In the global server, only weights of the locally trained model are available, but no local training data are available there. This ensures data security as well as confidentiality. Here, local clients do not need to share their private data, as they just transfer the weights of the model. As a result, the possibility of data breaching is reduced. Figure 4 depicts the federated learning procedure.

For example, when the MobileNet is trained in a federated way using the OASIS dataset, two local machines are trained with OAS_FL_01 and OAS_FL_02 training sets, and their corresponding test set is used as a local test. After each epoch, the weights are adjusted. The final model is tested using a global test set, which is the same as the conventional approach.

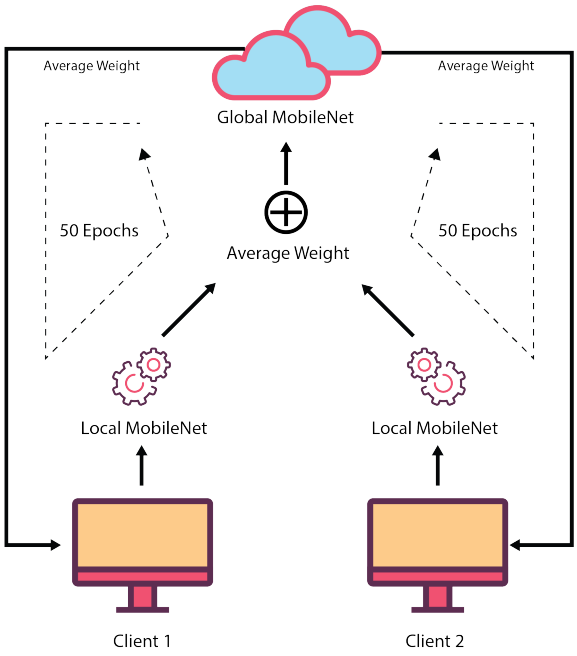

Figure 4. Federated learning procedure where two datasets are fed into two local machines: Weights are adjusted after each epoch. After 50 epochs, the final global model is found, which is tested with global test sets.

3.4. Evaluating the Robustness of the Current Approach

We conducted a number of tests using a variety of datasets in order to assess the robustness of our model as well as the conventional approach. The models are trained using the OASIS dataset, and then they are tested using the ADNI dataset, and vice versa. After that, we combined the datasets that have been trained, and after that, we test them using the OASIS and ADNI datasets. Additionally, the ADNI- and OASIS-trained models are utilized in an effort to validate this combined dataset. Further testing was performed on the merged dataset using the merged testing dataset. The experimental results help to evaluate the robustness of the models, along with the efficacy of the federated approach compared with the conventional approach. Figure 5 shows the procedures for checking the robustness and performance of the models trained in the conventional and federated approaches.

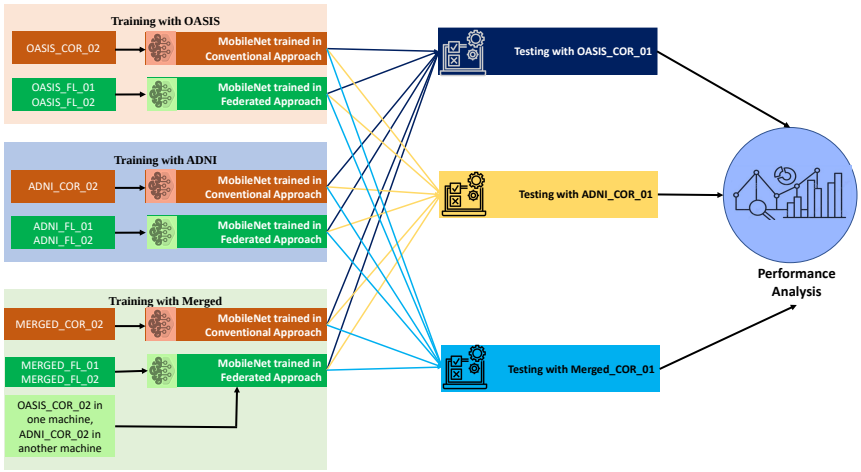

Figure 5. The overall procedures for testing the robustness of the models: Three different types of training procedures with different datasets are implemented here. MobileNet with both federated and conventional learning procedures are tested here. Moreover, a comparative performance analysis is conducted between these two approaches.

3.5. Hyperparameters and Implementation Setup

All the models are trained with loss–binary cross-entropy, learning rate—0.001, and Adam Optimizer. These hyperparameters are selected after performing an ablation study. These models are trained on an NVidia K80 GPUs, Intel(R) Xeon(R) CPU @ 2.30 GHz, and 12 GB of RAM.

4. Result Analysis

In this paper, our aim is to analyze the efficiency of the federated approach in detecting Alzheimer's disease. For this, we need to find out an optimal model and orientation of MRI images at first. Therefore, we picked the OASIS dataset and trained the existing state-of-the-art transfer learning approaches, such as MobileNet, MobileNeV2, ResNet50V2, InceptionResNetV2, DenseNet121, and DenseNet201. To find out the best orientation of MRI images, we explored three different orientations that were already prepared in the OASIS-1 dataset (coronal, sagittal, and transverse). As the number of images in the training images were not large enough, we enhanced the training set by performing augmentation, which largely increased the performance.

After selecting the optimal model and orientation, we evaluated the robustness of the conventional training approach by using two different datasets (OASIS and ADNI). For this, we trained the optimal model with OASIS in both conventional and federated approaches and tested with the ADNI, OASIS, and merged test sets. We also evaluated the performance of the effective deep learning architectures by training with OASIS and the merged training sets, as well as testing with all three test sets. In this section, all the performance-related issues are described and analyzed in detail.

4.1. Performance Evaluation Metrics

After a model is constructed, it is crucial to evaluate its performance. This section focuses on the most popular evaluation metrics seen in research papers. First, we check the proportions of true positive, true negative, false positive, and false negative values. After that, we take a look at a variety of advanced indicators for assessment.

True Positive Values (TP): The TP rate in AD detection is the percentage of times an AD detection algorithm correctly identifies an image having AD.

True Negative Values (TN): The TN rate in AD detection is the percentage of times an AD detection algorithm correctly identifies an image not having AD.

False Positive Values (FP): The FP measures the number of times an image is incorrectly labeled AD by the model.

False Negative Values (FN): The FN measures the number of times an image is incorrectly labeled normal by the model.

Sensitivity (Sn): Sn measures how well an architecture can make predictions for positive samples.

$$Sn = Recall = \frac{TP}{TP + FN} \tag{1}$$

Specificity (Sp): Sp measures how well an architecture can make predictions for negative samples.

$$Sp = \frac{TN}{FP + TN} \tag{2}$$

Precision (Pn): Pn is the degree to which a measurement or result is accurate or precise.

$$Pn = \frac{TP}{TP + FP} \tag{3}$$

F1 score (F1): $F1$ is calculated by taking the harmonic average of the accuracy and recall scores, with lower values being given greater consideration in the calculation.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

Accuracy: It can be defined as the percentage of the total number of samples that corresponds to the correct prediction of the number of samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

4.2. Selecting the Optimal Model and Orientation

First, we trained the MobileNet, MobileNeV2, ResNet50V2, InceptionResNetV2, DenseNet121, and DenseNet201 models with the OASIS dataset, which included 352 images for training and 84 images for testing (OAS_COR_01, OASIS_TRA_01, OASIS_SAG_01). Next, we tested the accuracy of the models using the OASIS dataset. These images were available in all three orientations (coronal, sagittal, and transverse). Table 3 presents an in-depth comparison of the three different transfer learning architectures' performances in each of the three different orientations. When it is evaluated using the coronal plane, MobileNet provided the highest results in terms of sensitivity (92%), F1 score (87%), and accuracy (90%), respectively. Other models reached an accuracy in the range of 86% to 89%, while having a significantly lower sensitivity. Note that sensitivity is extremely important in the identification of AD. The DenseNet201 model exhibited the worst performance, as detailed in Table 3. In the case of the transverse plane, the accuracy achieved by ResNet50V2 is 89%, which is the maximum possible score. The average performance of the coronal plane is significantly higher than that of any of the other models, in the range of 81–85%. The performance of MobileNet in the sagittal plane (75%) is significantly lower than the performance of MobileNet evaluated on the coronal plane (90%). MobileNet's highest

accuracy in the sagittal plane is 86%. It is therefore evident that all the models performed well in the coronal plane, with MobileNet providing the best performance.

Table 3. Comparison of performance of transfer learning architectures (conventional training approach) in the coronal, transverse, and sagittal plane before image augmentation.

CORONAL PLANE					
Model	**Precision (%)**	**Sensitivity (%)**	**F1 Score (%)**	**Specificity (%)**	**Accuracy (%)**
DenseNet121	81.00%	73.00	76.00	95.00	86.00
DenseNet201	74.00	60.00	61.00	96.00	81.00
InceptionResNetV2	82.00	82.00	82.00	92.00	88.00
MobileNet	85.00	92.00	87.00	89.00	90.00
MobileNetV2	87.00	79.00	82.00	96.00	89.00
ResNet50V2	88.00	71.00	76.00	98.00	87.00
TRANSVERSE PLANE					
Model	**Precision (%)**	**Sensitivity (%)**	**F1 Score (%)**	**Specificity (%)**	**Accuracy (%)**
DenseNet121	77.00	62.00	65.00	96.00	82.00
DenseNet201	72.00	70.00	71.00	89.00	81.00
InceptionResNetV2	76.00	71.00	73.00	92.00	83.00
MobileNet	77.00	78.00	77.00	89.00	85.00
MobileNetV2	72.00	64.00	66.00	93.00	81.00
ResNet50V2	84.00	85.00	84.00	92.00	89.00
SAGITTAL PLANE					
	Precision (%)	**Sensitivity (%)**	**F1 Score (%)**	**Specificity (%)**	**Accuracy (%)**
DenseNet121	78.00	72.00	74.00	93.00	85.00
DenseNet201	77.00	76.00	77.00	90.00	85.00
InceptionResNetV2	76.00	69.00	72.00	93.00	83.00
MobileNet	80.00	75.00	77.00	93.00	86.00
MobileNetV2	74.00	60.00	61.00	96.00	81.00
ResNet50V2	78.00	67.00	70.00	95.00	83.00

Since the number of images in the training set was insufficient to train these deep learning architectures, we only augmented the training set (with a total of 2416 images, 1187 of which were labeled as AD MRI and 1229 of which were not AD MRI). The testing set consists of the same number of images as before, including 18 that were labeled as AD, 66 that were tagged as non-AD (for a total of 84), OAS_COR_02, OASIS_TRA_02, and OASIS_SAG_02, as specified in Section 3.1. Table 4 gives information regarding the performance of various transfer learning architectures after image augmentation. In the case of the coronal plane, augmentation increased MobileNet's accuracy from 90% to 92%. Previously, it was at 90%. After the augmentation, the accuracy of ResNet50V2 raised to 92%. In the case of the transverse plane, the accuracy achieved by MobileNet after augmentation is at its maximum (86%). In the sagittal plane, MobileNet accomplished the highest level of accuracy (85%). MobileNet is the model that performed the best out of all of the models in this case, and the performance of all of the models is better in the coronal plane than in the transverse and sagittal planes.

MobileNet performed better in most of the cases, whereas models performed best in the coronal plane when compared with the performances of the transfer learning algorithms before and after augmentation. Moreover, after data augmentation, MobileNet achieved a better performance score than the nonaugmented dataset. As a result, we consider MobileNet as our best model and the coronal plane as our preferred orientation in this work.

Table 4. Comparison of performance of transfer learning architectures (conventional training approach) in the coronal, transverse, and sagittal plane after image augmentation.

CORONAL PLANE					
Model	**Precision (%)**	**Sensitivity (%)**	**F1 Score (%)**	**Specificity (%)**	**Accuracy (%)**
DenseNet121	74.00	85.00	73.00	70.00	76.00
DenseNet201	80.00	82.00	81.00	91.00	87.00
InceptionResNetV2	70.00	80.00	66.00	59.00	68.00
MobileNet	92.00	83.00	86.00	98.00	92.00
MobileNetV2	80.00	75.00	77.00	93.00	86.00
ResNet50V2	88.00	87.00	87.00	95.00	92.00
TRANSVERSE PLANE					
Model	**Precision (%)**	**Sensitivity (%)**	**F1 Score (%)**	**Specificity (%)**	**Accuracy (%)**
DenseNet121	11.00	50.00	18.00	0.00	21.00
DenseNet201	72.00	79.00	74.00	80.00	80.00
InceptionResNetV2	61.00	52.00	21.00	3.00	24.00
MobileNet	87.00	69.00	73.00	98.00	86.00
MobileNetV2	75.00	75.00	75.00	89.00	83.00
ResNet50V2	74.00	74.00	74.00	87.00	82.00
SAGITTAL PLANE					
Model	**Precision (%)**	**Sensitivity (%)**	**F1 Score (%)**	**Specificity (%)**	**Accuracy (%)**
DenseNet121	77.00	78.00	77.00	89.00	85.00
DenseNet201	81.00	78.00	80.00	83.00	85.00
InceptionResNetV2	77.00	87.00	80.00	80.00	83.00
MobileNet	77.00	82.00	79.00	86.00	85.00
MobileNetV2	80.00	65.00	68.00	96.00	83.00
ResNet50V2	76.00	71.00	73.00	92.00	83.00

4.3. Training with OASIS and Testing with ADNI, OASIS, and Merged Dataset

From Section 4.2, it is evident that MobileNet is the go-to model, and the coronal plane is the selected orientation of MRI images. We evaluated the performance of the federated learning and conventional approach for different types of experiments. For example, we trained our model in the ADNI dataset and tested our model in the OASIS dataset, and vice versa. We also trained our model using the merged dataset and tested it with the ADNI and OASIS datasets. In every case, we implemented these experiments in both the federated and conventional approaches and then compared their results in terms of the different performance evaluation metrics mentioned in Section 4.1.

4.3.1. Conventional Approach (Training with OASIS)

To check the robustness of the MobileNet in Alzheimer's disease detection, at first, we trained the MobileNet model with an augmented OASIS training set (OAS_COR_02_train) and tested it with the ADNI (ADNI_COR_01_test), OASIS (OAS_COR_01_test), and merged test set (AD_OAS_MERGED_01_test) in the conventional approach. Table 5 provides the performance of the conventional and federated approaches when the models are trained with the OASIS dataset. Here, in this approach, MobileNet achieved 91.67% accuracy and 66.67% sensitivity in the OASIS test set, 65.28% accuracy and 12.50% sensitivity in the ADNI test set, and 76.92% accuracy and 28.50% sensitivity in the merged test set. In all the cases, there is a huge difference between sensitivity and specificity, which shows that models were not well-generalized after training in the conventional approach. Table 5 depicts the performance between the conventional approach and the federated approach.

4.3.2. Federated Approach (Training with OASIS)

To evaluate the performance of the model trained in the federated approach, we trained two MobileNet models with the OAS_FL_01_Local1 and OAS_FL_01_Local2 training sets in

local machines. We finally obtained the trained MobileNet model with optimal weight. This model was tested using the ADNI (ADNI_COR_01_test), OASIS (OAS_COR_01_test), and merged test set (AD_OAS_MERGED_01_test), which is the same as for the conventional approach. MobileNet acquired 92.86% accuracy and 94.44% sensitivity when tested with the OASIS test set, 77.78% accuracy and 75.00% sensitivity in the ADNI test set, and 83.33% accuracy and 69.05% sensitivity in the merged test set. In each case, the federated approach acquired better accuracy and sensitivity and showed better generalization capability. In this approach, the model is more capable of distinguishing between two classes than the conventional training procedure, as is evident in Table 5.

Table 5. Performance comparison between the conventional and federated approaches when trained with the OASIS dataset.

Test Set	Approach	Precision (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
OASIS	Conventional	92.31	66.67	98.48	91.67
	Federated	77.27	94.44	92.42	92.86
ADNI	Conventional	42.86	12.50	91.67	65.28
	Federated	64.29	75.00	79.17	77.78
Merged	Conventional	66.67	28.57	94.74	76.92
	Federated	69.05	69.05	88.60	83.33

4.4. Training with ADNI and Testing with ADNI, OASIS, and Merged Dataset

4.4.1. Conventional Approach (Training with ADNI)

Here, MobileNet was trained with the ADNI_COR_02_training set and tested with the ADNI (ADNI_COR_01_test), OASIS (OAS_COR_01_test), and merged test sets (AD_OAS_MERGED_01_test). Table 6 provides the performance of the conventional and federated approaches when the models were trained with the ADNI dataset. MobileNet achieved 78.57% accuracy and 55.56% sensitivity in the OASIS test set, 75% accuracy and 75% sensitivity in the ADNI test set, and 75% accuracy and 66.67% sensitivity in the merged test set. Table 5 compares the performance between the conventional approach and the federated approach.

4.4.2. Federated Approach (Training with ADNI)

To evaluate the performance of the federated approach, we trained two local MobileNet models with ADNI_FL_01_Local1 and ADNI_FL_01_Local2 training sets, where we used the ADNI (ADNI_COR_01_test), OASIS (OAS_COR_01_test), and merged test set (AD_OAS_MERGED_01_test) to test the performance of the global MobileNet model. MobileNet acquired 78.57% accuracy while testing with the OASIS test set, 81.94% in the ADNI test set, and 77.56% accuracy in the merged test set. In the case of sensitivity, the performance was not up to the mark. However, the federated approach performed better than the conventional approach in the case of accuracy (see Table 6).

Table 6. Performance comparison between the conventional and federated approaches when trained with the ADNI dataset.

Test Set	Approach	Precision (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
OASIS	Conventional	43.48	55.56	80.30	75.00
	Federated	50.00	44.44	87.88	78.57
ADNI	Conventional	60.00	75.00	75.00	75.00
	Federated	78.95	62.50	91.67	81.94
Merged	Conventional	52.83	66.67	78.07	75.00
	Federated	65.22	35.71	92.98	77.56

4.5. Training with Merged and Testing with ADNI, OASIS, and Merged Dataset

4.5.1. Conventional Approach (Training with Merged)

Here, the MobileNet was trained with the AD_OAS_MERGED_01_train set and tested with OASIS (OAS_COR_01_test), ADNI (ADNI_COR_01_test), and merged (AD_OAS_MERGED_01_test) test sets. In this case, the MobileNet acquired 85.71% accuracy and 55.76% sensitivity in the OASIS test set, 70.83% accuracy and 62.50% sensitivity in the ADNI test set, and 78.85% accuracy and 64.29% sensitivity using the merged test set. Table 7 shows the performance of the conventional approach when trained using the merged set.

4.5.2. Federated Approach (Training with Merged)

Here, the merged training set is divided into two portions (AD_OAS_FL_MERGED_01_Local1, AD_OAS_FL_MERGED_01_Local2) to train the two local MobileNet models. The derived global MobileNet model was tested with the OASIS (OAS_COR_01_test), ADNI (ADNI_COR_01_test), and merged (AD_OAS_MERGED_01_test) test sets. In the case of OASIS, it achieved the highest (95.24%) accuracy and (83.33%) sensitivity, 81.94% accuracy and 62.50% sensitivity, and 83.97% accuracy and 61.90% sensitivity. Almost in all the cases, the proposed federated approach performed better, as shown in Table 7.

4.5.3. Federated Approach (Training with ADNI and OASIS in Two Different Machines)

Here, two MobileNet models were trained with the ADNI (ADNI_COR_02_train) and OASIS (OAS_COR_02_train) sets separately. The found global model was tested with the same test sets mentioned in the previous sections. This model acquired 91.67%, 77.78%, and 82.69% accuracy in the OASIS, ADNI, and merged test set, respectively. These results are lower than those of the federated approach trained with the merged set but greater than the conventional approach. However, in this approach, the model was found to be more consistent. We found this model to have better capability of distinguishing between two classes compared with the other two approaches. This is because the gap between sensitivity and specificity is lower than in the other two approaches, as depicted in Table 7.

Table 7. Performance comparison between the conventional and federated approaches when trained with the merged dataset.

Test Set	Approach	Precision (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
OASIS	Conventional	71.43	55.56	93.94	85.71
	Federated	93.75	83.33	98.48	95.24
	Federated with two different training sets in two local machines	73.91	94.44	90.91	91.67
ADNI	Conventional	55.56	62.50	75.00	70.83
	Federated	78.95	62.50	91.67	81.94
	Federated with two different training set in two local machines	66.67	66.67	83.33	77.78
Merged	Conventional	60.00	64.29	84.21	78.85
	Federated	74.29	61.90	92.11	83.97
	Federated with two different training sets in two local machines	67.44	69.05	87.72	82.69

4.6. Discussion

The proposed MobileNet acquired more than 95%, 78%, and 83% accuracy in the OASIS, ADNI, and merged tests, respectively. Figure 6 shows that the MobileNet model acquired the highest accuracy when trained with the merged dataset in the federated approach. From this figure, it is also clear that the MobileNet model performed better than the conventional approach in all the parameters when trained with the ADNI, OASIS, or merged training sets.

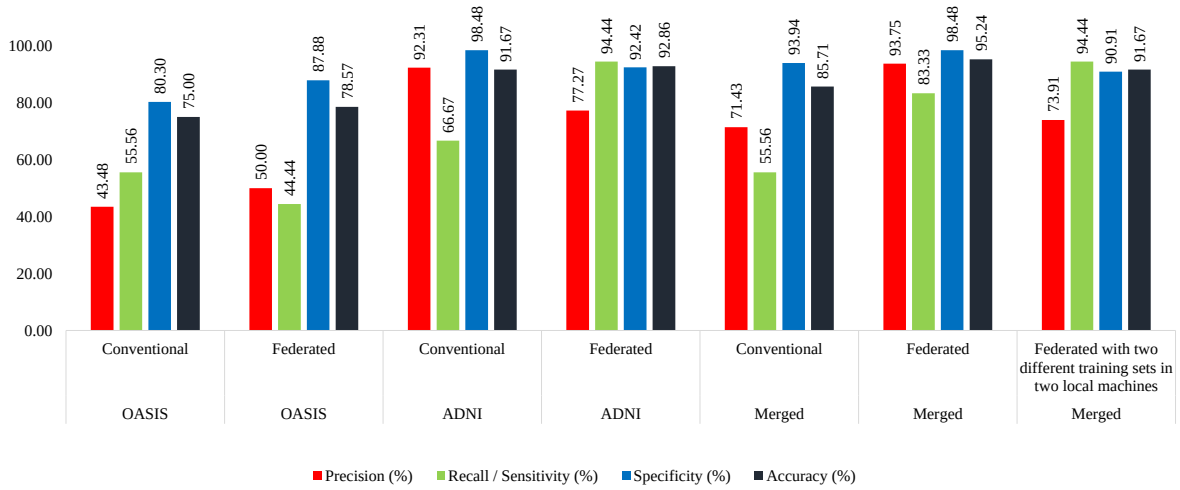

Figure 6. Performance analysis of the conventional and federated approaches when tested using the OASIS dataset.

In the case of testing with the ADNI, the MobileNet shows a similar pattern as the previous one. Here, the model was able to distinguish better between the classes in the case of training in the federated approach. When trained with OASIS, the federated approach achieved 77.78% accuracy, whereas the model could acquire only 65.28% accuracy in the conventional approach. Trained with the ADNI and merged sets, the models were seen to perform better in the case of those trained with the federated approach (Figure 7).

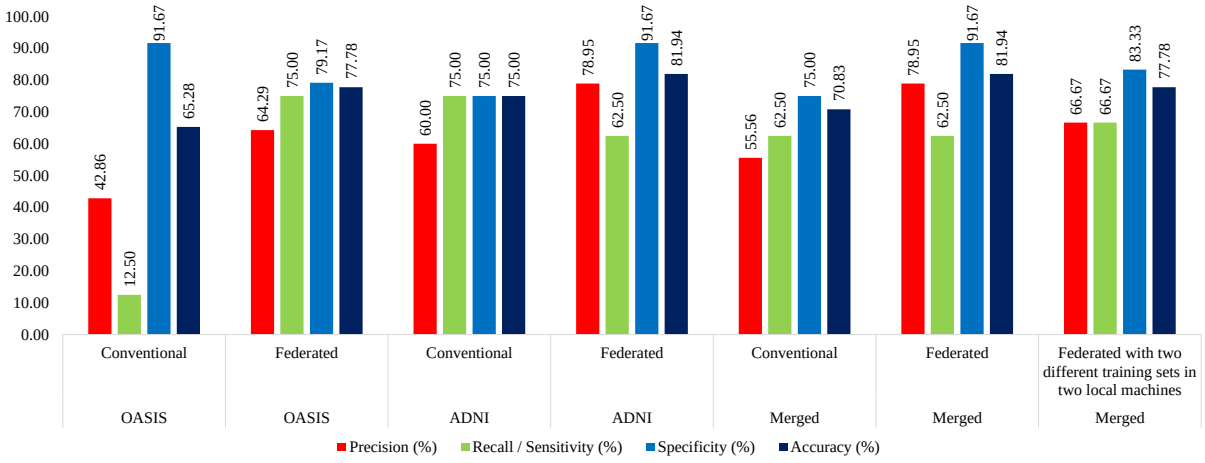

Figure 7. Performance analysis of the conventional and federated approaches when tested using the ADNI dataset.

While testing with the merged test set, the model performs better when trained in the federated approach in all the cases. If the model is trained with the OASIS dataset, it achieves 83.33% accuracy in the federated approach, while 76.92% in the conventional approach. In the case of training with the merged training set, the model achieved 83.98% accuracy in the federated approach, while only 78.85% in the conventional approach (Figure 8).

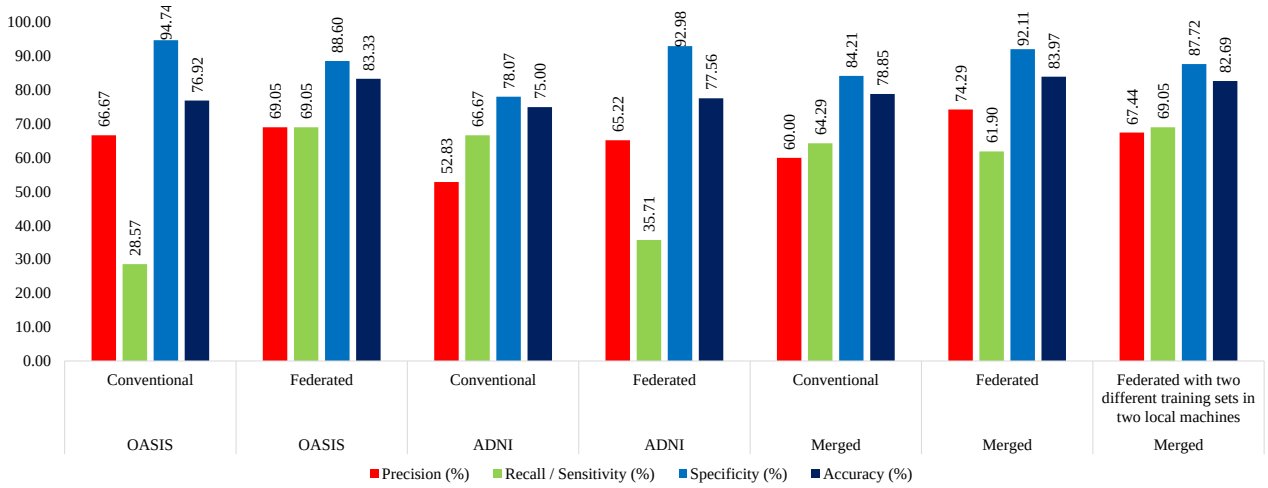

Figure 8. Performance analysis of the conventional and federated approaches while testing using the merged dataset.

Therefore, it is evident that models perform better if trained with the federated approach. We also found the models to be more robust than the conventional approach, since performance deviation was not significant enough in the case of a change in the training set. Nevertheless, the models acquired more consistent results in terms of sensitivity and specificity. This determines the model's capability in distinguishing between two classes. Hence, we found the federated approach to be better performing and more robust and consistent than the conventional approach.

5. Conclusions

Many researchers from all around the world are working to identify the most effective cure for AD. In this study, we pensively looked into the potential applications of federated learning for the detection of AD cases. In addition, we used traditional DL-based methods to find this disease. Out of three distinct planes—coronal, sagittal, and transverse, we chose coronal to be the optimal orientation plane for MRI brain images. The MobileNet architecture was discovered to be the optimal model, providing the best performance in all the cases. To avoid model bias, we used data augmentation strategies. This aids the model's understanding of the characteristics of AD MRI images. While training with the OASIS dataset, the conventional approach achieved an accuracy of 91.67%, whereas the FL approach achieved 92.86% accuracy. For the ADNI and merged datasets, the FL approach achieved 77.78% and 83.33%, respectively. The conventional method achieved 65.28% and 76.92% accuracy in this case, which is less than the FL method. Nevertheless, we also evaluated the robustness of the model by testing it on the OASIS, ADNI, and merged datasets. In most cases, the FL approach outperformed the conventional approach in terms of different evaluation parameters. Here, FL achieved 95.24%, 81.94%, and 83.97% accuracy on the OASIS, ADNI, and merged datasets, respectively. Our study reveals that the FL method outperforms all the conventional methods and ensures the privacy of the data, since it does not require the collection of any data. In building an efficient and robust deep learning architecture in the medical domain, especially in Alzheimer's disease detection, collecting medical images is the main barrier. By adopting the proposed federated approach, this barrier is reduced. We strongly believe that this will encourage more hospitals to allow researchers to train deep learning models at their hospital due to reduced privacy concerns. Furthermore, models will be trained on more diverse data, leading to the enhancement of deep learning models' robustness and performance. For further improvement of the model performance, extensive data preprocessing techniques can be exercised. In what follows, the robustness of the model might be improved by integrating more diverse datasets containing a greater number of samples, especially positive samples, into the

model. This may also enhance the efficiency of the approach. Formal verification of the deep learning models [36,37] can be another dimension that may add value in this domain. It may enhance trust in AI approaches in medical practitioners, as the system's correctness can be verified in this way. The explainability of the proposed model might be another important area of study.

Author Contributions: Conceptualization , T.G. and M.I.A.P.; formal analysis, T.G. and M.I.A.P.; funding acquisition, M.A.Y., M.A.H., M.M.M. and M.O.A.; investigation, T.G. and M.I.A.P.; methodology, T.G. and M.I.A.P.; supervision, M.A.Y.; visualization, T.G. and M.I.A.P.; writing—original draft, T.G. and M.I.A.P.; writing—review and editing, M.A.Y., M.A.H., M.M.M. and M.O.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Institutional Fund Projects under grant no. (IFPIP: 544-611-1443). The authors gratefully acknowledge technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Data Availability Statement: Publicly available datasets were used in this research. They are available at https://www.oasis-brains.org/, and https://adni.loni.usc.edu/. Data was shared by:

- **OASIS Cross-Sectional:** Principal Investigators: D. Marcus, R, Buckner, J, Csernansky J. Morris; P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 RR021382
- **ADNI:** Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuroimaging at the University of Southern California.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aging, N.I. Alzheimer's Disease Fact Sheet. 2023. Available online: https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet (accessed on 7 March 2023).
2. Kim, B.; Noh, G.O.; Kim, K. Behavioural and psychological symptoms of dementia in patients with Alzheimer's disease and family caregiver burden: A path analysis. *BMC Geriatr.* **2021**, *21*, 1–12. [CrossRef] [PubMed]
3. Vemuri, P.; Gunter, J.L.; Senjem, M.L.; Whitwell, J.L.; Kantarci, K.; Knopman, D.S.; Boeve, B.F.; Petersen, R.C.; Jack, C.R., Jr. Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies. *Neuroimage* **2008**, *39*, 1186–1197. [CrossRef]
4. O'Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, G.V.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep learning vs. traditional computer vision. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Las Vegas, CA, USA, 2–3 May 2019* ; Springer: Cham, Switzerland, 2020; Volume 1, pp. 128–144.
5. Hulsen, T. Sharing is caring—Data sharing initiatives in healthcare. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3046. [CrossRef] [PubMed]
6. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Artificial Intelligence and Statistics, PMLR, Ft. Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.

7. Ghosh, T.; Banna, M.H.A.; Nahian, M.J.A.; Kaiser, M.S.; Mahmud, M.; Li, S.; Pillay, N. A Privacy-Preserving Federated-MobileNet for Facial Expression Detection from Images. In Proceedings of the Applied Intelligence and Informatics: Second International Conference, AII 2022, Reggio Calabria, Italy, 1–3 September 2022; Springer: Cham, Switzerland, 2023; pp. 277–292.
8. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
9. Ahamed, K.U.; Islam, M.; Uddin, A.; Akhter, A.; Paul, B.K.; Yousuf, M.A.; Uddin, S.; Quinn, J.M.; Moni, M.A. A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images. *Comput. Biol. Med.* **2021**, *139*, 105014. [CrossRef] [PubMed]
10. Aurna, N.F.; Yousuf, M.A.; Taher, K.A.; Azad, A.; Moni, M.A. A classification of MRI brain tumor based on two stage feature level ensemble of deep CNN models. *Comput. Biol. Med.* **2022**, *146*, 105539. [CrossRef]
11. Hossain, M.M.; Hasan, M.M.; Rahim, M.A.; Rahman, M.M.; Yousuf, M.A.; Al-Ashhab, S.; Akhdar, H.F.; Alyami, S.A.; Azad, A.; Moni, M.A. Particle Swarm Optimized Fuzzy CNN With Quantitative Feature Fusion for Ultrasound Image Quality Identification. *IEEE J. Transl. Eng. Health Med.* **2022**, *10*, 1–12. [CrossRef]
12. Mim, T.R.; Amatullah, M.; Afreen, S.; Yousuf, M.A.; Uddin, S.; Alyami, S.A.; Hasan, K.F.; Moni, M.A. GRU-INC: An inception-attention based approach using GRU for human activity recognition. *Expert Syst. Appl.* **2023**, *216*, 119419. [CrossRef]
13. Alinsaif, S.; Lang, J.; Alzheimer's Disease Neuroimaging Initiative. 3D shearlet-based descriptors combined with deep features for the classification of Alzheimer's disease based on MRI data. *Comput. Biol. Med.* **2021**, *138*, 104879. [CrossRef]
14. Puente-Castro, A.; Fernandez-Blanco, E.; Pazos, A.; Munteanu, C.R. Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Comput. Biol. Med.* **2020**, *120*, 103764. [CrossRef]
15. Chui, K.T.; Gupta, B.B.; Alhalabi, W.; Alzahrani, F.S. An MRI scans-based Alzheimer's disease detection via convolutional neural network and transfer learning. *Diagnostics* **2022**, *12*, 1531. [CrossRef] [PubMed]
16. Folego, G.; Weiler, M.; Casseb, R.F.; Pires, R.; Rocha, A. Alzheimer's disease detection through whole-brain 3D-CNN MRI. *Front. Bioeng. Biotechnol.* **2020**, *8*, 534592. [CrossRef] [PubMed]
17. Liu, J.; Li, M.; Luo, Y.; Yang, S.; Li, W.; Bi, Y. Alzheimer's disease detection using depthwise separable convolutional neural networks. *Comput. Methods Programs Biomed.* **2021**, *203*, 106032. [CrossRef]
18. An, N.; Ding, H.; Yang, J.; Au, R.; Ang, T.F. Deep ensemble learning for Alzheimer's disease classification. *J. Biomed. Inform.* **2020**, *105*, 103411. [CrossRef]
19. Xia, Z.; Yue, G.; Xu, Y.; Feng, C.; Yang, M.; Wang, T.; Lei, B. A novel end-to-end hybrid network for Alzheimer's disease detection using 3D CNN and 3D CLSTM. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE, Iowa City, IA, USA, 3–7 April 2020; pp. 1–4.
20. Koh, J.E.W.; Jahmunah, V.; Pham, T.H.; Oh, S.L.; Ciaccio, E.J.; Acharya, U.R.; Yeong, C.H.; Fabell, M.K.M.; Rahmat, K.; Vijayananthan, A.; et al. Automated detection of Alzheimer's disease using bi-directional empirical model decomposition. *Pattern Recognit. Lett.* **2020**, *135*, 106–113. [CrossRef]
21. Kang, W.; Lin, L.; Zhang, B.; Shen, X.; Wu, S.; Alzheimer's Disease Neuroimaging Initiative. Multi-model and multi-slice ensemble learning architecture based on 2D convolutional neural networks for Alzheimer's disease diagnosis. *Comput. Biol. Med.* **2021**, *136*, 104678. [CrossRef]
22. Kaplan, E.; Dogan, S.; Tuncer, T.; Baygin, M.; Altunisik, E. Feed-forward LPQNet based automatic alzheimer's disease detection model. *Comput. Biol. Med.* **2021**, *137*, 104828. [CrossRef]
23. Bringas, S.; Salomón, S.; Duque, R.; Lage, C.; Montaña, J.L. Alzheimer's disease stage identification using deep learning models. *J. Biomed. Inform.* **2020**, *109*, 103514. [CrossRef]
24. Murugan, S.; Venkatesan, C.; Sumithra, M.; Gao, X.Z.; Elakkiya, B.; Akila, M.; Manoharan, S. DEMNET: A deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images. *IEEE Access* **2021**, *9*, 90319–90329. [CrossRef]
25. Lodha, P.; Talele, A.; Degaonkar, K. Diagnosis of alzheimer's disease using machine learning. In Proceedings of the 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), IEEE, Pune, India, 16–18 August 2018; pp. 1–4.
26. Li, F.; Liu, M.; Alzheimer's Disease Neuroimaging Initiative. A hybrid convolutional and recurrent neural network for hippocampus analysis in Alzheimer's disease. *J. Neurosci. Methods* **2019**, *323*, 108–118. [CrossRef]
27. Basheera, S.; Ram, M.S.S. A novel CNN based Alzheimer's disease classification using hybrid enhanced ICA segmented gray matter of MRI. *Comput. Med. Imaging Graph.* **2020**, *81*, 101713. [CrossRef] [PubMed]
28. Bi, X.; Li, S.; Xiao, B.; Li, Y.; Wang, G.; Ma, X. Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology. *Neurocomputing* **2020**, *392*, 296–304. [CrossRef]
29. Jabason, E.; Ahmad, M.O.; Swamy, M. Classification of Alzheimer's disease from MRI data using an ensemble of hybrid deep convolutional neural networks. In Proceedings of the 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, Dallas, TX, USA, 4–7 August 2019; pp. 481–484.
30. Helaly, H.A.; Badawy, M.; Haikal, A.Y. Deep learning approach for early detection of Alzheimer's disease. *Cogn. Comput.* **2022**, *14*, 1711–1727. [CrossRef] [PubMed]
31. Venugopalan, J.; Tong, L.; Hassanzadeh, H.R.; Wang, M.D. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci. Rep.* **2021**, *11*, 3254. [CrossRef]

32. Marcus, D.S.; Wang, T.H.; Parker, J.; Csernansky, J.G.; Morris, J.C.; Buckner, R.L. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.* **2007**, *19*, 1498–1507. [CrossRef]
33. Mueller, S.G.; Weiner, M.W.; Thal, L.J.; Petersen, R.C.; Jack, C.; Jagust, W.; Trojanowski, J.Q.; Toga, A.W.; Beckett, L. The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin.* **2005**, *15*, 869–877. [CrossRef]
34. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
35. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
36. Krichen, M.; Mihoub, A.; Alzahrani, M.Y.; Adoni, W.Y.H.; Nahhal, T. Are Formal Methods Applicable To Machine Learning And Artificial Intelligence? In Proceedings of the 2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 9–11 May 2022; pp. 48–53. [CrossRef]
37. Raman, R.; Gupta, N.; Jeppu, Y. Framework for Formal Verification of Machine Learning Based Complex System-of-Systems. *Insight* **2023**, *26*, 91–102. [CrossRef]