

Article

Alzheimer Disease Classification through Transfer Learning Approach

Noman Raza ¹, Asma Naseer ¹, Maria Tamoor ²  and Kashif Zafar ^{1,*} 

¹ Department of Computer Science, National University of Computer and Emerging Sciences, Lahore 54770, Pakistan

² Department of Computer Science, Forman Christian College, Lahore 54600, Pakistan

* Correspondence: kashif.zafar@nu.edu.pk

Abstract: Alzheimer’s disease (AD) is a slow neurological disorder that destroys the thought process, and consciousness, of a human. It directly affects the development of mental ability and neurocognitive functionality. The number of patients with Alzheimer’s disease is increasing day by day, especially in old aged people, who are above 60 years of age, and, gradually, it becomes cause of their death. In this research, we discuss the segmentation and classification of the Magnetic resonance imaging (MRI) of Alzheimer’s disease, through the concept of transfer learning and customizing of the convolutional neural network (CNN) by specifically using images that are segmented by the Gray Matter (GM) of the brain. Instead of training and computing the proposed model accuracy from the start, we used a pre-trained deep learning model as our base model, and, after that, transfer learning was applied. The accuracy of the proposed model was tested over a different number of epochs, 10, 25, and 50. The overall accuracy of the proposed model was 97.84%.

Keywords: gray matter; convolutional neural network; Alzheimer’s disease classification; dense-net



Citation: Raza, N.; Naseer, A.; Tamoor, M.; Zafar, K. Alzheimer Disease Classification through Transfer Learning Approach. *Diagnostics* **2023**, *13*, 801. <https://doi.org/10.3390/diagnostics13040801>

Academic Editors: Jaafar M. Alghazo and Ghazanfar Latif

Received: 1 December 2022

Revised: 13 January 2023

Accepted: 18 January 2023

Published: 20 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Alzheimer’s disease (AD) is a slow neurological disorder that destroys the thought process, and consciousness, of a human. It directly affects the development of mental ability and neurocognitive functionality [1]. AD is more common among older people, and sometimes it becomes the primary cause of their deaths [2,3]. People who have other medical problems at later age, such as diabetes, cardio-problems, and hypertension [1], are at higher risk of developing Alzheimer’s disease. There is no complete treatment for AD disease. However, early detection of this disease can help in taking preventative action at an early stage and can improve the symptoms of AD [4]. It is a primary reason for dementia in the elderly, due to the destruction of neurons related to the human memory [4], and to questioning, and learning functions [5]. This neurological disorder starts with slow deterioration and symptoms worsen with each passing day [6]. Memory issues are generally one of the first symptoms of AD [2], although preliminary symptoms may also vary from individual to individual [3]. Changes in different factors of thinking, inclusive of locating the proper words, spatial issues, and impaired judgment, might also become the early symptoms of AD [7]. The initial stage of mild cognitive impairment (MCI) is an early signal that a human may be an AD patient [8]. The number of AD patients is increasing day by day and it is expected that one out of eighty-five humans could be suffering from Alzheimer’s disease by 2050 [9,10].

The progression of Alzheimer’s disease can be diagnosed using clinical measures, but it is a very time-consuming process and expert persons are required to detect the symptoms [11,12]. Early diagnosis is very difficult to assess by an expert person, unless symptoms become very obvious. Early detection of AD can assist in lowering the risk of neuron disorder [13,14]. Early diagnosis can ensure that the patient is aware of the need to take precautionary measures to lower the risk of advancement of the disease from MCI to

AD [1]. In recent times, different machine learning and deep learning methods have been proposed for the prediction of the stages of AD through self-regulating analysis of magnetic resonance imaging (MRI) images, which provide efficient and improved diagnosis results for AD [14–16]. The major factors, or parameters, that researchers use are the cortical thickness of the human brain, gray matter (GM) density in the brain, ventricle expansions, and brain shrivel. Many research studies claim that a correlation exists between grey matter reduction and certain brain diseases, like Alzheimer’s disease [13]. The hippocampus is the part of the brain that is affected at the initial stage of Alzheimer’s disease. White matter (WM), gray matter (GM), and cerebrospinal liquid are the major and most primitive tissues in human brain images. Out of these three fundamental tissues of the brain, researchers have discovered that GM shrivel corresponds more with corporeal diminishment in mild cognitive impairment [17].

Different statistical and machine learning methods, such as Support vector machine (SVM) [18] are used for automated recognition of Alzheimer’s disease. Lately, deep learning techniques, such as CNN and sparse auto-encoders, have surpassed the SVM learning techniques [11]. There are some limitations to using deep learning methods. The training from scratch of deep learning models requires a lot of computation and a large dataset of annotated medical images [19,20]. The collection of such a massive amount of medical images is a bottleneck for researchers. To overcome issues of large dataset requirements and computational issues associated with complex methods, the transfer learning approach can be used for different medical imaging modalities. Transfer learning is very useful, even for cross-domain applications, such as the fact that a model trained on images of natural things could be used for medical images, like X-rays, CT scans, MRIs, and many others [21]. The idea behind transfer learning is to utilize a pre-trained model on a new, small image, dataset of a different nature [22,23]. Transfer learning is an emerging deep learning technique in which an architecture devised for one task is reused as the initial point for a second task. The important motive behind transfer learning is the saving of knowledge attained while resolving a specific dilemma, followed by applying that gained knowledge to solve different problems.

In this research, a CNN-based transfer learning technique is used for the classification of neuro-medical AD scans into the following four categories: Alzheimer’s disease (AD), late mild cognitive impairment (LMCI), mild cognitive impairment (MCI), and normal cognition (NC). The elemental motivation for using transfer learning was to forward the features extracted from natural brain images to AD images, and to investigate a new approach for the categorization of AD that could be helpful for clinicians in proper diagnosis and decision-making. In this way, the patient can be guided in taking precautionary measures to lower the risk of MCI progressing to AD. The main objective is to produce improved results, even when using a small-sized dataset. Transfer learning gives us the flexibility to drastically reduce training time by reusing pre-trained models for new data and for obtaining better results without over-fitting. The proposed model uses a deep learning architecture as a base model where the last few convolutional layers are retained and then merged with fully connected layers. By applying the transfer learning approach with retraining of the last two layers, promising results are obtained on multi-class categorizations, such as AD, LMCI, EMCI, MCI, and NC. This research is primarily based on two-dimensional Gray matter (GM) images of the human brain, that are extracted from raw Magnetic resonance images (MRIs), these being more helpful in the detection of the early stage of AD. As illustrated in Figure 1, the MRI scans of the patients are converted to 2D slices in the pre-processing step, where the GM slices are focused for the extraction and training of the proposed model.

The major research contributions are mentioned below:

- A customized convolutional neural network with transfer learning is proposed for the classification of Alzheimer’s disease.
- A new corpus, consisting of four different types of AD, is developed. Each type consists of 1254 images.

- Extraction of 2D GM slices, using “SPM12”, which is very familiar in medical image pre-processing.
- Higher accuracy, 97.84%, with a lower number of epochs, is achieved for the multi-class classification of AD.



Figure 1. MRI to 2D GM slice.

The rest of the paper is organized in such a way that Section 2 describes the background of the techniques used in this research. Section 3 provides a detailed view of the related work, in which some related techniques of AD classifications are discussed. Section 4 describes the methodology and details of the experimental results. Section 5 elaborates the discussions, while section 6 contains the conclusion of the research and future directions.

2. Background

Early diagnosis of Alzheimer’s disease can assist in lowering the peculiar disorder of the neurons. In this research, a model that takes a 2D GM slice as input to classify the stage of AD is discussed. A brief introduction for each of the basic concepts used in this research is provided below.

2.1. Convolutional Neural Network

Among many algorithms that enable a machine or a computer to train, perceive and learn for different medical classification problems, one is the convolutional neural network (CNN) [11,15,16]. CNN has revolutionized machine learning and artificial intelligence domains to make them akin to human brains. One of the important features of CNN is its characteristics of learning the features of images with training. In the human brain, the neurons are attached to each other in a specific pattern. CNN models are designed in the same manner. CNN models are multi-layered structures that perform in a group. The design of CNN comprises the following four kinds of layers [24]:

- Convolution
- Pooling
- Fully Connected
- Softmax

2.2. Transfer Learning

Transfer learning is an emerging deep learning technique in which an architecture trained for one task is reused as the initial point for a second task. The basic motivation behind this concept is to reduce training time, and to overcome issues associated with the need for large training sets. Pre-trained models are used for new problems without compromising the performance results.

2.3. Gray Matter

Gray matter and white matter are the two major tissues of which the central nervous system is composed [2,3]. Gray matter is the main part of the brain that is used to process information sent in the form of signals by the sensory organs of the human body.

2.4. DenseNet

DenseNet is an efficient variant of a convolutional neural network, in which each layer has extra inputs from all previous layers and forwards its very own characteristic maps to all the next layers. Due to concatenation, each layer in the DenseNet obtains complete and comprehensive knowledge from all previous layers. In this architecture, each layer obtains extra inputs from all preceding layers and passes on its feature-maps to all subsequent layers, as illustrated in Figure 2 [25]. Instead of summation, as applies to ResNet, layers are combined using concatenation. Due to this layer-wise propagation of knowledge, DenseNet can be thinner and have better computational performance [26].

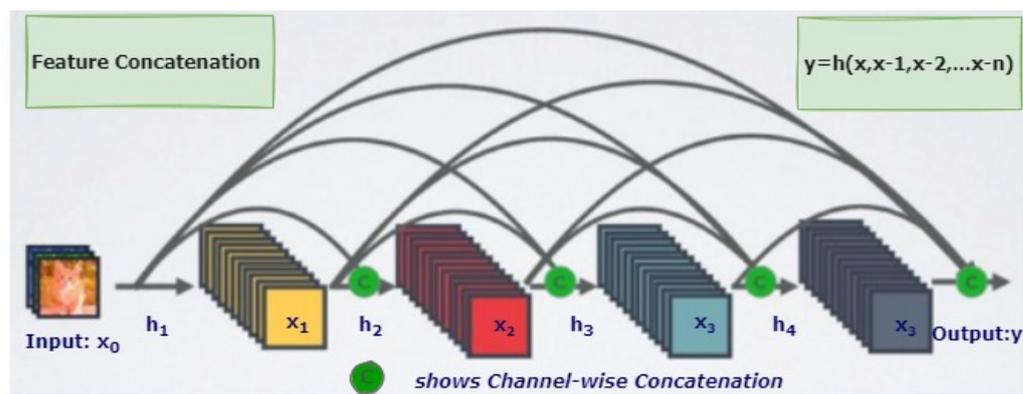


Figure 2. One Dense Block in DenseNet [26].

3. Related Work

In the paper, [27] a transfer learning approach was introduced by using the two most famous deep CNN architectures (Inception and VGG16) with the already trained and fine-tuned weights of ImageNet data. Using a pre-trained model on ImageNet, the researchers trained the last fully connected layer with a small number of training MRI scans. To overcome the over-fitting of the small training dataset, image entropy was applied to MRI images, to extract the most informative portions. An OASIS cross-sectional dataset with 416 subjects was used in an experiment aimed at the binary classification of AD. Five-fold cross-validation was applied with an 80 percent and 20 percent split between training and testing in the fully connected layer retraining. To compare the results VGG16 was also trained from scratch, as well as with transfer learning. Due to the small training set, the VGG16 trained from scratch performed less well in terms of accuracy, 74.12%, while the VGG16 with transfer learning provided 92.3% accuracy. Finally, Inception V4 was used with transfer learning that provided promising results with 96.25% accuracy.

In the paper [28] CNN with LeNet-5 was utilized for the classification of the brain with AD and the normal brain, by using functional MRI 4D data. In the first step, the 4D data was transformed into 2D by using the neuroimaging packages Nibabel and OpenCV. Then, 2D images were labeled as AD vs NC. The LeNet model, based on CNN, was then used for the binary classification of the images. The results were compared with the famous support vector machine model and, in contrast to it, the proposed model provided better results, with 96.86% accuracy.

In the paper [29] a framework with the combination of sparse auto-encoders (SAEs) and a softmax logistic regression was used, along with autoencoders, to use unlabeled data. Two data sets, MR and positron emission tomography (PET) from the ADNI database, were used. The main target of this research was to use SAE for high-level feature selection in the unsupervised pre-training stage. As a result of two different neuroimaging modalities, a zero-masking technique was used for the extraction of complimentary details from these different datasets. Features extracted from SAE, using unsupervised data, were then manipulated with a softmax regression. The performance of the model was tested on the

classification of AD. In comparison with other advanced models, like SVM and other deep learning methods, the proposed model performed very well with 91.4 percent accuracy just because of its capability to extract features in one setting and its requiring of less labeled data.

In the paper, [30], a customized 2D-CNN model, with 9 depth-wise separable convolutional and normalization layers, was used, along with Inception V3 and Xception models for transfer learning. In this research, the classification of AD patients, class imbalance, and data leakage issues were discussed. The second fully connected layer used the sigmoid function as an activation function to categorize the data into two classes. An OASIS dataset, with T1-weighted structural MRI images, was used and the dataset was divided into 3 portions: set 1, set 2, and set 3. Cross-validations with 2-folds, 5-folds, and 9-folds were applied to the partitioned datasets, respectively. Dataset 1 was used for the prediction of AD, dataset 2 was used for class imbalance and dataset 3 was used for data leakage problems. For the AD classification on dataset 1, 45 subjects were used for training and validation purposes. Stochastic gradient descent (SGD) was used as an optimization algorithm. For loss function, binary cross entropy was used. In comparison with other deep learning models, the proposed model, which was based on transfer learning, provided promising results.

In another paper, [31], a deep learning framework with a softmax output layer and stacked autoencoders are used for the detection of Alzheimer's disease and its initial stage MCI. MRI data of 311 patients available on the ADNI database was used. Gray matter (GM) was extracted from the MRI images, which made the baseline for the detection of MCI and the CMRGlC patterns using PET. Elastic Net is then used to extract the high-level features. In individual cross-fold, 90 percent of subjects are used for training and the rest of the 10 percent for testing. SK-SVM and MK-SVM are considered for comparison with the proposed model. The model gives 87.76% accuracy in the binary classification of AD patients.

In another piece of research, [32] AlexNet, a fine-tuned pre-trained CNN, was used for the binary and multi-class classification of 3D MRI images. The proposed model was trained on the already pre-processed data in which WM, GM, and CSF were segmented and, then, the testing of the model was conducted on the unsegmented 3D MRI scans of the human brain. An OASIS dataset, consisting of 382 subjects, was used for training and testing. After the training of the proposed model on the segmented dataset, the retrained convolutional neural network was then used for the validation over the unsegmented 3D MRI images. For multi-class classification, the proposed model outperformed the binary classification, with 92.8% accuracy versus 89.6%.

In another paper [33], a modified Siamese CNN model, inspired by Oxford Net (VGG16), was used for the classification of AD stages. The basic idea behind the proposed model was to use the augmentation technique, with an extra convolutional layer in VGG16. Augmentation was applied to an OASIS dataset after the pre-processing phase. Two parallel layers of modified VGG16 worked for the extraction of the most important features. Batch normalization was applied to increase the learning rate, which gradually decreased, due to changing the parameter in individual layers of the CNN model. In comparison with the other state-of-the-art models, the proposed model provided 99.05% accuracy, and it also reduced the problems of over-fitting and regularization. In [34], a layer-wise transfer learning approach and tissue segmentation were used for the classification of AD. The dataset used in this research was collected from the ADNI database. In the pre-processing step, the skull stripping, and extraction of GM, WM, and CSF were conducted using SPM12. The VGG-19 network was customized by modifying the last two fully connected and classification layers. Instead of freezing the trained fully connected layers, the researchers divided the model into two groups and then they gradually fixed CNN layers in different blocks. The training of the proposed model was done on both augmented and non-augmented datasets. In the first group, 8 CNN layers with 3 max-pooling layers were kept fixed, and in the second group 12 CNN layers along with 4 max-pooling layers

were kept fixed. In experiments after the augmentation, the classification results of the proposed model were 98.73%, 83.72%, and 80% on AD vs NC, EMCI vs LMCI, and other classes, respectively.

In another paper [14], a cross-model technique, using the transfer learning technique, was used to reduce the over-fitting problem, while the training was done on a small set of MRI images. The proposed model was trained on the structural MRI data collected from the ADNI database and then tested on the DTI dataset. The outcome of the model on the two different cross-modalities was outstanding, with 92 percent accuracy on NC vs AD, 80 percent on NC vs MCI, and 85 percent on MCI vs AD.

In [16], the deep learning models, GoogLeNet and ResNet, were trained from scratch on structural MRI data sets available on the ADNI database. The main target of this research was to segment the gray matter (GM) and then train the CNN networks on these segmented GM images. The addition of the augmentation layer proved to be a useful step in the classification of the four stages of AD.

In another research [35], a convolutional network, with some extra parameters like gender, Mental-state exam score, and age, were trained on different datasets. The first data set was composed of clinically diagnosed Alzheimer's patients and the second data set was extracted from the ADNI. The validation of the model was done on three different datasets, which included Australian Imaging, Biomarker and Lifestyle Flagship Study of Ageing, and the National Alzheimer's Coordinating Center. The outcome of the model on multi-modal datasets was good in comparison to the other CNN models.

In a paper [36], ResNet18 is used for the classification of AD stages. The main purpose of this research was to utilize the Resting-state fMRI data, extracted from the ADNI database, which is a very useful neuroimaging technology used for the observation of neurodegenerative diseases. The concept of transfer learning was applied to the convolutional model, which was trained from scratch. The results of the proposed model were extracted with and without augmentations. They were also compared with other advanced CNN models. The outcomes of the proposed model showed promising results in the classification of AD stages.

4. Methodology

In this research, the segmentation of Alzheimer's MRI images and their multi-class classification through transfer learning is proposed. For Alzheimer's detection through MRI, Convolution Neural Network (CNN) was customized by using the specific GM-segmented images of the brain. Instead of training the model from scratch, a pre-trained deep learning model, Densenet-169, was used as a base model. Gradually, transfer learning was applied to this base model for Alzheimer's detection. An overview of the proposed methodology is illustrated in Figure 3 which shows how MRI images were pre-processed and GM slices extracted. These GM slices were used for training the pre-trained deep learner and, finally, multi-class classification was performed.

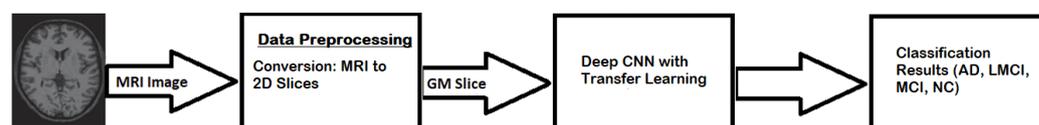


Figure 3. Proposed model workflow.

4.1. DataSet

The dataset [34], used in this research was acquired from the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) database [37]. In 2004, this dataset was collected to analyze multiple sensitive techniques on various scanned images, like PET scan, sMRI, MRI and fMRI, etc. To evaluate the initial stages of Alzheimer's disease [5] this dataset is considered a benchmark dataset. It contains images of four categories i.e.,

Alzheimer's Disease (AD), Non-Cognitive (NC), Late Mild Cognitive Impairment (LMCI), and Mild Cognitive Impairment (MCI). Each category has 1254 images in total.

4.2. Data Preprocessing

In preprocessing, a series of steps were applied to the T1 weighted MRI images extracted from the publicly available database ADNI. For pre-processing, Statistical Parametric Mapping 12 (SPM12) was used which is quite often used in practice for medical and natural images. All the MRI images which were in Neuroimaging Informatics Technology Initiative (NIFTI) format were pre-processed to identify and extract the Gray Matter (GM) slices. The proposed model was trained on GM to detect and examine the initial changes in AD patients. The major steps of pre-processing encompassed segmentation, skull stripping, spatial normalization, and re-scaling. All the major steps, performed in the pre-processing of MRI images, are illustrated in Figure 4.

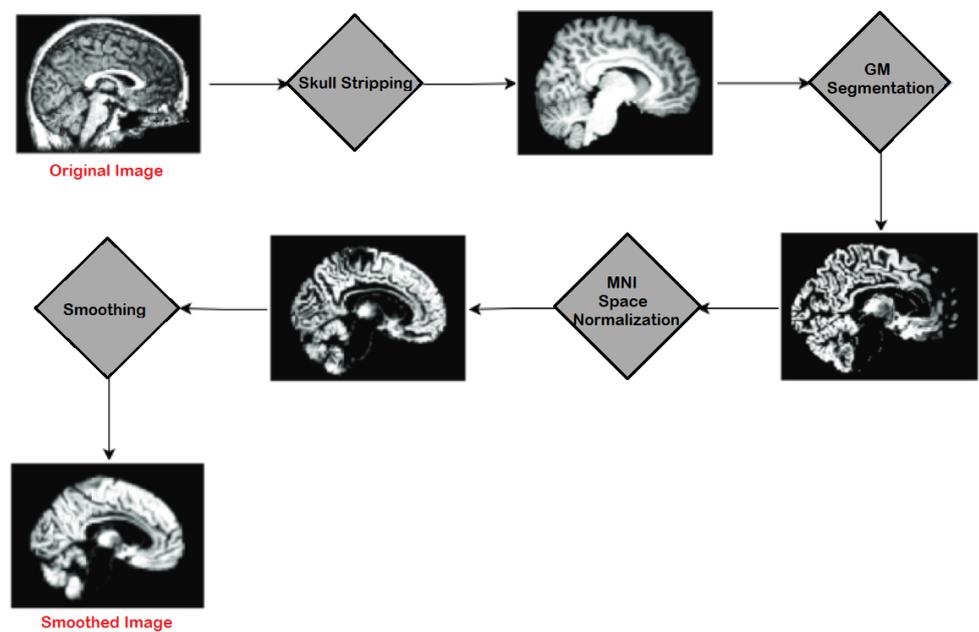


Figure 4. Major Steps of Pre-processing on MRIs.

4.2.1. Skull Stripping

For the morphometric analysis and study of brain MRIs, a preliminary process of skull stripping was performed, in which the tissues of the brain, i.e., cerebellum and cortex, were segmented from the surrounding zone consisting of non-brain and skull area. Among the multiple automated skull stripping approaches [38], the most commonly used technique, histogram thresholding, was followed by the employment of certain morphological processes [39]. The morphological procedures included erosion, and were based on anisotropic filters. The Snake method of contouring was also applied, to remove the eyes. The gradual process of skull stripping involved background and noise removal, followed by contour identification of the brain and, finally, refinement of the brain contour.

Head segmentation required noise removal, which was performed with the help of a threshold level, considering that the maximum noise was generated by Rayleigh probability distribution. The noise was removed by applying Equation (1):

$$Rnoise(f) = \frac{f}{\sigma^2} \exp\left(\frac{-f^2}{2\sigma^2}\right) \quad (1)$$

where f is the intensity of noise and σ is standard deviation of Rayleigh noise.

After that, the best-fit Rayleigh curve was subtracted from the histogram volume, as given in Equation (2):

$$G(f) = h(f) - r(f) \quad (2)$$

where $h(f)$ is histogram volume and $r(f)$ is the best-fit curve. The value for $r(f)$ was determined by minimizing error as given in Equation (3):

$$\varepsilon_\tau = \sum_{f=0}^{\tau-1} g(f) + \sum_{f=\tau}^{\infty} r(f) \quad (3)$$

The said procedure produced speckle, which was outside the head region. This error was removed by using morphological operations in a 5×5 kernel.

The initial brain mask was generated in three steps. The first one was the smoothing of the brain and attenuation of non-brain regions for which anisotropic diffusion, given in Equation (4), was used in a nonlinear way:

$$\frac{\partial}{\partial t} M(x, t) = \nabla \cdot (C(x, t) \nabla M(x, t)) \quad (4)$$

where M is MRI, x represents its coordinates, t is the iteration step and $C(x, t)$ is the function for diffusion, given in Equation (5).

Next, the MRI volume was defused with the help of automatic thresholding, and a diffusion function, given in Equation (5), was used to produce a binary mask:

$$C(x, t) = \exp\left(-\left(\frac{|\nabla M(x, t)|}{\sqrt{2}f}\right)^2\right) \quad (5)$$

where f was constant for diffusion, which was set as 128.

Finally, non-brain regions were removed by using spatial and morphological information obtained from the head mask. All the holes within the mask region were filled by applying binary erosion with kernel size 10×10 . As a result, the kernel separated the eyes from the brain. After erosion, all the regions where the centroid was outside the bounding box were eliminated, which was decided after applying heuristics. After erosion, binary dilation was applied with the same size as the kernel that was used for erosion. This helped in recovering the darkest pixels which were eliminated from the edge of the brain, due to threshold.

Two sample images before and after skull stripping are given in Figure 5 for axial and sagittal planes.

4.2.2. Segmentation

Segmentation is one of the major steps of pre-processing. For the diagnosis of certain brain disorders, segmenting brain tissues in Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) makes the diagnosis process efficient.

The human brain MRI needs to be segmented into three main parts i.e., WM, GM, and CSF [40]. MRIs are mainly generated with the help of T1 and T2 weighted scan sequences. The main differences between the two are in Repetition Time (TR) and Time to Echo (TE), which is shorter for T1 and longer for T2. The contrast and brightness of MRIs are determined by T1 and T2, which affect the intensity levels of GM, WM, and SCF. For segmenting brain MRIs in GM, WM and SCF intensity probability distribution (IPD) was used, along with bias regularization, which was set to an extremely small regularization value i.e., 0.0001. Cutoff of 60 mm was used as Full Width at Half Maximum (FWHM), which was considered the width of a line structure at half of its highest amplitude. Figure 6 illustrates the results of segmenting an original brain MRI into GM, WM and SCF.

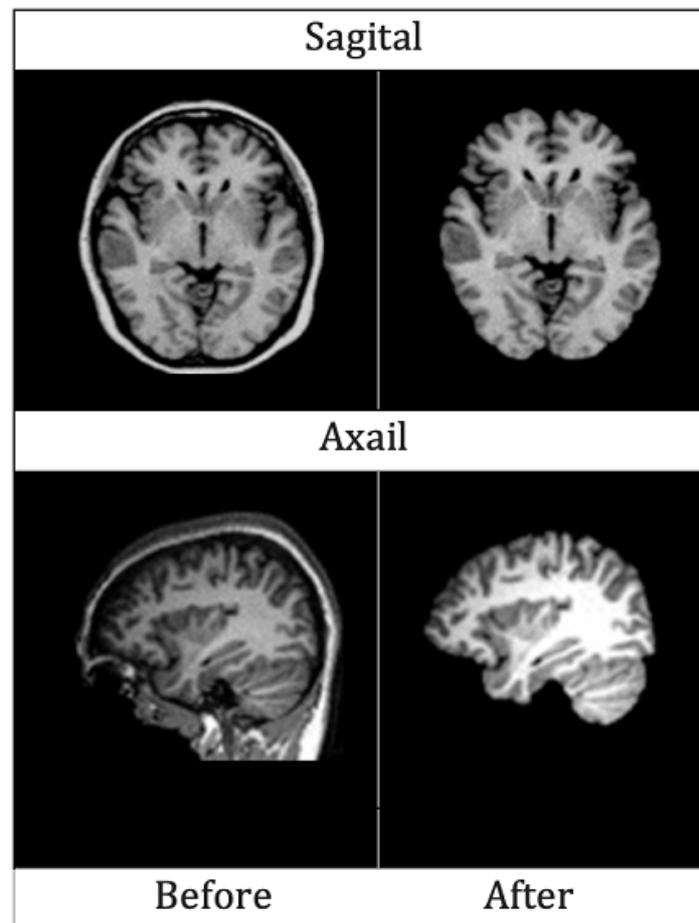


Figure 5. MRIs of Axial and Sagittal planes before and after skull stripping.

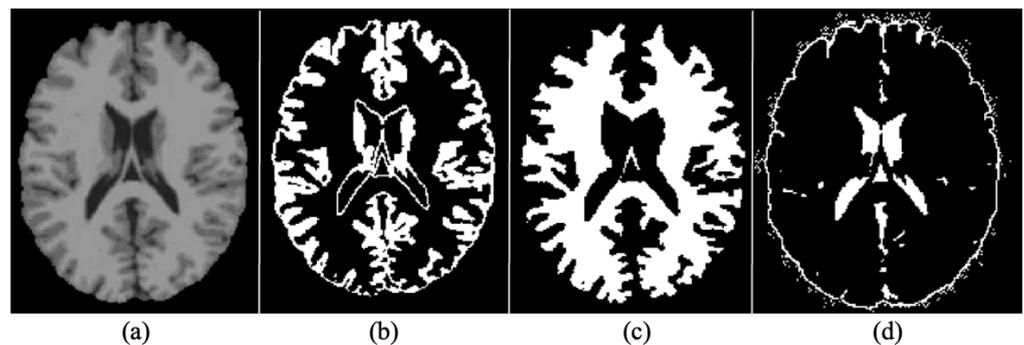


Figure 6. (a) Original (b) Grey Matter (GM) (c) White Matter (WM) (d) Cerebrospinal Fluid.

4.2.3. Normalization

Spatial normalization is a process of mapping images containing gray matter to the same reference space. In this research, Voxel-based morphometry [41] was employed for spatial normalization. In Neuroanatomy, Voxel-based morphometry measures the dissimilarity between local tissue clusters of the brain and several other brain images. The differences are calculated via a voxel-wise comparison between the two brain images and the images are mapped to a reference brain space. For this purpose, the Montreal Neurological Institute (MNI) space was used to identify the boundaries around the brain.

4.2.4. Rescaling

After the segmentation of raw MRI images, 256×240 shaped data samples were extracted. All these images were then re-scaled to 224×224 , that were, finally, used

to train, test, and evaluate the proposed model. The re-scaling adjusted numbers for comparing values even if they were out of scope. Furthermore, it reduced the number of trainable parameters of deep learners and they trained faster.

4.2.5. Smoothing

After rescaling, MRIs were smoothed by applying a 3D Gaussian kernel with voxel size (2 2 2). Each voxel was updated by the weighted mean value, attained from the adjacent voxels. The weight was decided based on the Gaussian shape that surrounded the voxel. The variance or width of the Gaussian curve determined how much smoothing was applied. With higher variance, more neighboring values were considered in calculating the mean value, which was expressed as FWHM. Figure 7 illustrates the Gaussian Kernel.

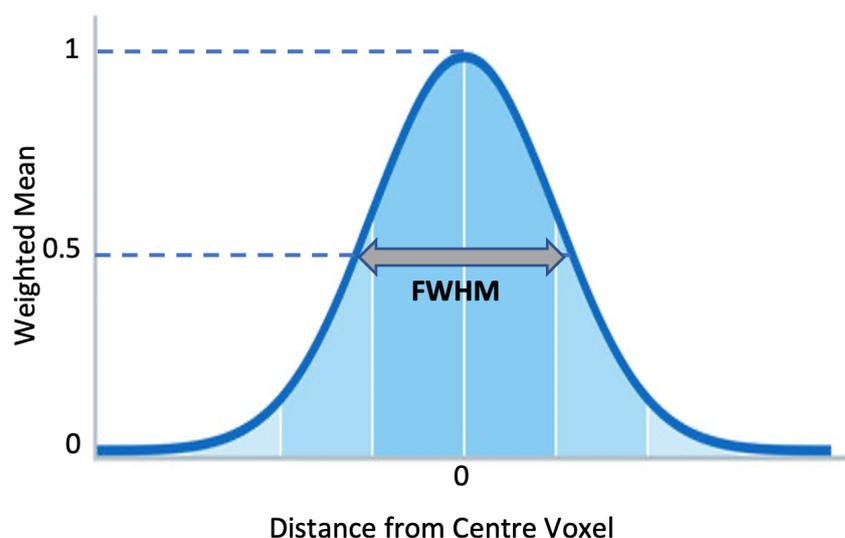


Figure 7. Gaussian Smoothing Kernel.

4.2.6. Augmentation

The last step of data pre-processing was augmentation which was adopted to avoid over-fitting. For generating augmented MRIs, five augmentation techniques, i.e., rescaling, rotation, zooming, horizontal flip, and vertical flip, were applied. In re-scaling, images were re-scaled to a range between 0 and 1 by multiplying each cell with $\frac{1}{255}$. The MRIs were rotated at 30° in an anti-clockwise manner. All the MRIs were also zoomed in with a ratio of 0.2. Finally, vertical and horizontal flips were applied and, in this way, five images in total were generated from one pre-processed MRI. Figure 8 illustrates the results of augmentation.

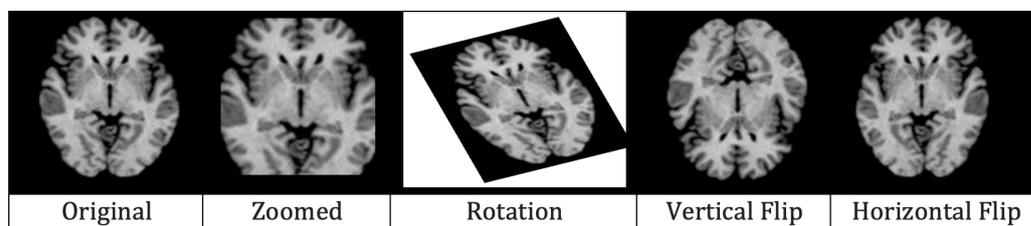


Figure 8. Effects of augmentation on MRI.

4.3. Architecture

After pre-processing, the pre-processed images were provided to the proposed Convolutional Neural Network, shown in Figure 9. The proposed CNN model consisted of a 7×7 convolution layers, 3×3 max-pooling layers, four dense blocks with a different number of convolution layers (6, 12, 32, 32) in respective blocks, three transitional layers,

and a fully connected classification layer. Three major components, i.e., the convolution layer, dense blocks, and transition layers, were used to extract the features from the pre-processed MRIs. The convolution layer with a kernel matrix of size 7×7 and a stride value set as 2, was used for the extraction of macro-features [42]. These extracted macro-level features were then forwarded to dense blocks, where more features were extracted. Each dense block contained 1×1 convolutions, which were used to reduce the number of input feature maps, followed by 3×3 convolution layers. In each dense block, the extracted features were passed to all the next layers. Between the dense blocks, there were transition layers, along with the activation functions, like batch normalization and ReLu, etc. The transition layer contained 1×1 convolution and a 2×2 max pooling layer with a stride value set as 2. Once the transition layer reduced the dimensions of the features, they were forwarded to the fully connected classification layer where Softmax was used for the classification. Initially, all the initial blocks of layers were kept frozen and the last two blocks were retrained.

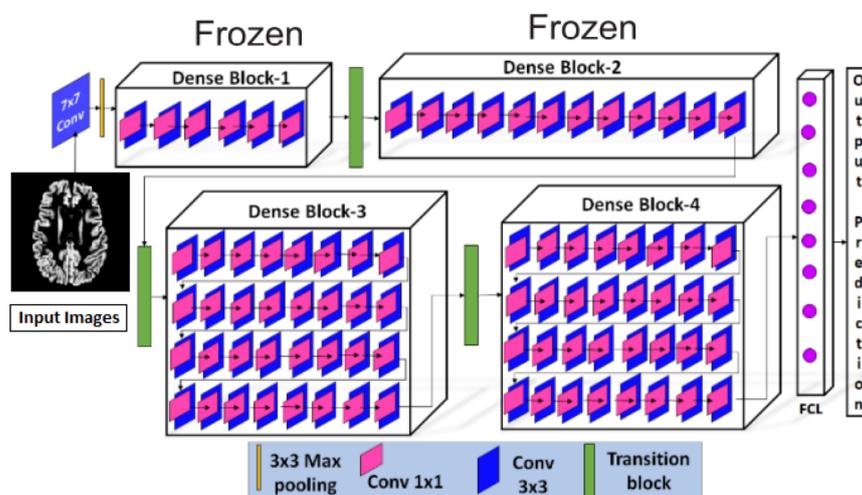


Figure 9. Proposed Model Architecture.

The details of all hyperparameters are listed in Table 1.

Table 1. Hyperparameters of the proposed model.

HYPERPARAMETERS	
Activation Function	ReLU
Epochs	50
Batch Size	128
Optimizer	Adam
Loss Function	Categorical Cross Entropy
Drop out	0.4

The concept of dense connectivity used in this research is represented in Equation (6):

$$c_l = H_l[(c_0, c_1, c_2, \dots, c_{l-1})] \tag{6}$$

where H_l is the non-linear transformation of L^{th} layer and $[c_0, c_1, c_2, c_3, \dots]$ are the feature maps forwarded by all the previous layers to the L^{th} layer.

The growth rate at the L^{th} layer, which revealed how much information was added by the previous layers, was calculated via Equation (7):

$$g_l = [g_0 + g \times (l - 1)] \tag{7}$$

where the hyper parameter g is the growth rate.

Convolutional layers were used to apply the filter on the input images to extract the feature maps. Once the feature maps were extracted by applying different filters, they were passed through the ReLU activation function. To obtain the non-linear input to the unit M_{ij}^l , all the previous layers' benefaction were summed up, as described in Equation (8):

$$M_{ij}^l = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \mu ab Y_{(i+a)(j+b)}^{l-1} \tag{8}$$

To reduce the dimensionality of the extracted feature maps, the max pooling layer was used. Considering the dimensions of feature maps as $[m_h, m_w, m_c]$, representing height, width, and channels of the feature maps, the dimensions were reduced by applying Equation (9) and resultant feature maps were achieved.

$$max_p = \frac{(m_h - f + 1)}{s} \times \frac{(m_w - f + 1)}{s} \times m_c \tag{9}$$

where f represents the size of the filter and s represents the strides.

In the dense layers, a neuron received the input from all the neurons of its preceding layer to perform the matrix multiplication. The standard equation for the matrix multiplication in the dense layer is shown in Equation (10):

$$N.\lambda = \begin{pmatrix} n_{11} & n_{12} & \dots & n_{1y} & p_1 \\ n_{21} & n_{22} & \dots & n_{2n} & p_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ n_{x1} & n_{x2} & \dots & n_{xy} & p_y \end{pmatrix} \tag{10}$$

where the N represents matrix dimensions of $x \times y$ and another matrix P with dimensions $1 \times y$ and λ is the matrix of the trained parameters of the previous layer.

The λ , representing the matrix of the trained parameters of the previous layer, was updated by using the back-propagation in the training phase. The back-propagation was used to adjust the weights w_k associated with a layer k over the learning rate η , as shown in Equation (11):

$$w^k = w^k - \eta \times dw^k \tag{11}$$

where dw is the partial derivatives of the loss function.

The partial derivative dw of the loss function of w was acquired by employing Equation (12):

$$dw^k = \frac{\partial L}{\partial w^k} = \frac{1}{n} dZ^k A^{[k-1]T} \tag{12}$$

where Z^k is the activation function and A^k is the non-linear activation function at layer k [43].

Figure 10 illustrates all the trainable and untrainable parameters of the proposed model, DenseNet, along with the output shape at each layer.

Layer (type)	Output Shape	Param #
densenet169 (Functional)	(None, 7, 7, 1664)	12642880
dropout (Dropout)	(None, 7, 7, 1664)	0
flatten (Flatten)	(None, 81536)	0
batch_normalization (BatchNo	(None, 81536)	326144
dense (Dense)	(None, 2048)	166987776
batch_normalization_1 (Batch	(None, 2048)	8192
activation (Activation)	(None, 2048)	0
dropout_1 (Dropout)	(None, 2048)	0
dense_1 (Dense)	(None, 1024)	2098176
batch_normalization_2 (Batch	(None, 1024)	4096
activation_1 (Activation)	(None, 1024)	0
dropout_2 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 4)	4100
Total params: 182,071,364		
Trainable params: 175,172,868		
Non-trainable params: 6,898,496		

Figure 10. Output size and model parameters at each layer of the proposed DenseNet.

4.4. Experimental Setup and Training

For the performance evaluation for Alzheimer’s classification, four transfer learning-based models were trained with the same architecture, as illustrated in Figure 9. The first model was trained without using the feature of transfer learning and was provided with the images without augmentation. The second one was again trained without the feature of transfer learning but was provided with augmented MRIs. The next two models were used with the feature of transfer learning and were trained on only original MRIs and augmented MRIs, respectively.

Gray matter (GM) images, as shown in Figure 11, were provided to the models for training. Three views of these MRIs, i.e., axial, coronal, and sagittal, were considered for training the models and, in total, 1254 images from each category, i.e., AD, NC, MCI, and LMCI, were used. In two steps, the data were divided into training, validation, and testing sets. In the first step, 20% of the data were split and separated for testing. The remaining 80% of the data were further divided into training and validation sets, as explained in Figure 12.

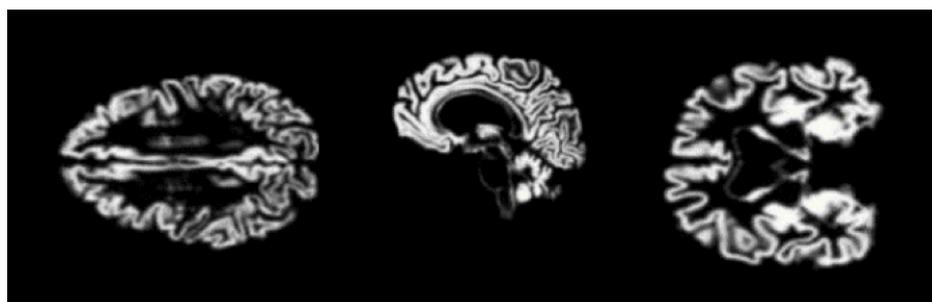


Figure 11. Gray matter images for proposed model.

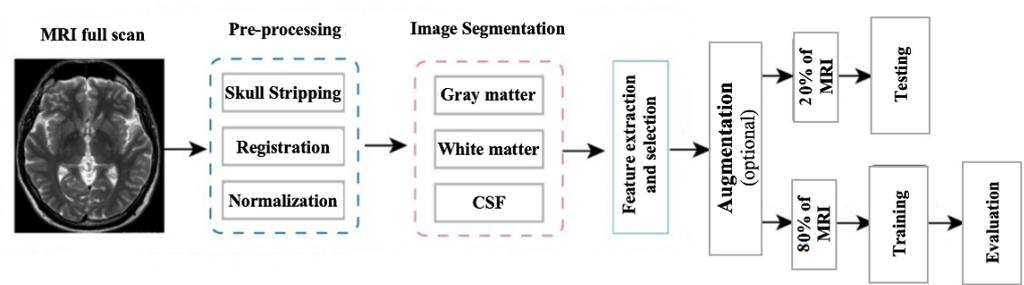


Figure 12. Framework of the proposed methodology.

During the training of the model, the number of epochs was gradually increased to verify the training accuracy and validation loss. The gradual increase in training accuracy and reduction in training loss with a different number of epochs is illustrated in Table 2.

Table 2. Training accuracy and validation loss for transfer learning-based model, trained on augmented MRIs.

	Epochs 10	Epochs 25	Epochs 50
Loss	6.89%	3.18%	2.16%
Accuracy	93.11%	96.82%	97.84%

5. Results

The four deep learning-based models were tested over the 20% data from the ADNI dataset. To check the performance of the proposed models, multiple measures, i.e., sensitivity, Equation (13), specificity, Equation (14), and accuracy, Equation (15), were used. All of these measures were elaborated in terms of True Positive (TP), False Positive (FP), False Negative (FN), and False Positive (FP) in their respective equations.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (14)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

The class-wise outcomes of these measures are depicted in Table 3, while the confusion matrix of the four classes for the best achieved accuracy is given in Tables 4 and 5. We can see that the proposed model performed well in the classification of all categories, especially in the case of AD vs NC, MCI, and LMCI, with higher accuracy of 99.68%, sensitivity of 99.65%, and specificity of 99.69%.

Table 3. Confusion Matrix.

		Predicted			
		MCI	AD	NCI	LMCI
Actual	MCI	248	1	1	0
	AD	1	247	1	1
	NCI	2	1	247	0
	LMCI	0	1	0	249

Table 4. Class-wise performance comparison of proposed model.

Image Classes	Specificity	Sensitivity	Accuracy
MCI vs. AD, NC and LMCI	99.89	98.42	99.52
LMCI vs. AD, NC and MCI	97.76	99.36	98.16
NC vs. AD, MCI and LMCI	99.78	94.32	98.33
AD vs. NC, MCI and LMCI	99.69	99.65	99.68

Table 5. Accuracy achieved by the four models with and without Augmentation and Transfer Learning.

Transfer Learning	Augmented Data	Accuracy
No	No	81.58%
No	Yes	86.04%
Yes	No	93.30%
Yes	Yes	97.84%

6. Discussions

Numerous studies have been conducted recently on early diagnosis of Alzheimer's disease by utilizing deep learning techniques, especially CNN. Most of the diagnosing systems are developed by training deep learning models from scratch and with a huge dataset that is trained over a large number of epochs. This study proposed a transfer learning assisted fine-tuning approach for the detection of four stages of Alzheimer's disease. The designed model was evaluated by using preprocessed GM sliced images with a different number of epochs. The performance of the proposed was evaluated based on three measures, which were sensitivity, accuracy, and specificity. All of these three measures play a huge role in the correct classification of healthy versus ill persons. Gray matter (GM) is the main fundamental part of the human brain which processes information sent in the form of signals by different sensory organs of the body. GM is more helpful in the early diagnosis of AD. This research focused on the segregation of AD patients from healthy people. The proposed model was gradually trained and tested with a different number of epochs. Initially, the model was trained and tested over 10 epochs, with 93.11% accuracy. In the next phase, the number of epochs increased from 10 to 25, and then in the next phase the epochs increased from 25 to 50, with an overall accuracy of 97.84%, which elaborated the effectiveness of the proposed study (Figure 13).

Table 6 depicts the comparison with the base article [32] and the other studies that have contributed to the early diagnosis of Alzheimer's disease. However, the proposed model performs outstandingly with 97.84% accuracy in the multi-class classification of AD with a very less number of epochs.

Early detection of Alzheimer's disease, using MRIs, can be challenging, due to the presence of different artifacts. like noisy background, low contrast and partial volume [45]. To overcome these issues, more recent technologies, like Functional magnetic resonance imaging (fMRI), can be used in the future for the early detection of AD.

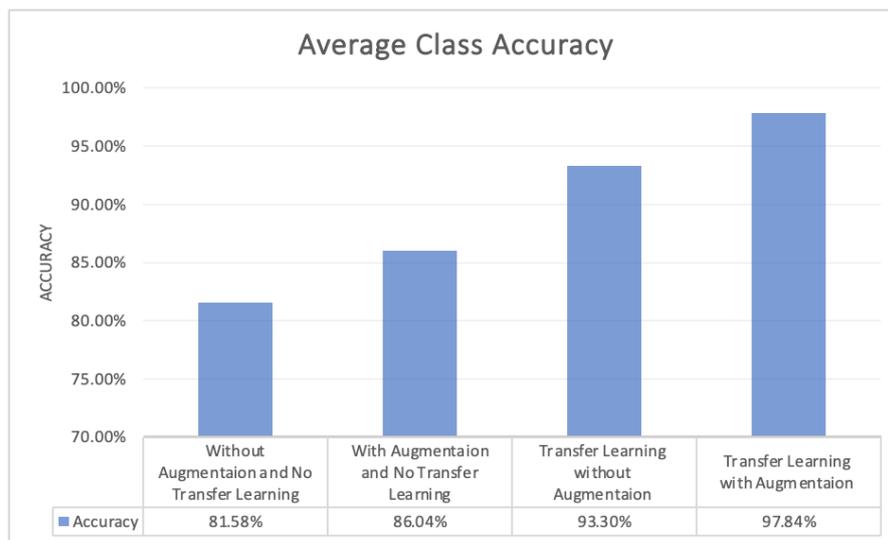


Figure 13. Average Class Accuracy of the four Models.

Table 6. Comparison of the proposed model with existing models.

Model	Year	Accuracy
Inception V4 [27]	2017	96.25%
AlexNet [32]	2019	92.80%
GoogLeNet [44]	2022	96.39%
ResNet-18 [15]	2021	96.88%
Proposed Model		97.84%

7. Conclusions

Alzheimer’s disease is a slow neurological disorder that destroys the thought processes, and consciousness, of a human and, mostly, the aptness to perform straightforward tasks. Many deep learning models have been proposed to detect and classify the stages of AD. Gray matter is the main fundamental part of the human brain that is mainly affected by Alzheimer’s disease and it plays a major role in the gradual destruction of neurons. In this research, a deep learning model, Dense-Net, with transfer learning, was applied to the MRI dataset. The main objective of this research was to classify the stages of Alzheimer’s disease, based on the extracted gray matter (GM), the main fundamental part of the human brain. MRI scans were segmented into 3 parts, GM, WM, and CSF, using the SPM12 for pre-processing. The 2D GM slices were used as input for the training and testing of the model. A pre-trained DenseNet model, with retraining of the last two blocks, was applied over the segmented GM slices. The proposed model provided promising results, with 97.84% accuracy in the multi-class classification of AD.

There were certain pros and cons of the proposed technique. DenseNet is well-reputed, due to its applications in medical imaging [46]. The analysis of medical images by DenseNet is remarkable, as it acquires a comprehensive and complete knowledge of image details from all previous layers. In addition to that, the propagation which takes place layer-wise can be made thinner to make it computationally economical without compromising accuracy, due to the fact that the final prediction depends on features obtained from all layers. As the proposed methodology was based on transfer learning, so the significant limitation of this approach was negative transfer, due to which, for initial training, the target and initial problems were supposed to be adequately homogeneous.

Still, considering the outstanding performance of the proposed model with respect to other CNN models, in the future, this model could be applied for the detection and classification of lungs, and other diseases.

Author Contributions: Conceptualization, N.R. and A.N.; methodology, N.R. and A.N.; software, N.R.; validation, M.T., A.N. and K.Z.; formal analysis, M.T. and K.Z.; investigation, N.R., A.N., M.T. and K.Z.; resources, A.N.; data curation, N.R.; writing—original draft preparation, N.R.; writing—review and editing, A.N. and M.T.; visualization, N.R.X.; supervision, A.N. and K.Z.; project administration, A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset [34], used in this research, is acquired from the publicly available Alzheimer’s Disease Neuroimaging Initiative (ADNI) database [37].

Acknowledgments: We would like to acknowledge all data sources who made all datasets publicly available [34]. In addition to availability of the dataset, all codes for deep CNN architectures are publicly available.

Conflicts of Interest: There is no conflict of interest among the authors. This paper will be of interest to the readership of this Journal. As the corresponding author and first author of the research work, I hereby confirm that the manuscript is entirely original, has not been copyrighted, published, submitted, or accepted for publication elsewhere.

References

1. Srivastava, S.; Ahmad, R.; Khare, S.K. Alzheimer’s disease and its treatment by different approaches: A review. *Eur. J. Med. Chem.* **2021**, *216*, 113320. [[CrossRef](#)] [[PubMed](#)]
2. Alzheimer’s Association. 2015 Alzheimer’s disease facts and figures. *Alzheimer’s Dement.* **2015**, *11*, 332–384. [[CrossRef](#)] [[PubMed](#)]
3. Alzheimer’s Association. 2016 Alzheimer’s disease facts and figures. *Alzheimer’s Dement.* **2016**, *12*, 459–509. [[CrossRef](#)] [[PubMed](#)]
4. Breijyeh, Z.; Karaman, R. Comprehensive review on Alzheimer’s disease: Causes and treatment. *Molecules* **2020**, *25*, 5789. [[CrossRef](#)]
5. Jack, C.R., Jr.; Bernstein, M.A.; Fox, N.C.; Thompson, P.; Alexander, G.; Harvey, D.; Borowski, B.; Britson, P.J.; Whitwell, J.L.; Ward, C.; et al. The Alzheimer’s disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging Off. J. Int. Soc. Magn. Reson. Med.* **2008**, *27*, 685–691. [[CrossRef](#)]
6. Khan, Y.F.; Kaushik, B. Computer vision technique for neuro-image analysis in neurodegenerative diseases: A survey. In Proceedings of the IEEE 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 9–11 March 2020; pp. 346–350.
7. Khan, Y.F.; Kaushik, B. Neuro-image classification for the prediction of alzheimer’s disease using machine learning techniques. In *Proceedings of the International Conference on Machine Intelligence and Data Science Applications*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 483–493.
8. Ungar, L.; Altmann, A.; Greicius, M.D. Apolipoprotein E, gender, and Alzheimer’s disease: An overlooked, but potent and promising interaction. *Brain Imaging Behav.* **2014**, *8*, 262–273. [[CrossRef](#)]
9. Brookmeyer, R.; Johnson, E.; Ziegler-Graham, K.; Arrighi, H.M. Forecasting the global burden of Alzheimer’s disease. *Alzheimer’s Dement.* **2007**, *3*, 186–191. [[CrossRef](#)]
10. Naseer, A.; Yasir, T.; Azhar, A.; Shakeel, T.; Zafar, K. Computer-aided brain tumor diagnosis: Performance evaluation of deep learner CNN using augmented brain MRI. *Int. J. Biomed. Imaging* **2021**, *2021*, 5513500. [[CrossRef](#)]
11. Kundaram, S.S.; Pathak, K.C. Deep learning-based Alzheimer disease detection. In *Proceedings of the 4th International Conference on Microelectronics, Computing And Communication Systems*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 587–597.
12. Naseer, A.; Tamoor, M.; Azhar, A. Computer-aided COVID-19 diagnosis and a comparison of deep learners using augmented CXRs. *J. X-ray Sci. Technol.* **2021**, *30*, 1–21. [[CrossRef](#)]
13. Dadar, M.; Manera, A.L.; Ducharme, S.; Collins, D.L. White matter hyperintensities are associated with grey matter atrophy and cognitive decline in Alzheimer’s disease and frontotemporal dementia. *Neurobiol. Aging* **2022**, *111*, 54–63. [[CrossRef](#)]
14. Aderghal, K.; Khvostikov, A.; Krylov, A.; Benois-Pineau, J.; Afdel, K.; Catheline, G. Classification of Alzheimer disease on imaging modalities with deep CNNs using cross-modal transfer learning. In Proceedings of the 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), Karlstad, Sweden, 18–21 June 2018; pp. 345–350.
15. Ebrahimi, A.; Luo, S.; Initiative, A.D.N. Disease Neuroimaging Initiative for the Alzheimer’s. Convolutional neural networks for Alzheimer’s disease detection on MRI images. *J. Med. Imaging* **2021**, *8*, 024503. [[CrossRef](#)] [[PubMed](#)]
16. 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.
17. Khedher, L.; Ramírez, J.; Górriz, J.M.; Brahim, A.; Segovia, F.; Alzheimer’s Disease Neuroimaging Initiative. Early diagnosis of Alzheimer’s disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing* **2015**, *151*, 139–150. [[CrossRef](#)]

18. Plant, C.; Teipel, S.J.; Oswald, A.; Böhm, C.; Meindl, T.; Mourao-Miranda, J.; Bokde, A.W.; Hampel, H.; Ewers, M. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. *Neuroimage* **2010**, *50*, 162–174. [[CrossRef](#)] [[PubMed](#)]
19. Tamoor, M.; Younas, I. Automatic segmentation of medical images using a novel Harris Hawk optimization method and an active contour model. *J. X-ray Sci. Technol.* **2021**, *29*, 721–739. [[CrossRef](#)] [[PubMed](#)]
20. Malik, Y.S.; Tamoor, M.; Naseer, A.; Wali, A.; Khan, A. Applying an adaptive Otsu-based initialization algorithm to optimize active contour models for skin lesion segmentation. *J. X-ray Sci. Technol.* **2022**, *30*, 1–16. [[CrossRef](#)] [[PubMed](#)]
21. Gao, M.; Bagci, U.; Lu, L.; Wu, A.; Buty, M.; Shin, H.C.; Roth, H.; Papadakis, G.Z.; Depeursinge, A.; Summers, R.M.; et al. Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2018**, *6*, 1–6. [[CrossRef](#)]
22. Liu, X.; Wang, C.; Bai, J.; Liao, G. Fine-tuning pre-trained convolutional neural networks for gastric precancerous disease classification on magnification narrow-band imaging images. *Neurocomputing* **2020**, *392*, 253–267. [[CrossRef](#)]
23. Naseer, A.; Zafar, K. Meta-feature based few-shot Siamese learning for Urdu optical character recognition. In *Computational Intelligence*; Wiley: Hoboken, NJ, USA, 2022.
24. Sezer, A.; Sezer, H.B. Convolutional neural network based diagnosis of bone pathologies of proximal humerus. *Neurocomputing* **2020**, *392*, 124–131. [[CrossRef](#)]
25. Kim, H. Organize Everything I Know. 2018. Available online: <https://oi.readthedocs.io/en/latest/computer{ }vision/cnn/densenet.html> (accessed on 25 November 2022).
26. Tsang, S.H. Review: DenseNet—Dense Convolutional Network (Image Classification). 2018. Available online: <https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803> (accessed on 25 November 2022).
27. 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, MO, USA, 13–16 November 2017; pp. 1166–1169.
28. Sarraf, S.; Tofighi, G. Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data. In Proceedings of the IEEE 2016 Future Technologies Conference (FTC), San Francisco, CA, USA, 6–7 December 2016; pp. 816–820.
29. Liu, S.; Liu, S.; Cai, W.; Che, H.; Pujol, S.; Kikinis, R.; Feng, D.; Fulham, M.J. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Trans. Biomed. Eng.* **2014**, *62*, 1132–1140. [[CrossRef](#)]
30. Tufail, A.B.; Ma, Y.K.; Zhang, Q.N. Binary classification of Alzheimer's disease using sMRI imaging modality and deep learning. *J. Digit. Imaging* **2020**, *33*, 1073–1090. [[CrossRef](#)]
31. Liu, S.; Liu, S.; Cai, W.; Pujol, S.; Kikinis, R.; Feng, D. Early diagnosis of Alzheimer's disease with deep learning. In Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April–2 May 2014; pp. 1015–1018.
32. 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](#)] [[PubMed](#)]
33. Mehmood, A.; Maqsood, M.; Bashir, M.; Shuyuan, Y. A deep Siamese convolution neural network for multi-class classification of Alzheimer disease. *Brain Sci.* **2020**, *10*, 84. [[CrossRef](#)] [[PubMed](#)]
34. 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](#)] [[PubMed](#)]
35. 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](#)]
36. 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.* **2020**, *44*, 1–16. [[CrossRef](#)]
37. Weiner, D.M.W. Alzheimer Disease Neuroimaging Initiative. 2004. Available online: <https://adni.loni.usc.edu/> (accessed on 25 November 2022).
38. Kalavathi, P.; Prasath, V. Methods on skull stripping of MRI head scan images—A review. *J. Digit. Imaging* **2016**, *29*, 365–379. [[CrossRef](#)]
39. Atkins, M.S.; Mackiewicz, B.T. Fully automatic segmentation of the brain in MRI. *IEEE Trans. Med. Imaging* **1998**, *17*, 98–107. [[CrossRef](#)]
40. Young, J.; Modat, M.; Cardoso, M.J.; Mendelson, A.; Cash, D.; Ourselin, S.; Alzheimer's Disease Neuroimaging Initiative. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *NeuroImage Clin.* **2013**, *2*, 735–745. [[CrossRef](#)]
41. Ashburner, J.; Friston, K. Voxel-Based Morphometry—The Methods. *Neuroimage* **2000**, *11 Pt 1*, 805–821. [[CrossRef](#)]
42. Bose, S.R.; Pearline, A.S.; Bose, R. *Real-Time Plant Species Recognition Using Non-Averaged Densenet-169 Deep Learning Paradigm*; SSRN: Rochester, NY, USA, 2022.
43. Vulli, A.; Srinivasu, P.N.; Sashank, M.S.K.; Shafi, J.; Choi, J.; Ijaz, M.F. Fine-Tuned DenseNet-169 for Breast Cancer Metastasis Prediction Using FastAI and 1-Cycle Policy. *Sensors* **2022**, *22*, 2988. [[CrossRef](#)]
44. Shanmugam, J.V.; Duraisamy, B.; Simon, B.C.; Bhaskaran, P. Alzheimer's disease classification using pre-trained deep networks. *Biomed. Signal Process. Control* **2022**, *71*, 103217. [[CrossRef](#)]

45. Chamarajan, G.; Charishma, Y.; Ghomi, S.M.T.; Marchiori, E. Alzheimer's Disease: A Survey. *Int. J. Artif. Intell.* **2021**, *8*, 33–39.
46. Zhou, T.; Ye, X.; Lu, H.; Zheng, X.; Qiu, S.; Liu, Y. Dense Convolutional Network and Its Application in Medical Image Analysis. *BioMed Res. Int.* **2022**, *2022*, 2384830. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.