Article

An Intelligent System for Early Recognition of Alzheimer's Disease Using Neuroimaging

Modupe Odusami [1], Rytis Maskeliūnas [1] and Robertas Damaševičius [2,*]

1 Department of Multimedia Engineering, Kaunas University of Technology, 51368 Kaunas, Lithuania; modupe.odusami@ktu.edu (M.O.); rytis.maskeliunas@ktu.lt (R.M.)
2 Department of Software Engineering, Kaunas University of Technology, 51368 Kaunas, Lithuania
* Correspondence: robertas.damasevicius@ktu.lt

Abstract: Alzheimer's disease (AD) is a neurodegenerative disease that affects brain cells, and mild cognitive impairment (MCI) has been defined as the early phase that describes the onset of AD. Early detection of MCI can be used to save patient brain cells from further damage and direct additional medical treatment to prevent its progression. Lately, the use of deep learning for the early identification of AD has generated a lot of interest. However, one of the limitations of such algorithms is their inability to identify changes in the functional connectivity in the functional brain network of patients with MCI. In this paper, we attempt to elucidate this issue with randomized concatenated deep features obtained from two pre-trained models, which simultaneously learn deep features from brain functional networks from magnetic resonance imaging (MRI) images. We experimented with ResNet18 and DenseNet201 to perform the task of AD multiclass classification. A gradient class activation map was used to mark the discriminating region of the image for the proposed model prediction. Accuracy, precision, and recall were used to assess the performance of the proposed system. The experimental analysis showed that the proposed model was able to achieve 98.86% accuracy, 98.94% precision, and 98.89% recall in multiclass classification. The findings indicate that advanced deep learning with MRI images can be used to classify and predict neurodegenerative brain diseases such as AD.

Keywords: intelligent systems; image processing; expert systems; Alzheimer's disease; MRI; deep learning; explainability

1. Introduction

Alzheimer's disease (AD) is a brain disease that has become more prevalent over time and it is now the fourth leading cause of mortality in industrialized nations. Memory loss and cognitive decline are the most common signs of AD, and are caused by damage to and death of nerve cells related to memory in the brain area [1]. Mild cognitive impairment (MCI) is a stage that occurs between normal and AD [2]. AD progresses gradually through the prodromal stage of MCI, and finally, to AD dementia. According to studies, people with MCI acquire AD at a rate of 10–15% every year [3]. Early identification of patients with MCI can delay or prevent the progression of the disease from the MCI stage to AD. The morphological differences in the brain lesions in patients with intermediate stages of MCI are very small [3]. Furthermore, they have similar clinical manifestations; thus, to act early in the diagnosis and treatment of AD, the diagnosis and prognosis of MCI disease have been analyzed using magnetic resonance imaging (MRI) studies [4], which can capture alterations in the brain anatomy and function [5]. In recent years, several machine learning algorithms have been used for the early diagnosis of AD using magnetic resonance imaging. A support vector machine (SVM) with particle swarm optimization was applied to the classification of AD and MCI [6]. The authors of [7] suggested the exploration of deep learning in their future studies. The authors of [8,9] reported the use of clustering

techniques for clustering AD stages for early detection. The study in [10] revealed that deep learning and SVM produced higher accuracies in the classification of AD stages. However, given the very high dimensions of the input image, SVM may not be advantageous [11]. Deep learning methods have now been proposed for automatic recognition of dementia diseases [12,13] due to their ability to create an effective prediction model; in particular, convolutional neural networks (CNN) have shown great success in the computer-assisted diagnosis of AD and MCI, and are beneficial for accurate classification of AD.

In the latest studies, pre-trained CNNs have proved to be excellent in the automatic diagnosis of cognitive disease from brain MR images. AlexNet [14,15], VGG16 [16,17], VGG11 [18], ResNet-34 [19], ResNet-50 [20], U-Net [21,22], SqueezeNet and InceptionV3 [23], and DenseNet201 are examples of pre-trained deep neural networks that have been effectively used in MRI analysis. Compared to a model that comprises a single network pre-trained on MRI slices, multiple pre-trained networks on a large scale with MRI may gather potentially useful functional and structural information for discriminating the AD stages. An informative feature should keep all the important information from the input image. Comparatively, a model based on multiple pre-trained networks has shown outstanding performance in the classification of AD and MCI classification [24]. Due to the complex structure of MRIs and because the functional changes in the brain in the AD and MCI intermediate classes are closely related, multiple pre-trained networks have been designed to obtain meaningful information randomly from different layers of the pre-trained networks. This information is then combined or concatenated to extract multi-scale information from the input images [25,26]. Concatenation produces a discriminant and appropriate descriptor to further improve the power of the features of the classifier model. For example, the authors of [24] exploit the voting technique to get the required information for classification from the decision of the hybrid model. Although the model extracted the more correlated and diversified features, it could not generalize well on unseen data.

In this paper, we propose a randomized concatenation of deep features approach for automatic discrimination of patients with AD, early MCI (EMCI), late MCI (LMCI), MCI, and cognitively normal (CN) using deep learning architectures to take advantage of random information from MRI brain imaging data. This method was used to develop a definite categorization descriptor. First, each pre-trained model receives the discriminating information for the different classes from the MRI imaging data. A concatenation technique in which two fully connected layers are appended to combine the features learned, then a constant (weight) is added in concatenation. The idea behind the proposed method is that weight can reduce the value of a part of feature maps in concatenation, and then the multiplied convolutional weight will automatically enlarge the useful feature maps.

The primary contributions of this paper may be summarized as follows.

- A hybrid pre-trained CNN model for the early diagnosis of AD.
- A deep feature concatenation method for merging deep features collected from various pre-trained CNNs.
- Weight randomization to lessen the gap between feature maps in the concatenation of fully connected layers.
- Gradient-weighted class activation map to visualize discriminative regions of the image to explain the model's decision.

The remainder of the paper is structured as follows. Section 2 describes the recent research on the early detection of AD. Section 3 illustrates the proposed hybrid model for diagnosing AD. Section 4 shows the result of the experiment on the ADNI dataset, and the discussion is detailed in Section 5. Section 6 shows the comparison of the model with previous research, while we conclude the study in Section 7.

2. Related Work

Several CNN-based techniques that were pre-trained have been utilized for multiclass classification and binary classification for the early detection of AD. The authors of [26] adopted Resnet 152 to obtain a highly discriminative feature representation for

the classification of AD stages. A four-way classifier was implemented to classify AD, MCI, LMCI, and CN using ADNI imaging data. The proposed technique resulted in a prediction accuracy of 98.8%. The authors of [27] utilized VGG16-trained learning transfer for the multiclass classification of AD based on four stages (CN, AD, MCI, LMCI). The proposed model gave a testing accuracy of 95.31%. The study [28] employed multiple deep sequence-based models using 3D ResNet18 with data augmentation Resnet18 to extract features for accurate AD classification and achieved a classification accuracy of 96.88%.

In [29], the authors developed a deep learning approach based on modified Resnet18 for the binary classification of AD, including EMCI versus LMCI, AD vs. CN, CN vs. EMCI, CN vs. LMCI, EMCI vs. AD, LMCI vs. AD, MCI vs. EMCI, and MCI vs. EMCI. Their method provided an accuracy of 99.99% for EMCI vs. AD. The authors of [30] used a transfer learning technique for the three-way classification of MRI images using three pre-trained convolutional neural networks (CNN), namely: ResNet-101, ResNet-50, and ResNet-18. The experimental results showed that ResNet-101 gave an accuracy of 98.37%.

The authors of [31] proposed a 2D CNN approach based on ResNet50 with the addition of different batch normalization and activation functions to classify brain slices into three classes: cognitively normal (NC), mild cognitive impairment (MCI), and AD. The proposed model achieved an accuracy of 99.82%. Another study [32] introduced the SegNet-based deep learning method to extract AD-related brain morphological local characteristics, which are required to classify the AD condition. Resnet101 performed a three-class classification, and the results showed that the use of a deep learning technique with a pre-trained model proved highly helpful in improving the performance of the classifier. A 3D CNN was used in [33] to develop a classifier that could discriminate between AD and CN from resting-state fMRI images while achieving the accuracy of 97.77%.

The authors of [34] developed a unique method for diagnosing AD stages using probability-based fusion of various CNN models based on the DenseNet network. The proposed model was able to perform a four-way classification of CN, EMCI, LMCI and AD, on the ADNI dataset. The experimental results showed that the proposed model gave an accuracy of 83.33%. In [35], the authors used a three-dimensional net-121 architecture with a dropout rate of 0.7 for training, using the entire set of data to obtain an accuracy of 87%.

For the diagnosis of AD and MCI, the authors of [36] suggested an ensemble of densely linked 3D convolutional networks in which dense connections were introduced to improve the use of features. The ensemble model achieved an obvious improvement in its accuracy (with a 97.52% accuracy) compared to just using the simple average of the networks' predictions. The authors of [37] used convolutional network topologies for binary classification using freeze characteristics taken from the ImageNet source data set. The results of the study showed that the VGG architecture gave an accuracy of 99.27% (MCI/AD), 98.89% (AD/CN) and 97.06% (MCI/CN).

The authors of [38] proposed a layer-wise transfer learning method using VGG 19 that segregated between normal control (NC), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI) and AD. The authors of [39] suggested a transfer learning technique to reliably categorize brain sMRI slices into three classes: AD, CN and MCI. The authors used a pre-trained VGG16 network as a feature extractor and for transfer learning. The proposed model achieved a classification accuracy of 95.73%.

The authors of [40] proposed an approach for efficient classification of AD using a pre-trained AlexNet network to efficiently extract significant information from the MRI data. Their model achieved a classification accuracy of 98.35%. The authors of [41] used transfer learning to classify images from the OASIS database by fine-tuning a pre-trained AlexNet. The model showed promising results with an accuracy of 92.85% for multiclass classification. Furthermore, the authors of [42] employed a modified AlexNet with parameters to adjust for the classification of four stages of AD. The proposed strategies showed results with an accuracy of 95.70%.

The authors of [43] used fMRI data sets to classify different stages of the disease using the architecture of a CNN AlexNet. The severity of AD was divided into five phases according to the proposed model. Low- to high-level characteristics were learned during classification, resulting in an average accuracy of 97.64%. Their conclusion was that the incorporation of additional pre-trained models and transfer learning could increase classification accuracy. Similarly, the authors of [44] used modified AlexNet, ResNet-18 and GoogLeNet for the classification of brain images of CN, EMCI, MCLI, LMCI, and AD. The experimental results also showed that ResNet18 performed better.

In another study, the authors of [45] utilized different pre-trained models using a fine-tuned approach to transfer learning in the ADNI dataset for three-way AD classification (AD, MCI, and NC). The experimental results showed that DenseNet outperformed the others, achieving a maximal average accuracy of 99.05%. The authors of [46] performed an experimental analysis of some deep neural networks for the classification of AD. DenseNet-121 was found to be better than all other models used for the analysis with a classification accuracy of 90.22%.

3. Proposed Randomized Concatenated Deep Features Approach

As previously stated, the suggested technique seeks to provide an accurate diagnosis of Alzheimer's disease by weight-randomizing concatenated deep features taken from the ResNet18 and DenseNet121 networks. Figure 1 shows the pipeline of the proposed randomized concatenated deep features-based classification system to identify patients with AD, MCI, EMCI, LMCI, and NC clinical status using MRI neuroimaging data.
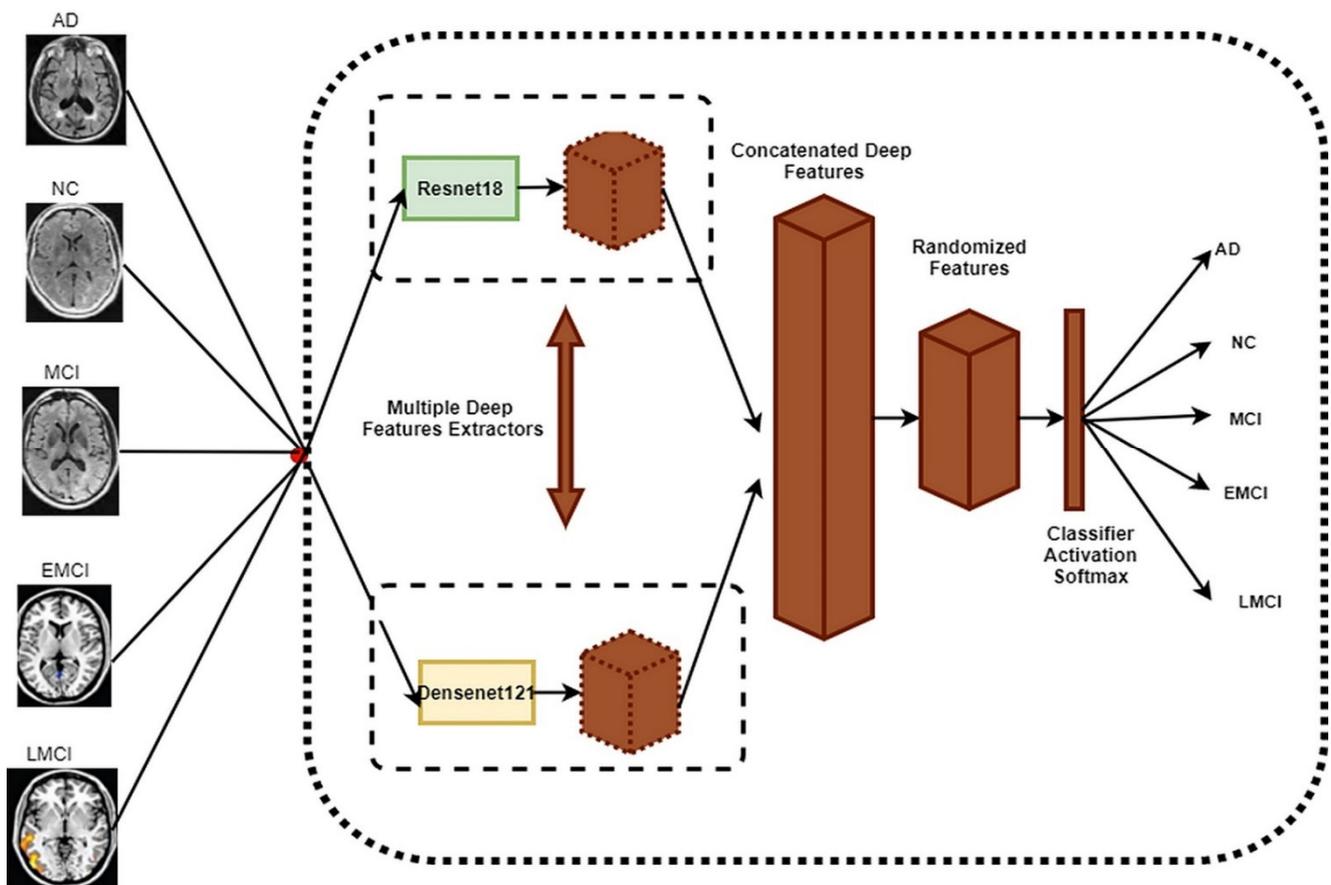

Figure 1. The main framework for the proposed method.

3.1. Dataset

The data for this study were obtained from the ADNI database. ADNI is a long-term research project aimed at developing clinical, imaging, genetic, and biochemical indicators for the early diagnosis and tracking of Alzheimer's disease. There are 138 MRI scans in the dataset, with 25 CN, 25 SMC, 25 EMCI, 25 LMCI, 13 MCI, and 25 AD scans [18]. The participants are above the age of 71, and each has been diagnosed with Alzheimer's disease and assigned to one of these phases according to their cognitive test results [18]. Figure 2 shows the class distribution of the MRI dataset. A total of 7509 processed MRI images were evaluated from the ADNI database and the data split utility (random split) in PyTorch divided the data into 5256 and 2253 images for training and validation, respectively. Details of the data splitting are given in Figure 3. A new set of samples were extracted from subjects that were separate from the subjects used for training and validation to test the model for generalizability.
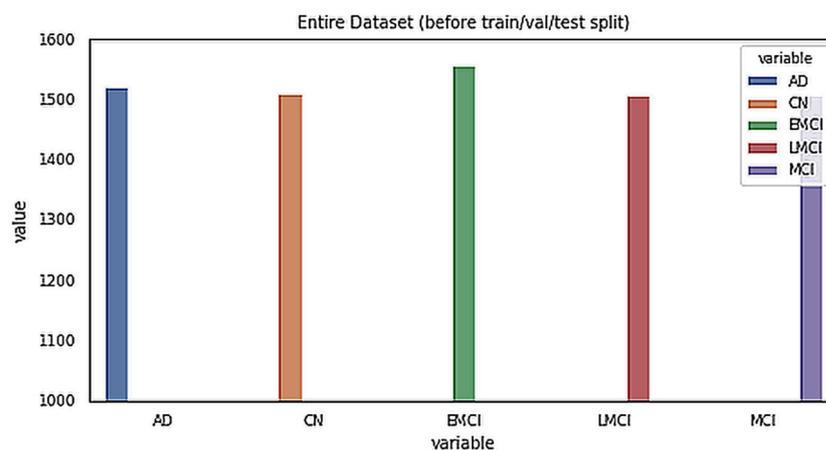

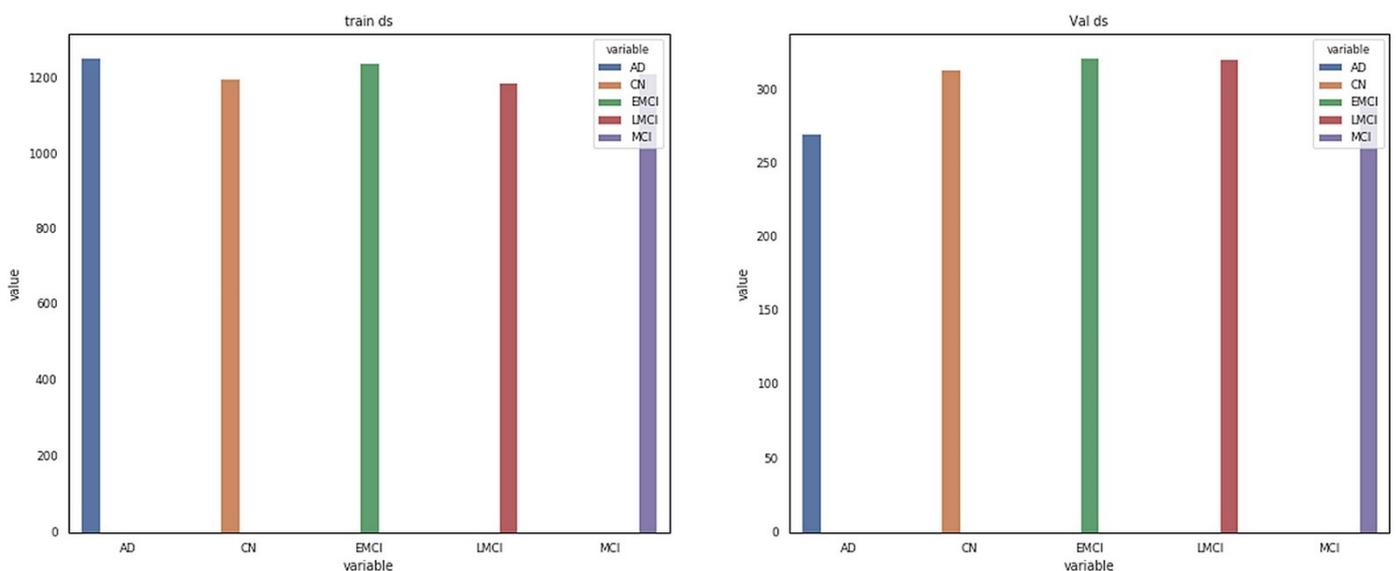**Figure 2.** Class distribution of the MRI dataset.

Figure 3. Details of dataset splitting per class.

3.2. Deep Feature Extraction

The ResNet18 and DenseNet121 deep CNN architectures were used in the feature extraction procedure from MRI images. In CNN designs, multiple layers are used to extract features. The offered layers include convolutional pooling, batch normalization, rectified linear unit (ReLU), SoftMax, and fully connected (FC) [47]. The FC layer was kept the same

for all models. The convolutional layer was made up of several weights H^l kernels for each layer y to take the input X^{y-1} and extract the local characteristics. Both Resnet18 and DenseNet121 were set with weights that had been pre-trained using natural photos from ImageNet [48]. Then, we ran each model one time on our training and validation images to extract the deep features.

3.3. Concatenation of Deep Features

Concatenation of the recovered deep features is an effective approach to combine multiple characteristics to improve the classification process [49]. In this study, the concatenation procedure was achieved by extracting characteristics from images using Resnet18 and DenseNet121 features. The high-level features of the entire connected layers of both Resnet18 and Densenet21 were linked into a single vector to form the classification descriptor as shown in Equation (1).

$$Feature\ Descriptor = F^{(Resnet-18)} \cup F^{(Densenet-121)} \tag{1}$$

3.4. Weight Randomization and Classification

Combining information by concatenating or adding leads to different information being mixed in the fusion layer. However, some of this information may be useless, and as such, narrowing the gap among the fused layers can result in better classification accuracy. In this study, an inbuilt weight initialization in PyTorch was implemented, including a uniform Kaiming distribution [50,51] and Xavier (Glorot). The Kaiming weight initialization was created to perform non-symmetrical activation functions such as ReLUs, while the Xavier initialization was designed for layers with sigmoid activators. Each layer's weights are created using a normal distribution. The resulting tensor using the Kaiming weight will have values sampled from $N\ (-bound,\ bound)$, where *bound* is described in Equation (2).

$$bound = gain \cdot \sqrt{\frac{3}{fan\_mode}} \tag{2}$$

The resulting tensor using Xavier will have values sampled from $U(-a,\ a)$ where *a* is described in Equation (3).

$$a = gain \cdot \sqrt{\frac{6}{fan_{in} + fan\_out}} \tag{3}$$

The following parameters depicted in Table 1 are considered in the Kaiming weight initialization and the Xavier weight initialization.

Table 1. Parameters for Kaiming and Xavier weight initialization.

Parameters	Description
Tensor	*n*-dimensional torch.Tensor
a	Negative slope of the rectifier
mode	"Fan _in" conserves the degree of the variance of the weights in the forward pass
nonlinearity	The non-linear function (nn.functional name)

The last description of the feature determines whether the input image is classified as MCI or LMCI or EMCI or AD or CN. The fully connected layer turns the input data into a 1D vector, and the SoftMax layer computes the five class scores.

3.5. Gradient-Weighted Class Activation Map (Grad-CAM)

Grad-CAM is the generalization of the class activation map (CAM), which locates the discriminative region(s) for a CNN prediction by computing class activation maps with gradient information. Grad-CAM creates a map of the working class by incorporating

gradient data into the final decision layers. Grad-CAM weighs the 2D activations with the average gradient. It helps to understand what the network sees, and which neurons are firing in a particular layer, given an input image [51]. The last class gradient related to a channel is used to measure all channels, following the final convolution layer to create a localization map showing the critical locations in an image that have a significant impact on the model forecast. To obtain the class discrimination localization map, the class score gradient was calculated relative to the feature maps of the convolutional layer [52]. Then, to obtain the key weights of the neuron, we took a global average of these gradients as described in Equation (4).

$$\propto_u^v = \frac{1}{Y} \sum_{i=0}^{s} \sum_{j=0}^{t} \frac{\partial w^v}{\partial X_{ij}^u} \tag{4}$$

where s is width for any class, v is height for any class, ∂w^v is gradient for the class score, X^U are feature maps, and $\propto_u^v$ are neuron weights.

Finally, the Grad-CAM map was obtained by linearly combining the weights with the activation map of ReLU.

3.6. Implementation Details

Our proposed study was created on the NVIDIA Corporation Tu116 (Geforce GTX 1660) GPU using Python 3.6 with the Pytorch package. The proposed study was evaluated on the prepared dataset using the random split approach, and the details of the data split can be found in Table 1. The images of the dataset were resized to 224×224 pixels, and a batch size of 10 was kept due to memory usage with the number of epochs, which was 10. The number of loader worker processes was set to 4.

For optimization of the hyperparameters of the deep learning model, we used the AutoKeras 1.0.8 library, which optimizes both the architecture and hyperparameters as guided by the Bayesian optimization to select the best hyperparameter values [53]. Based on the hyperparameter optimization stage, the learning rate was fixed to 0.0001 for the training of the proposed model. A dropout of 0.4 and weight decay of 0.002 were also used to fit the proposed model for the purpose of fine-tuning.

The architecture of the prototype system developed for AD diagnostics is presented in Figure 4, while its deployment is shown in Figure 5. The doctor uses his/her personal computer to access the patient's records on the hospital server via a dedicated web service. The hospital server retrieves the patient's brain MRI images from the MRI server, which are stored in NIfTI (Neuroimaging Informatics Technology Initiative) format. The diagnostic decision is taken by the Alzheimer's disease classifier implementing the approach described in this paper. The decision is supported by Grad-CAM attention images that explain the suggested decision.
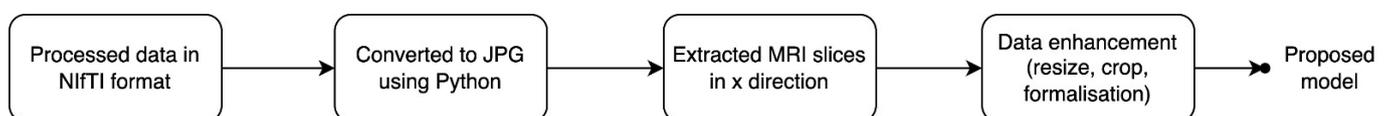

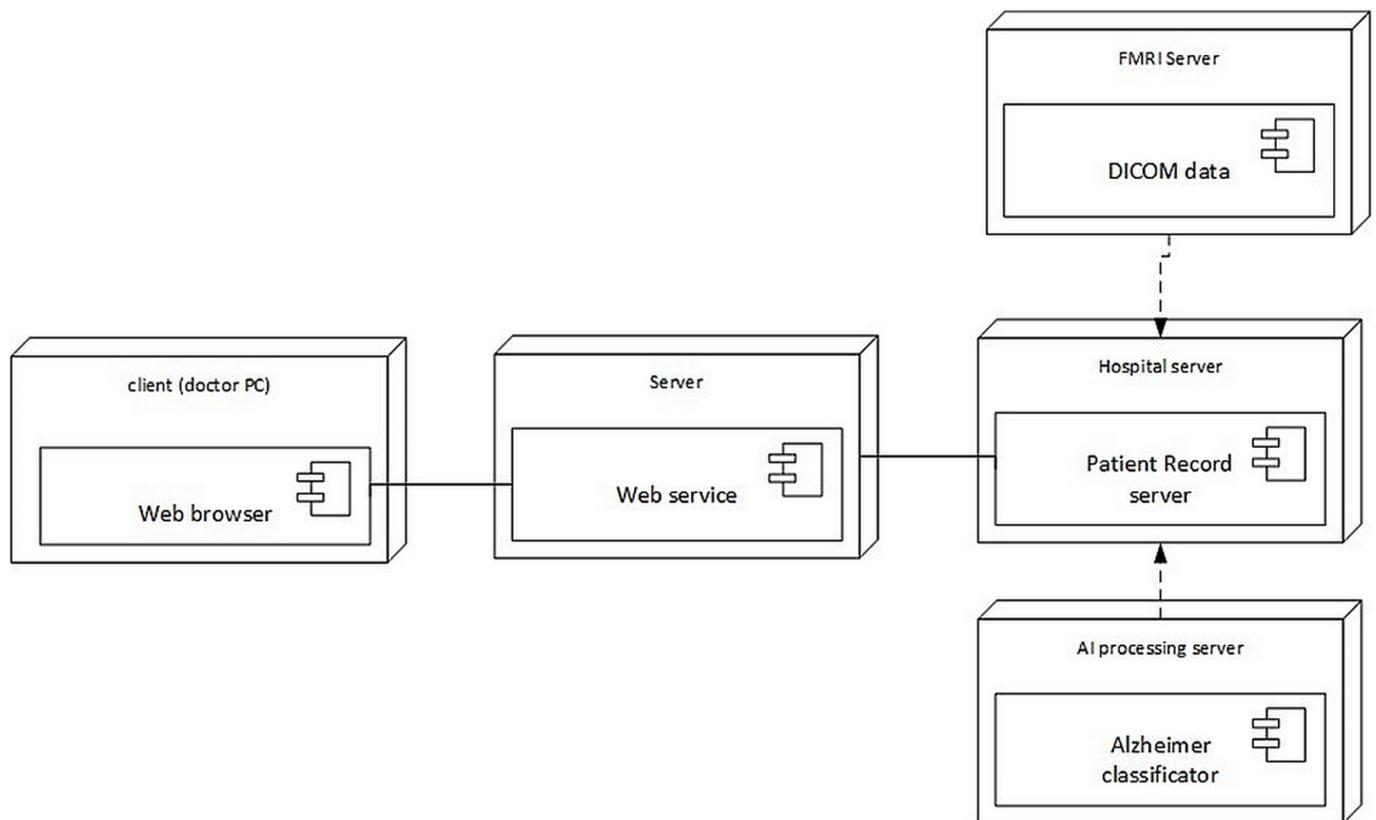**Figure 4.** Detailed workflow of data used in the proposed model.

Figure 5. Deployment of the developed system architecture.

4. Experimental Results

We provide the experimental results for one five-way benchmark multiclass classification problem [18], one four-way multiclass classification problem [33], and one three-way classification problem [45]. The suggested model's training efficiency was evaluated in terms of important parameters, i.e., training accuracy, validation accuracy, training loss, and validation loss at different epochs without dropout, with dropout, with dropout and weight decay. Learning rates of 0.0001 and 0.0003 were optimized with stochastic gradient descent (SGD) and adaptive moment estimation (ADAM). The proposed model performed better when using a learning rate of 0.0001 with the SGD optimizer. These parameters are calculated to estimate the trained models with a learning rate of 0.0001 optimized with SGD. These parameters are calculated to estimate the overfitting of the trained models. The graphs of the training loss/validation accuracy and training accuracy/validation accuracy of the proposed model with a random split are depicted in Figure 6. Furthermore, a confusion matrix was generated for the proposed model to quantify the performance metrics, i.e., precision, recall, F1 score, and accuracy. The results of the parameters with the five-way multiclass, the four-way multiclass, and the three-way multiclass problem using the Kaiming and Xaiver weight initialization are presented in Tables 2 and 3, respectively. Table 4 shows the results of the suggested model on the test dataset, while Table 5 shows the results of the proposed model on test data from different subjects. Each of the classes consists of 20 slices each.
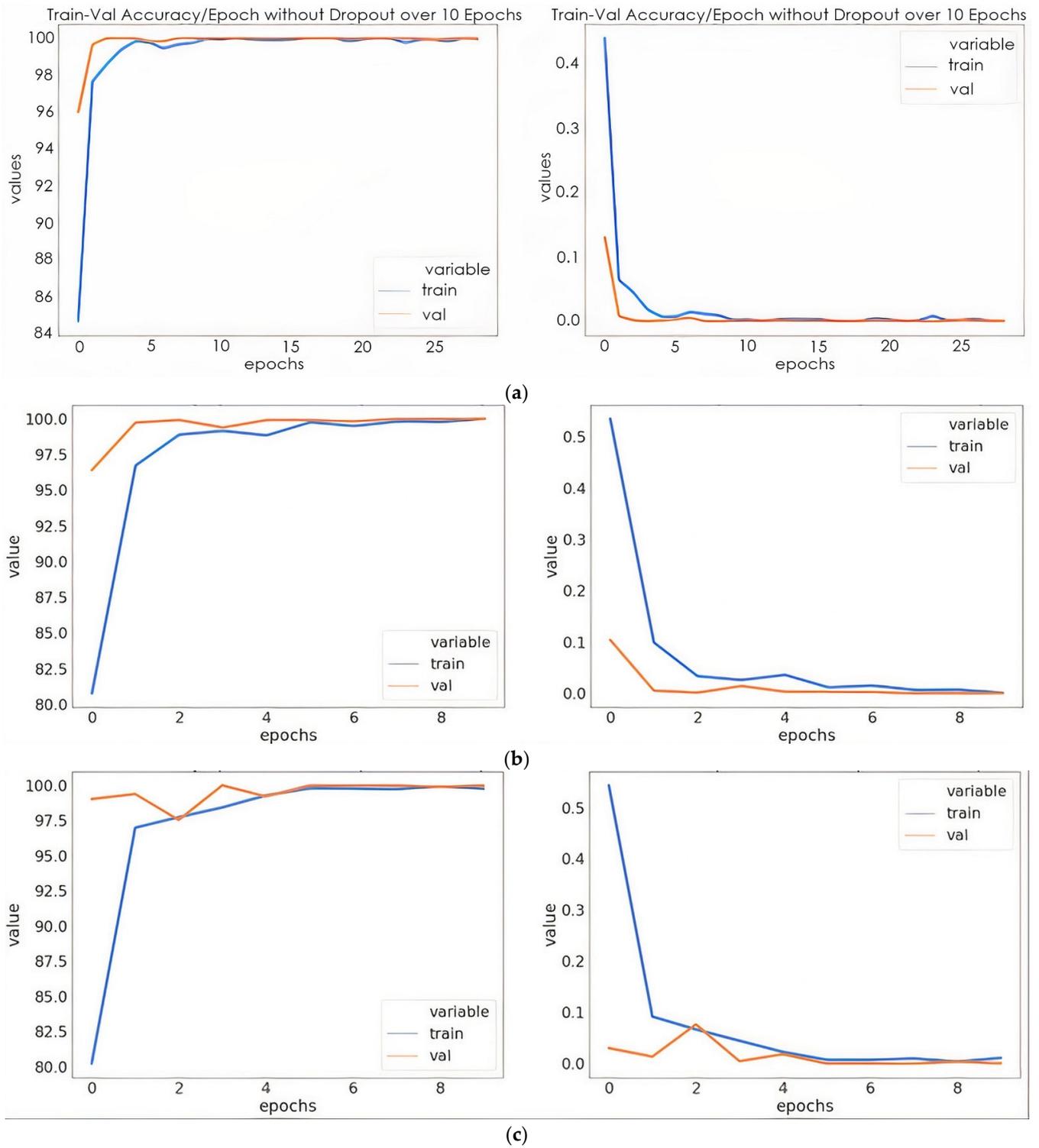
Figure 6. Training/validation accuracies and training loss/validation loss: (**a**) model without dropout; (**b**) mode with dropout; (**c**) model with dropout and weight decay discussions.

Table 2. The training performance of the proposed model on various multiclass classifications utilizing Kaiming weight initialization.

Ways of Multiclass	Epochs	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
5 Ways	1	54.93	91.76	1.09	0.32
	2	87.99	98.85	0.33	0.08
	3	96.11	99.87	0.14	0.02
	4	97.97	99.78	0.08	0.01
	5	98.97	99.91	0.06	0.01
	6	99.60	99.10	0.04	0.09
4 Ways	1	50.59	89.21	1.17	0.49
	2	86.86	98.89	0.42	0.12
	3	95.68	99.56	0.17	0.04
	4	96.99	99.61	0.10	0.02
	5	97.91	99.78	0.07	0.01
	6	99.30	98.90	0.05	0.16
3 Ways	1	68.26	93.15	0.76	0.25
	2	89.92	99.26	0.28	0.07
	3	95.36	99.78	0.14	0.03
	4	97.17	99.85	0.09	0.01
	5	97.90	99.93	0.07	0.01
	6	98.50	98.70	0.04	0.16

Table 3. The proposed model's training performance on various multiclass classifications utilizing Xaiver weight initialization.

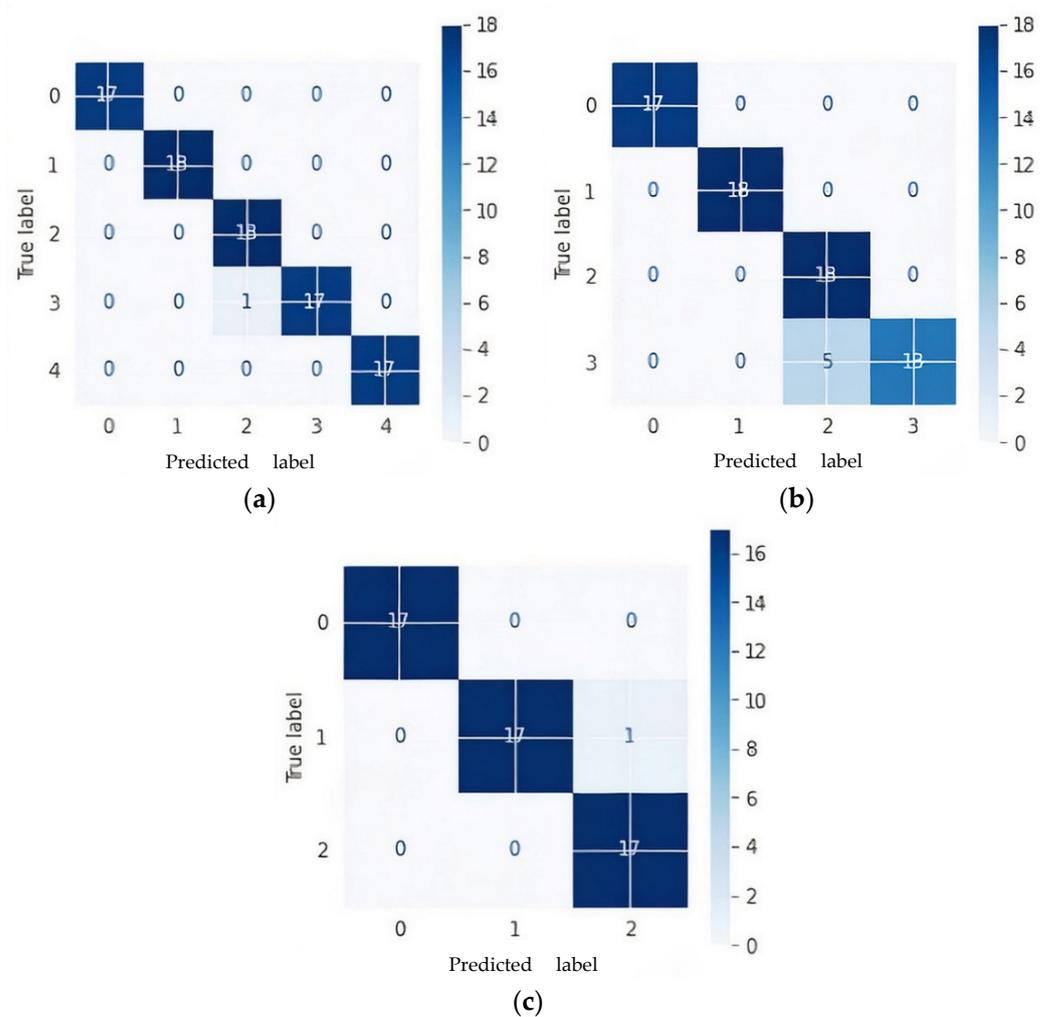| Ways of Multiclass	Epochs	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
5 Ways	1	51.20	81.68	1.22	0.67
	2	81.22	96.54	0.59	0.22
	3	93.53	99.07	0.25	0.08
	4	97.32	99.51	0.13	0.04
	5	98.61	99.69	0.08	0.02
	6	99.30	99.73	0.05	0.02
4 Ways	1	44.88	78.63	1.25	0.77
	2	79.29	95.08	0.64	0.28
	3	92.96	99.34	0.28	0.09
	4	96.75	99.83	0.14	0.04
	5	98.29	99.94	0.09	0.02
	6	98.84	99.98	0.06	0.01
3 Ways	1	62.12	86.16	0.85	0.39
	2	85.56	96.35	0.40	0.15
	3	93.26	99.18	0.22	0.07
	4	96.63	99.85	0.13	0.03
	5	97.33	99.78	0.09	0.02
	6	98.31	99.85	0.07	0.01

The confusion matrix was generated using a test dataset taking samples from one subject that was not used for training. The confusion matrix for a five-way multiclass, four-way multiclass, and three-way multiclass is represented in Figure 7 and 0, 1, 2, 3, and 4 labels are represented as AD, CN, EMCI, LMCI, and MCI, respectively.

Table 4. Test accuracy and test loss of the proposed model based on Xaiver and Kaiming weight initialization.

Weight Initialization	Ways of Multiclass	Test Accuracy (%)	Test Loss
Kaiming	5 ways	98.86	0.05
	4 ways	93.06	0.14
	3 ways	98.21	0.06
Xaiver	5 ways	87.50	0.43
	4 ways	88.89	0.24
	3 ways	96.21	0.04

Table 5. Results of the proposed model for test data with different subjects.

Ways of Multiclass	Accuracy (%) First Test Sample	Accuracy (%) Second Test Sample	Accuracy (%) Third Test Sample	Accuracy (%) Fourth Test Sample	Standard Deviation
5 ways	98.86	98.07	98.98	98.90	0.42
4 ways	93.06	93.02	94.10	93.20	0.50
3 ways	98.21	98.40	98.04	99.01	0.42

Figure 7. Confusion matrices of proposed model based on Kaiming weight initialization on test data: (**a**) 5-way multiclass; (**b**) 4-way multiclass; (**c**) 3-way multiclass.

The proposed model gave an average precision and recall of 98.94% and 98.89%, 89.61% and 88.89%, 98.14%, 98.14% for the five-way, four-way, and three-way multiclass, respectively. The per class classification report of the proposed model based on precision, recall, and F1-score is detailed in Figure 8. Furthermore, the proposed model was checked to see if the predicted label matched the actual label by examining the number of images predicted correctly and incorrectly.
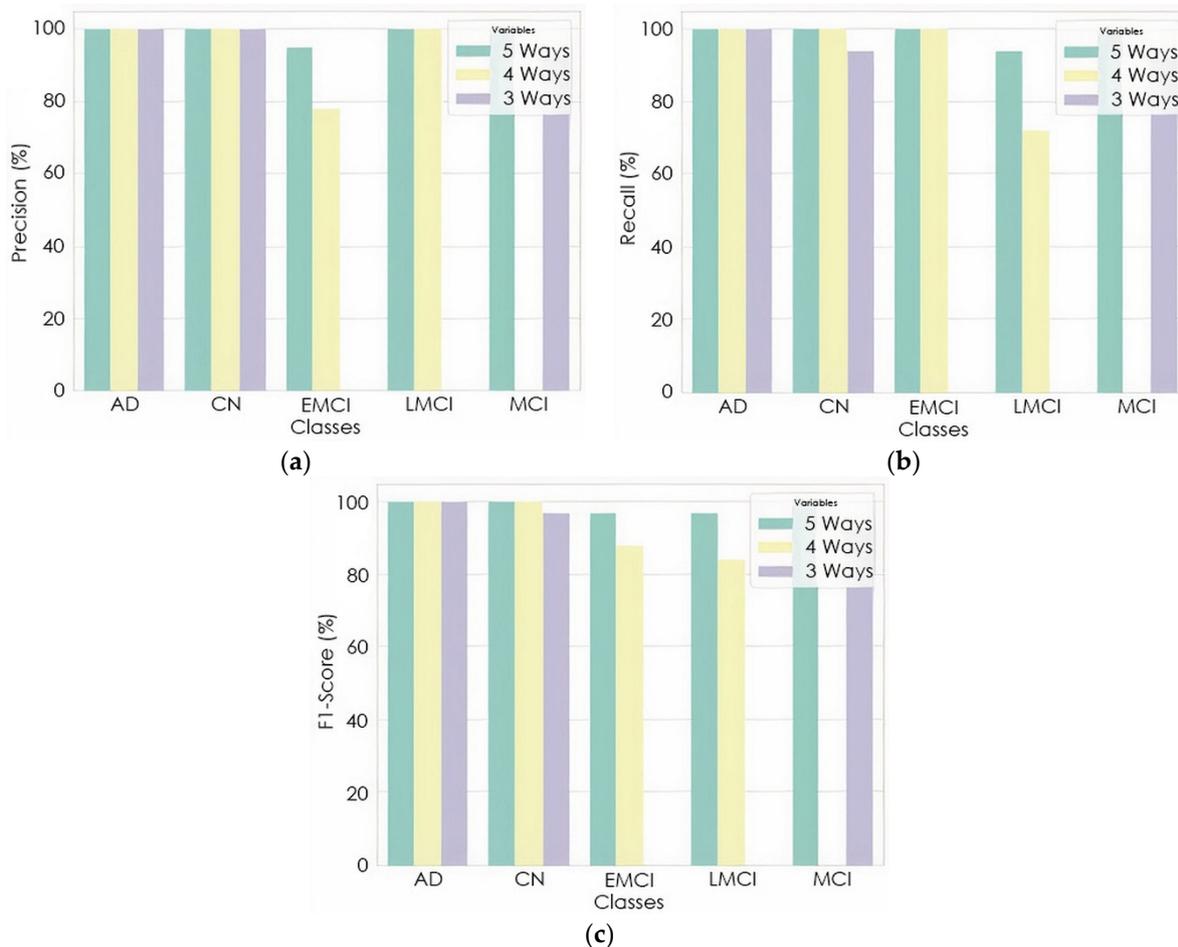

Figure 8. Per class classification report: (**a**) precision; (**b**) recall; (**c**) F1-Score.

To verify the efficacy of our proposed technique, the representation of the output data from the model before classification was featured in the form of cluster figures by the t-SNE projection, as shown in Figure 9 for the four-way multiclass and three-way multiclass classification. This was achieved by reducing the dimensional output layer down to five dimensions, four dimensions, and three dimensions for five-way, four-way, and three-way classification, respectively.
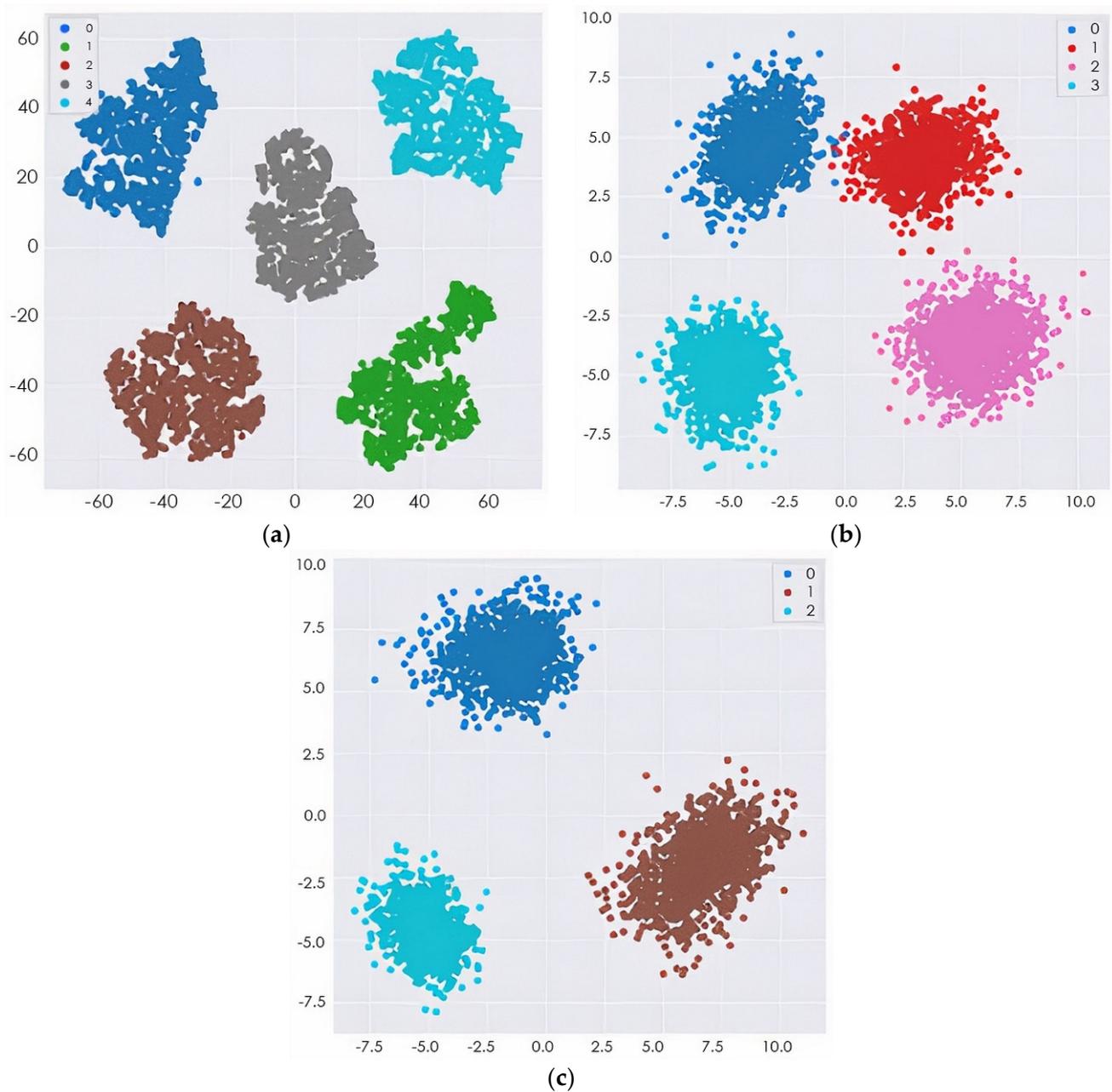
Figure 9. The t-SNE projection visualization of several MRI features: (**a**) 5-way multiclass classification; (**b**) 4-way multiclass classification; (**c**) 3-way multiclass classification.

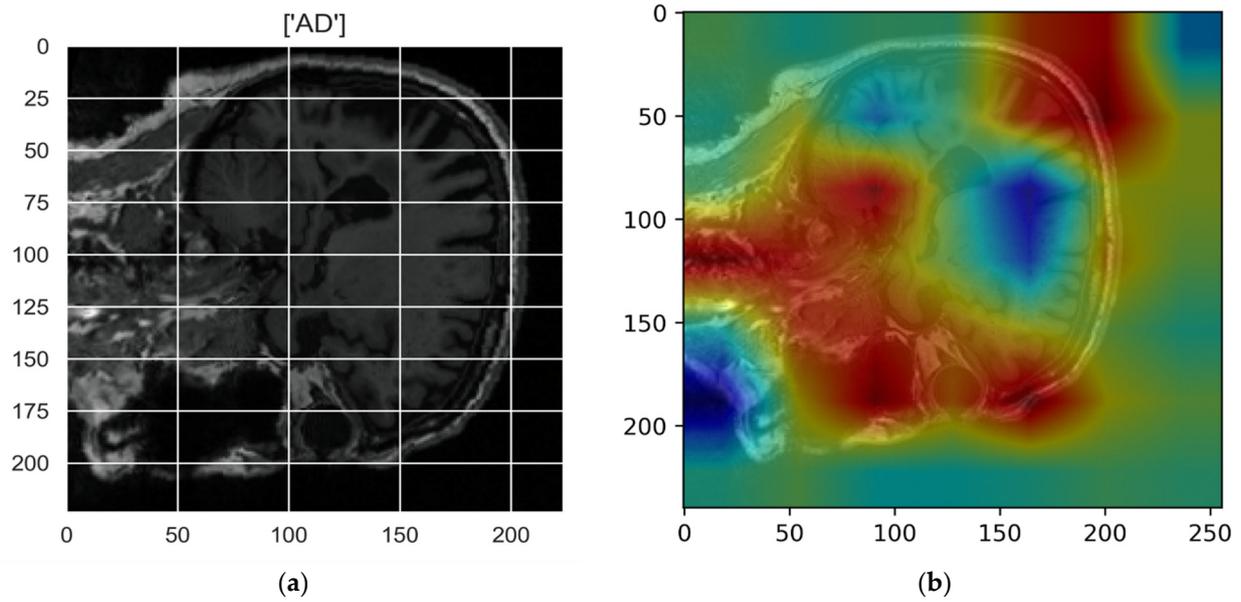Figures 10–14 show the result of the Grad-CAM on the predicted classes AD, EMCI, CN, MCI, and LMCI respectively.

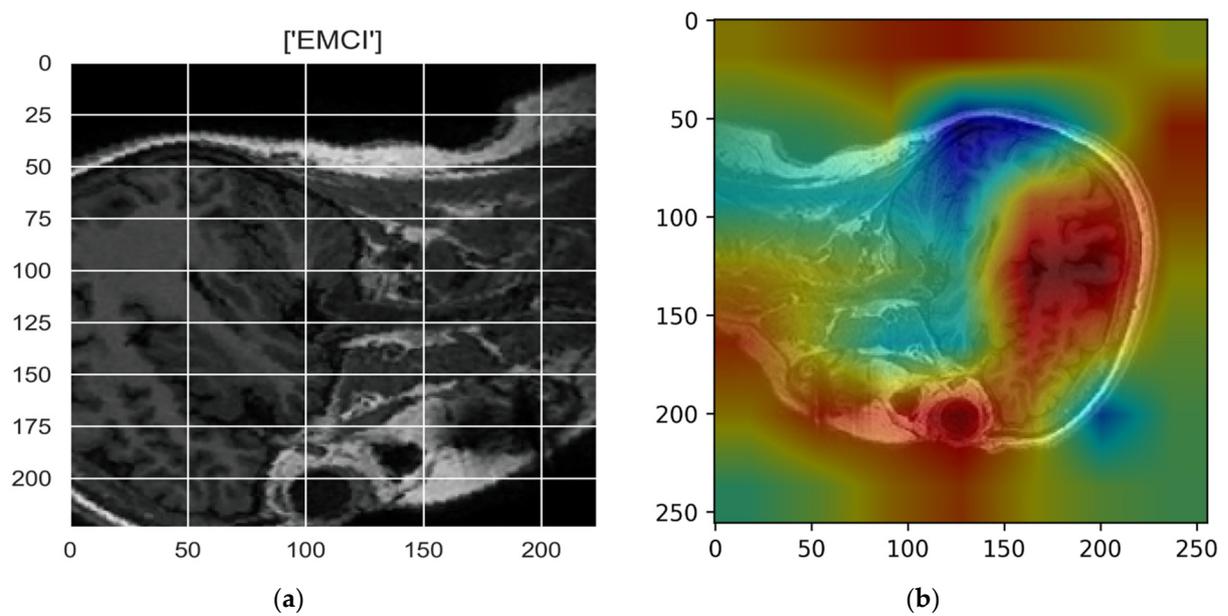**Figure 10.** Grad-CAM explanation for the prediction of class AD: (**a**) MRI image, (**b**) Grad-CAM attention map.

Figure 11. Grad-CAM explanation for the prediction of class EMCI: (**a**) MRI image, (**b**) Grad-CAM attention map.
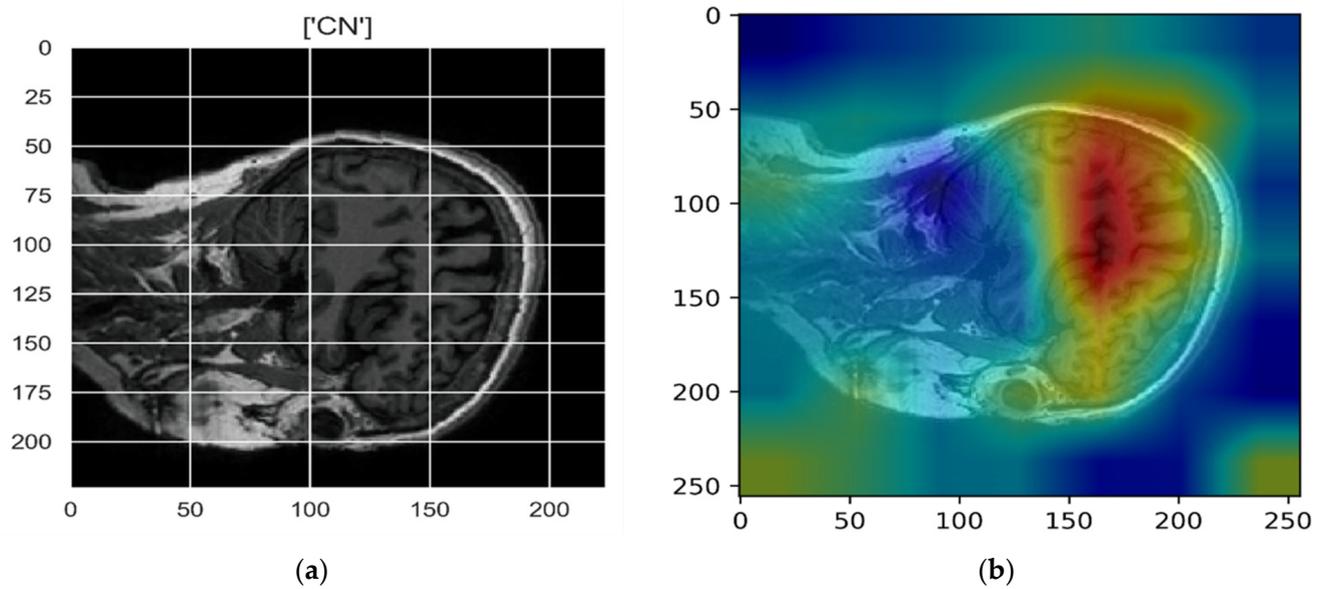
Figure 12. Grad-CAM explanation for the prediction of class CN: (**a**) MRI image, (**b**) Grad-CAM attention map.
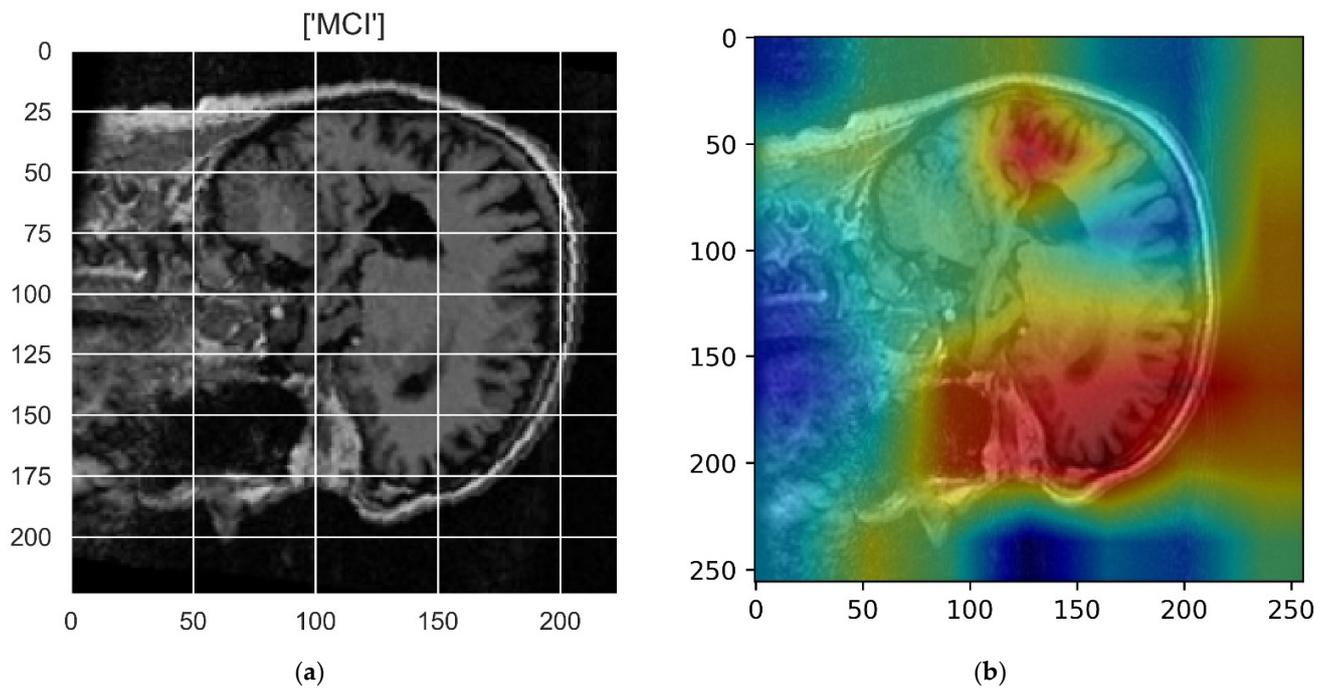

Figure 13. Grad-CAM explanation for the prediction of class MCI: (**a**) MRI image, (**b**) Grad-CAM attention map.
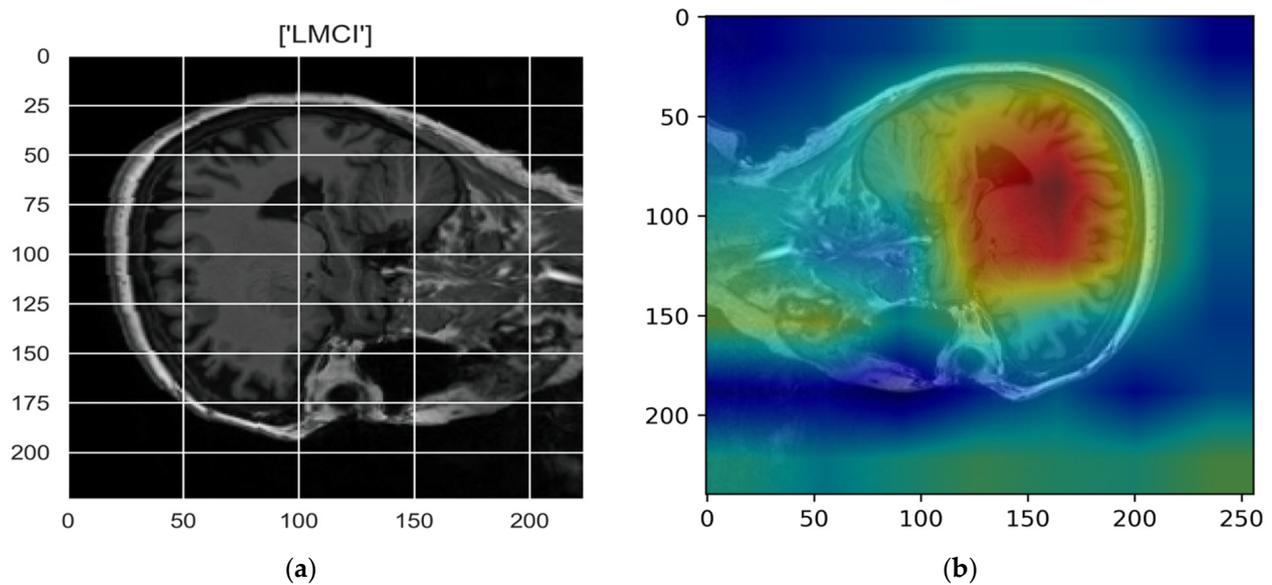
Figure 14. Grad-CAM explanation for the prediction of class LMCI: (**a**) MRI image, (**b**) Grad-CAM attention map.

5. Discussion

The proposed model has minimal training and validation loss and exhibits the best training and validation accuracy with dropout (0.4) and a weight decay of 0.02. In Figure 6a, the training and validation accuracy of the proposed model without dropout clearly show that the training accuracy was steady at 6 epochs. This indicates that the model was no longer learning. Likewise, the validation accuracy remained the same after 5 epochs. Although the model was able to fit adequately with the use of a 0.4 dropout, the accuracy was lower than that of the model with a dropout of 0.4 and weight decay of 0.002. The overfitting was reduced to minimal with the use of dropout and weight decay. However, the validation accuracy was slightly higher than the training accuracy with a very close margin as the number of epochs increased. In Tables 2 and 3, the proposed model provided the best training and validation accuracy at 5 epochs for all the ways of multiclass classification. As shown in Table 4, the proposed model based on the initialization of the Kaiming weight produced a testing precision with a three-way multiclass classification accuracy of 100%, a four-way multiclass classification accuracy of 95.83%, and a five-way multiclass classification accuracy of 98.86%. It was evident from the results that the proposed model accurately discriminates between three-way multiclass (AD/MCI/CN) and five-way multiclass (AD/CN/LMCI/EMCI/MCI) classifications of Alzheimer's disease. However, for the purpose of this study, more emphasis was on the five-way multiclass classification. Table 4 shows that for the Kaiming weight initialization technique, the three different multiclass methods demonstrate the highest test accuracies. Therefore, Kaiming weight initialization was set for this study. As we can see in Table 4, for all classification problems, the test accuracy of the proposed model using the Kaiming weight initialization is superior to the Xavier weight initialization technique. The result of the generalization of the proposed model on different test data is shown in Table 5 with a standard deviation of 0.42, 0.50 and 0.42 for five-way, four-way, and three-ways multiclass classification, respectively.

Figure 6 shows that one subject of LMCI was misclassified as EMCI as in the case of five multiclass classifications while five LMCI were incorrectly diagnosed as EMCI in four multiclass classifications. Although LMCI was found to be very close to being diagnosed as AD, from the confusion matrixes in Figure 6a,b, it was seen as EMCI. This is because a single modality cannot capture more disparate differences between EMCI and LMCI. One CN was misclassified as MCI in both five-way and three-way multiclass, which was an

indication of an effective model because in medical diagnosis, it is preferred to screen a person as diseased than to eliminate a diseased person by falsely predicting a negative.

As shown in Figure 7, a precision of 95% and 73% was achieved to distinguish EMCI from other classes in five-ways and four-ways multiclass, respectively, while a precision of 100% and 86% was also achieved to distinguish LMCI from other classes in five-way and four-way, respectively. The proposed model also achieved a precision of 94% and 94% to differentiate MCI from other classes in five-way and three-way multiclass classification, respectively. From the per class classification result in Figure 7, five-way multiclass had the highest recall of 100% and 94% in differentiating EMCI and LMCI, respectively, from other classes.

The proposed model was trained to be incredibly confident in its predictions, as an image with a confidence score of 0.687 had the predicted intent of being a solid subject for the predicted class. This result also showed that if the features of the different classes (EMCI and LMCI) could be more represented, the misclassification error would be greatly reduced, thus increasing the classification accuracy. The confidence in the predicted label was very high because some of the numerical properties provided by the initialization technique were not kept when the training process updated the weight values. The confidence level of the proposed model for the incorrectly classified image was 0.602. The designated classes of each multiclass classification look well separated, as depicted in Figure 8. This could improve the model classification performance.

In the Grad-CAM maps in Figures 10 and 11, the red regions highlight the most important discriminative regions, and other colored regions are less important. In MRI images with a sagittal plane view, the model focuses on the vermis of cerebellum [54], and the fourth ventricle areas of the posterior and anterior lobes [55] are involved in the prediction of the AD class. For the EMCI prediction, white matter hyperintensities (WMH) are found to be associated with EMCI class prediction [56]. In Figures 12 and 13, the model focused on the internal cerebra vein to identify the pattern for the prediction of CN and MCI, as cerebral basal vein dilation is related to the volume of the white matter hippocampus (WMH) [57].

6. Comparison with Existing Methods

To our knowledge, our approach is the first to have a concatenated randomized output from two pre-trained models for the early diagnosis of Alzheimer's disease based on a multiclass classification of five-ways, four-ways, and three-ways. To assess our method, we compared the proposed technique to similar studies that used the same parameters, as shown in Table 6. It can be clearly seen that the proposed method yielded the best results in regard to all metrics. The neuroimaging data of all these comparative studies are based on the ADNI website, and their methods are confined to pre-trained networks.

Table 6. Classification performance comparison.

Authors	Methodology	Multiclass	Accuracy (%)	Precision (%)	Recall (%)
Ramzan et al., (2019) [18]	Resnet 18 (Finetuning)	5 Ways AD/CN/EMCI/LMCI/MCI	97.88	98.10	97.89
Parmar et al., (2020) [33]	3D CNN	4 Ways AD/CN/EMCI/LMCI	93.00	93.18	-
Puete-Castro et al., (2020) [58]	Resnet18 and SVM	3 Ways AD/CN/MCI	78.72	68.96	58.66
Proposed	Resnet18 and DenseNet121 with Randomized weight	5 Ways AD/CN/EMCI/LMCI/MCI	98.86	98.94	98.89
Proposed		4 Ways AD/CN/EMCI/LMCI	93.06	94.56	93.05
Proposed		3 Ways AD/CN/MCI	98.21	98.14	98.14

As described in Section 2, the studies described in [18,33,58] utilized deep learning using the transfer learning method for the early diagnosis of Alzheimer's disease. As shown in Table 6, the multiclass classification of the proposed method in the five-way multiclass classification outperformed the results of [18], achieving an accuracy of 98.86%, which is 0.98% higher than that achieved by the study in [18], and the classification accuracy in the four-way multiclass classification is much higher than that of the study in [33], demonstrating the effectiveness of our model. Additionally, in our comparative analysis, we discovered that our model performed better than the study in [58], as our model achieved a classification accuracy of 98.21% and a precision of 98.14%.

7. Conclusions

In this study, we adopted two pre-trained models to learn features simultaneously from MRI images, and the learned features were concatenated for AD classification. The concatenation of features amounted to distant or irregular information in fully connected layers during the classification process, and we hypothesized that adopting weight randomization would reduce the gap between feature maps. We tested our hypothesis by performing detailed experiments using brain MRI images from the ADNI dataset, which serves as a benchmark, where 25 subjects were used from each of the five categories of AD, MCI, EMCI, LMCI, and NC (a total of 125 subjects). We investigated the effectiveness of different weight randomization in our application by using Kaiming weight initialization and Xavier weight initialization in our model, and we presented in-depth results and their relation to the number of classes for multiclass classification based on precision. Our results showed that the MRI features concatenated by our proposed model improve the discrimination between classes in five-way multiclass classification compared to four-way multiclass classification.

Finally, our findings were compared to four other state-of-the-art methodologies, where our proposed strategy outperformed them by a significant margin and resulting in a 0.98% and 0.06% improvement in the accuracy of AD vs. CN vs. EMCI vs. LMCI vs. MCI and AD vs. CN vs. EMCI vs. LMCI classification problems, respectively. Likewise, the model achieved an increase of 0.84% and 1.38% in precision for AD vs. CN vs. EMCI vs. LMCI vs. MCI and AD vs. CN vs. EMCI vs. LMCI multiclass classification, respectively. Although the weight mechanism has been proven to be an effective strategy to collect information from feature maps [17], the high classification accuracy of the test data obtained in this paper proved that the weight mechanism can act as a good strategy to automatically enlarge useful feature maps for clinical diagnosis.

There were some limitations in the present study. A multimodal input should be further considered to enhance the capture of more disparate differences between EMCI and LMCI. In addition, a base domain model learned from brain images is possibly more appropriate to the target domain. A large dataset is also required to drastically reduce overfitting. Tackling these limitations in future work will continue to improve the robustness of the recommended approach.

Author Contributions: Conceptualization, R.M.; methodology, R.M.; software, M.O.; validation, M.O., R.M. and R.D.; formal analysis, M.O., R.M. and R.D.; investigation, M.O., R.M. and R.D.; resources, R.M.; data curation, M.O.; writing—original draft preparation, M.O. and R.M.; writing—review and editing, R.D.; visualization, M.O.; supervision, R.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The ADNI database is available from http://adni.loni.usc.edu/ (accessed 23 December 2021).

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

References

1. Mohamed, T.M.; Youssef, M.A.M.; Bakry, A.A.; El-Keiy, M.M. Alzheimer's disease improved through the activity of mitochondrial chain complexes and their gene expression in rats by boswellic acid. *Metab. Brain Dis.* **2021**, *36*, 255–264. [CrossRef]
2. Tadokoro, K.; Yamashita, T.; Fukui, Y.; Nomura, E.; Ohta, Y.; Ueno, S.; Nishina, S.; Tsunoda, K.; Wakutani, Y.; Takao, Y.; et al. Early detection of cognitive decline in mild cognitive impairment and Alzheimer's disease with a novel eye tracking test. *J. Neurol. Sci.* **2021**, *427*, 117529. [CrossRef]
3. Zhang, T.; Liao, Q.; Zhang, D.; Zhang, C.; Yan, J.; Ngetich, R.; Zhang, J.; Jin, Z.; Li, L. Predicting MCI to AD Conversation Using Integrated sMRI and rs-fMRI: Machine Learning and Graph Theory Approach. *Front. Aging Neurosci.* **2021**, *13*, 429. [CrossRef]
4. Hojjati, S.H.; Ebrahimzadeh, A.; Khazaee, A.; Babajani-Feremi, A. Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM. *J. Neurosci. Methods* **2017**, *282*, 69–80. [CrossRef]
5. Dai, Z.; Lin, Q.; Li, T.; Wang, X.; Yuan, H.; Yu, X.; He, Y.; Wang, H. Disrupted structural and functional brain networks in Alzheimer's disease. *Neurobiol. Aging* **2019**, *75*, 71–82. [CrossRef]
6. Zeng, N.; Qiu, H.; Wang, Z.; Liu, W.; Zhang, H.; Li, Y. A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease. *Neurocomputing* **2018**, *320*, 195–202. [CrossRef]
7. Gamberger, D.; Lavrač, N.; Srivatsa, S.; Tanzi, R.E.; Doraiswamy, P.M. Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease. *Sci. Rep.* **2017**, *7*, 1–12. [CrossRef]
8. 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]
9. Afzal, S.; Maqsood, M.; Khan, U.; Mehmood, I.; Nawaz, H.; Aadil, F.; Song, O.Y.; Nam, Y. Alzheimer Disease Detection Techniques and Methods: A Review. *Int. J. Interact. Multimed. Artif. Intell.* **2021**, *6*, 26–38. [CrossRef]
10. Wen, J.; Thibeau-Sutre, E.; Diaz-Melo, M.; Samper-González, J.; Routier, A.; Bottani, S.; Dormont, D.; Durrleman, S.; Burgos, N.; Colliot, O.; et al. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. *Med. Image Anal.* **2020**, *63*, 101694. [CrossRef]
11. Noor, M.B.T.; Zenia, N.Z.; Kaiser, M.S.; Mamun, S.A.; Mahmud, M. Application of deep learning in detecting neurological disorders from magnetic resonance images: A survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain Inf.* **2020**, *7*, 11. [CrossRef]
12. Li, H.; Habes, M.; Wolk, D.A.; Fan, Y. A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimer's Dement.* **2019**, *15*, 1059–1070. [CrossRef]
13. Nawaz, H.; Maqsood, M.; Afzal, S.; Aadil, F.; Mehmood, I.; Rho, S. A deep feature-based real-time system for Alzheimer disease stage detection. *Multimed. Tools Appl.* **2020**, *80*, 35789–35807. [CrossRef]
14. Shakarami, A.; Tarrah, H.; Mahdavi-Hormat, A. A CAD system for diagnosing Alzheimer's disease using 2D slices and an improved AlexNet-SVM method. *Optik* **2020**, *212*, 164237. [CrossRef]
15. Hon, M.; Khan, N.M. Towards Alzheimer's disease classification through transfer learning. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 13–16 November 2017; pp. 1166–1169. [CrossRef]
16. Qiu, S.; Joshi, P.S.; Miller, M.I.; Xue, C.; Zhou, X.; Karjadi, C.; Chang, G.H.; Joshi, A.S.; Dwyer, B.; Zhu, S.; et al. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain* **2020**, *143*, 1920–1933. [CrossRef]
17. Kadry, S.; Taniar, D.; Damasevicius, R.; Rajinikanth, V. Automated detection of schizophrenia from brain MRI slices using optimized deep-features. In Proceedings of the 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021, Chennai, India, 25–27 March 2021. [CrossRef]
18. Ramzan, F.; Khan, M.U.G.; Rehmat, A.; Iqbal, S.; Saba, T.; Rehman, A.; Mehmood, Z. A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer's Disease Stages Using Resting-State fMRI and Residual Neural Networks. *J. Med. Syst.* **2019**, *44*, 37. [CrossRef]
19. Amin-Naji, M.; Mahdavinataj, H.; Aghagolzadeh, A. Alzheimer's disease diagnosis from structural MRI using Siamese convolutional neural network. In Proceedings of the 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran, Iran, 6–7 March 2019; pp. 75–79. [CrossRef]
20. Jabason, E.; Ahmad, M.O.; Swamy, M.N.S. 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), Dallas, TX, USA, 4–7 August 2019; pp. 481–484. [CrossRef]
21. Kadry, S.; Damasevicius, R.; Taniar, D.; Rajinikanth, V.; Lawal, I.A. U-net supported segmentation of ischemic-stroke-lesion from brain MRI slices. In Proceedings of the 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021, Chennai, India, 25–27 March 2021. [CrossRef]

22. Maqsood, S.; Damasevicius, R.; Shah, F.M. An efficient approach for the detection of brain tumor using fuzzy logic and U-NET CNN classification. In Proceedings of the International Conference on Computational Science and Its Applications, ICCSA 2021, Cagliari, Italy, 13–16 September 2021; pp. 105–118. [CrossRef]

23. Odusami, M.; Maskeliunas, R.; Damaševičius, R.; Misra, S. Comparable study of pre-trained model on alzheimer disease classification. In Proceedings of the International Conference on Computational Science and Its Applications, ICCSA 2021, Cagliari, Italy, 13–16 September 2021; pp. 63–74. [CrossRef]

24. Lu, D.; Popuri, K.; Ding, G.W.; Balachandar, R.; Beg, M.F. Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease. *Med. Image Anal.* **2020**, *46*, 26–342. [CrossRef]

25. Bae, J.B.; Lee, S.; Jung, W.; Park, S.; Kim, W.; Oh, H.; Han, J.W.; Kim, G.E.; Kim, J.S.; Kim, J.H.; et al. Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Sci. Rep.* **2020**, *10*, 22252. [CrossRef]

26. Farooq, A.; Anwar, S.; Awais, M.; Rehman, S. A deep CNN based multi-class classification of Alzheimer's disease using MRI. In Proceedings of the 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, China, 18–20 October 2017; pp. 1–6. [CrossRef]

27. Ebrahim, D.; Ali-Eldin, A.M.T.; Moustafa, H.E.; Arafat, H. Alzheimer Disease Early Detection Using Convolutional Neural Networks. In Proceedings of the 2020 15th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 15–16 December 2020; pp. 1–6. [CrossRef]

28. Ebrahimi, A.; Luo, S.; Chiong, R. Deep sequence modelling for Alzheimer's disease detection using MRI. *Comput. Biol. Med.* **2021**, *134*, 104537. [CrossRef]

29. Odusami, M.; Maskeliūnas, R.; Damaševičius, R.; Krilavičius, T. Analysis of Features of Alzheimer's Disease: Detection of Early Stage from Functional Brain Changes in Magnetic Resonance Images Using a Finetuned ResNet18 Network. *Diagnostics* **2021**, *11*, 1071. [CrossRef]

30. Prakash, D.; Madusanka, N.; Bhattacharjee, S.; Kim, C.-H.; Park, H.-G.; Choi, H.-K. Diagnosing Alzheimer's Disease based on Multiclass MRI Scans using Transfer Learning Techniques. *Curr. Med. Imaging* **2021**, *17*, 1460–1472. [CrossRef]

31. Yadav, K.S.; Miyapuram, K.P. A Novel Approach Towards Early Detection of Alzheimer's Disease Using Deep Learning on Magnetic Resonance Images. In Proceedings of the International Conference on Brain Informatics, Padova, Italy, 17–19 September 2021; pp. 486–495. [CrossRef]

32. Buvaneswari, P.R.; Gayathri, R. Deep Learning-Based Segmentation in Classification of Alzheimer's Disease. *Arab. J. Sci. Eng.* **2021**, *46*, 5373–5383. [CrossRef]

33. Parmar, H.; Nutter, B.; Long, L.; Antani, S.; Mitra, S. Spatiotemporal feature extraction and classification of Alzheimer's disease using deep learning 3D-CNN for fMRI data. *J. Med. Imaging* **2020**, *7*, 056001. [CrossRef]

34. Ruiz, J.; Mahmud, M.; Modasshir, M. Shamim Kaiser, and for the Alzheimer's Disease Neuroimaging Initiative, 3D DenseNet Ensemble in 4-Way Classification of Alzheimer's Disease. In Proceedings of the International Conference on Brain Informatics, Padova, Italy, 19 September 2020; pp. 85–96. [CrossRef]

35. Solano-Rojas, B.; Villalón-Fonseca, R.; Marín-Raventós, G. Alzheimer's Disease Early Detection Using a Low Cost Three-Dimensional Densenet-121 Architecture. In *The Impact of Digital Technologies on Public Health in Developed and Developing Countries, Proceedings of the International Conference on Smart Homes and Health Telematics, Hammamet, Tunisia, 24–26 June 2020*; Springer International Publishing: Cham, Switzerland, 2020; pp. 3–15. [CrossRef]

36. Wang, H.; Shen, Y.; Wang, S.; Xiao, T.; Deng, L.; Wang, X.; Zhao, X. Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing* **2019**, *333*, 145–156. [CrossRef]

37. Naz, S.; Ashraf, A.; Zaib, A. Transfer learning using freeze features for Alzheimer neurological disorder detection using ADNI dataset. *Multimed. Syst.* **2021**, 1–10. [CrossRef]

38. Mehmood, A.; Yang, S.; Feng, Z.; Wang, M.; Ahmad, A.S.; Khan, R.; Maqsood, M.; Yaqub, M. A Transfer Learning Approach for Early Diagnosis of Alzheimer's Disease on MRI Images. *Neuroscience* **2021**, *460*, 43–52. [CrossRef]

39. Jain, R.; Jain, N.; Aggarwal, A.; Hemanth, D.J. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cogn. Syst. Res.* **2019**, *57*, 147–159. [CrossRef]

40. Kumar, A.; Dadheech, P.; Dogiwal, S.R.; Kumar, S.; Kumari, R. Medical Image Classification Algorithm Based on Weight Initialization-Sliding Window Fusion Convolutional Neural Network. In *Computer-Aided Design and Diagnosis Methods for Biomedical Applications*; CRC Press: Boca Raton, FL, USA, 2021.

41. Maqsood, M.; Nazir, F.; Khan, U.; Aadil, F.; Jamal, H.; Mehmood, I.; Song, O.-Y. Transfer Learning Assisted Classification and Detection of Alzheimer's Disease Stages Using 3D MRI Scans. *Sensors* **2019**, *19*, 2645. [CrossRef]

42. Acharya, H.; Mehta, R.; Singh, D.K. Alzheimer Disease Classification Using Transfer Learning. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021; pp. 1503–1508. [CrossRef]

43. Kazemi, Y.; Houghten, S. A deep learning pipeline to classify different stages of Alzheimer's disease from fMRI data. In Proceedings of the 2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), St. Louis, MO, USA, 30 May–2 June 2018; pp. 1–8. [CrossRef]

44. Shanmugam, J.V.; Duraisamy, B.; Simon, B.C.; Bhaskaran, P. Alzheimer's disease classification using pre-trained deep networks. *Biomed. Signal Processing Control* **2022**, *71*, 103217. [CrossRef]

45. Hazarika, R.A.; Kandar, D.; Maji, A.K. An experimental analysis of different Deep Learning based Models for Alzheimer's Disease classification using Brain Magnetic Resonance Images. *J. King Saud Univ.—Comput. Inf. Sci.* **2021**. [CrossRef]

46. Jmour, N.; Zayen, S.; Abdelkrim, A. Convolutional neural networks for image classification. In Proceedings of the 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), Hammamet, Tunisia, 22–25 March 2018; pp. 397–402. [CrossRef]

47. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]

48. Nguyen, L.D.; Lin, D.; Lin, Z.; Cao, J. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018; pp. 1–5. [CrossRef]

49. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385. Available online: http://arxiv.org/abs/1512.03385 (accessed on 20 April 2021).

50. Lyu, Z.; ElSaid, A.; Karns, J.; Mkaouer, M.; Desell, T. An Experimental Study of Weight Initialization and Lamarckian Inheritance on Neuroevolution. In *Applications of Evolutionary Computation*; Springer: Cham, Switzerland, 2021; pp. 584–600. [CrossRef]

51. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.

52. Ashraf, A.; Naz, S.; Shirazi, S.H.; Razzak, I.; Parsad, M. Deep transfer learning for alzheimer neurological disorder detection. *Multimed. Tools Appl.* **2021**, *80*, 30117–30142. [CrossRef]

53. Jin, H.; Song, Q.; Hu, X. Auto-keras: An efficient neural architecture search system. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, Anchorage, AK, USA, 4–8 August 2019.

54. Toniolo, S.; Serra, L.; Olivito, G.; Marra, C.; Bozzali, M.; Cercignani, M. Patterns of Cerebellar Gray Matter Atrophy Across Alzheimer's Disease Progression. *Front. Cell. Neurosci.* **2018**, *12*, 430. [CrossRef]

55. Todd, K.L.; Brighton, T.; Norton, E.; Schick, S.; Elkins, W.; Pletnikova, O.; Fortinsky, R.H.; Troncoso, J.C.; Molfese, P.J.; Resnick, S.M.; et al. Ventricular and Periventricular Anomalies in the Aging and Cognitively Impaired Brain. *Front. Aging Neurosci.* **2018**, *9*, 445. [CrossRef] [PubMed]

56. Femir-Gurtuna, B.; Gürtuna, B.F.; Kurt, E.; Ulasoglu-Yildiz, C.; Bayram, A.; Yildirim, E.; Büyükişcan, E.S.; Bilgic, B. White-matter changes in early and late stages of mild cognitive impairment. *J. Clin. Neurosci.* **2020**, *78*, 181–184. [CrossRef] [PubMed]

57. Houck, A.L.; Gutierrez, J.; Gao, F.; Igwe, K.; Colon, J.; Black, S.; Brickman, A. Increased Diameters of the Internal Cerebral Veins and the Basal Veins of Rosenthal Are Associated with White Matter Hyperintensity Volume. *Am. J. Neuroradiol.* **2019**, *40*, 1712–1718. [CrossRef]

58. 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]