Residual block fully connected DCNN with categorical generalized focal dice loss and its application to Alzheimer's disease severity detection

Adi Alhudhaif [Corresp., 1], **Kemal Polat** [2]

[1] Department of Computer Science, College of Computer Engineering and Sciences in Al-kharj, Prince Sattam Bin Abdulaziz University, Al-kharj, Saudi Arabia

[2] Department of Electrical and Electronics Engineering, Faculty of Engineering, Bolu Abant Izzet Baysal University, Bolu, TURKEY

Corresponding Author: Adi Alhudhaif
Email address: a.alhudhaif@psau.edu.sa

Background: Alzheimer's disease (AD) is a disease that manifests itself with a deterioration in all mental activities, daily activities, and behaviors, especially memory, due to the constantly increasing damage to some parts of the brain as people age. Detecting AD at an early stage is a significant challenge. Various diagnostic devices are used to diagnose AD. MRI (Magnetic Resonance) devices are widely used to analyze and classify the stages of AD. However, the time-consuming process of recording the affected areas of the brain in the images obtained from these devices is another challenge. Therefore, conventional techniques cannot detect the early stage of AD.

Methods: This study proposes a deep learning model supported by a fusion loss model that includes fully connected layers and residual blocks to solve the above-mentioned challenges. The proposed model has been trained and tested on the publicly available T1-weighted MRI-based KAGGLE dataset. Data augmentation techniques were used after various preliminary operations were applied to the data set.

Results: The proposed model effectively classified four AD classes in the KAGGLE dataset. The proposed model reached the test accuracy of 0.973 in binary classification and 0.982 in multi-class classification thanks to experimental studies and provided a superior classification performance than other studies in the literature.

Residual Block Fully Connected DCNN with Categorical Generalized Focal Dice Loss and its Application to Alzheimer's Disease Severity Detection

Adi Alhudhaif[1*] and Kemal Polat[2]

[1*]Department of Computer Science, College of Computer Engineering and Sciences in Al-kharj, Prince Sattam bin Abdulaziz University, P.O. Box 151, Al-Kharj 11942, Saudi Arabia,
e-mail: A.alhudhaif@psau.edu.sa

Department of Electrical and Electronics Engineering, Faculty of Engineering, Bolu Abant Izzet Baysal University, Bolu, Turkey
e-mail: kpolat@ibu.edu.tr

Corresponding Author:
Dr. Adi Alhudhaif
Department of Computer Science, College of Computer Engineering and Sciences in Al-kharj, Prince Sattam bin Abdulaziz University, P.O. Box 151, Al-Kharj 11942, Saudi Arabia
Email address: A.alhudhaif@psau.edu.sa

Abstract

Background: Alzheimer's disease (AD) is a disease that manifests itself with a deterioration in all mental activities, daily activities, and behaviors, especially memory, due to the constantly increasing damage to some parts of the brain as people age. Detecting AD at an early stage is a significant challenge. Various diagnostic devices are used to diagnose AD. MRI (Magnetic Resonance) devices are widely used to analyze and classify the stages of AD. However, the time-consuming process of recording the affected areas of the brain in the images obtained from these devices is another challenge. Therefore, conventional techniques cannot detect the early stage of AD.

Methods: This study proposes a deep learning model supported by a fusion loss model that includes fully connected layers and residual blocks to solve the above-mentioned challenges. The proposed model has been trained and tested on the publicly available T1-weighted MRI-based KAGGLE dataset. Data augmentation techniques were used after various preliminary operations were applied to the data set.

Results: The proposed model effectively classified four AD classes in the KAGGLE dataset. The proposed model reached the test accuracy of 0.973 in binary classification and 0.982 in multi-class classification thanks to experimental studies and provided a superior classification performance than other studies in the literature.

Keywords: Alzheimer's disease; Categorical Generalized Focal Dice Loss; Deep learning; New hybrid models; classification

Introduction

The brain is a vital organ that contains memory and manages thoughts and decision-making [Armstrong et al., 2009]. AD is one of the causes of dementia. In AD, beta-amyloid and phosphorylated tau proteins accumulate excessively, which causes brain cell degeneration [Alzheimer's disease facts and figures, 2022]. AD is an irreversible disease that causes progressive brain deterioration. Memory cells gradually die, causing an increasing shrinkage in the brain [Alberdi et al., 2018]. AD is a fatal neurological disease with a life expectancy of 4-8 years after diagnosis [Babalola et al., 2009]. This disease, which can be seen at any age, is more common in people over 65 [Coppola et al., 2013; Chandra et al., 2019].

- -

According to official figures in the United States, 121,499 people died from AD in 2019. AD was ranked sixth among the causes of death in 2019 when Covid-19 was among the top 10 reasons. In addition, In 2020 and 2021, it was shown as the seventh leading cause of death. Total payments made in 2022 for health care, long-term care, and hospice services for people aged 65 and over with dementia are estimated at $321 billion [Alzheimer's disease facts and figures, 2022].

Several different computer-assisted neuroimaging methods can diagnose AD. In light of clinical experience, MRI has become almost standardized in these imaging modalities. Although AD is incurable, its progression can be slowed with early diagnosis and treatment [Gopinadhan et al., 2022].

In the last decades, many machine learning-based and specifically deep learning-based approaches to diagnosing AD have been presented. If we summarize the previous studies in the literature, classification based on segmentation of brain images was used for early diagnosis of AD by Mehmood et al. For the AD classification, the VGG network was used for transfer learning [Mehmood et al., 2021]. Lahmiri et al. used Deep Convolutional Neural Networks (DCNN) to increase the diagnostic sensitivity of AD. Besides, the KNN approach was used to filter the number of features [Lahmiri, 2023]. Shanmugam et al. analyzed the success of three pre-trained networks, GoogLeNet, AlexNet, and ResNet-18, in classifying AD's stages [Shanmugam et al., 2022]. Frizzell et al. conducted a systematic literature review covering studies conducted between 2009 and 2020 for the diagnosis of AI-based AD using the PubMed database. In the review, images consisting of 3 different classes, including normal aging, mild cognitive impairment, and AD, obtained from MRI, were examined, and comparatively analyzed with the proposed artificial intelligence algorithms [Frizzell et al., 2022]. Kumar et al. have proposed a classification model that uses the AlexNet framework to diagnose AD at an early stage from MR images [Kumar et al., 2022]. Jung et al. proposed a new conditional generative adversarial network (cGAN) capable of synthesizing high-quality

3D MR images. The proposed model consists of an attention-based 2D generator, a 2D parser, and a 3D splitter that can synthesize 2D slices from 3D MR images [Jung et al., 2023]. Liu et al. proposed the Monte Carlo Aggregated Neural Network model combining ResNet50 and Monte Carlo sampling for early AD detection. The proposed model is trained on 2D slices obtained from 3D MR images [Liu et al., 2023]. Sharma et al. proposed an artificial neural network model using the VGG16 deep learning network as a feature extractor for the classification of 4 different stages of AD [Sharma et al., 2022]. Alorf et al. proposed a multi-label spoofing model using the Stacked Sparse Autoencoder and Brain Connectivity Graph Convolutional Network models for the six stages of AD obtained from the rs-fMR imaging device [Alorf et al., 2022]. Raghavaiah et al. trained a DCNN-based model on fMRI and MRI to diagnose AD from specific sound control information. It is fed to the proposed model via an image converter with decomposed parameters from fMRI and MRI [Ravagiah et al., 2021]. Loddo et al. proposed a deep learning model combining AlexNet, ResNet101, and Inception-ResNet-V2 for AD classification [Loddo et al., 2021]. Studies in the literature achieved test success of up to 99% in ADNI and OASIS data sets with deep learning models. As far as we know in the literature, only Loddo et al. and Sharma et al. tested their models on the Kaggle dataset. The Kaggle dataset is a dataset that challenges models due to both its small size and difficulties in distinguishing between classes.

Manual classification of AD on brain images obtained from MRI devices is time-consuming. At the same time, AD is very similar to what happens in the brain with aging. Therefore, it is a challenging task to diagnose AD by clinicians. Thus, Deep learning-based computer-aided systems, which often achieve higher success than clinicians, gain more significant importance for AD classification.

In summary, the main contributions of this paper to the literature are given as follows:

- In this study, we divided an inspired by the encoder layer of the U-shaped segmentation algorithm and connected the model we designed to the fully connected layer. Therefore, we have fused the Generalised Dice Loss(GDL) function used in multi-class segmentation with the Focal Loss(FL) function in such a network.
- Non-Local Means and Estimate sigma algorithms are first fused to eliminate noise in MR images by us.
- Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm used for histogram equalization to enhance MR images.
- A fully connected Deep Convolutional neural network with Residual blocks is proposed to classify AD.
- A fusion loss function increased the network's test accuracy in classifying AD from MR images.

The second section, consisting of the material and method section, gives detailed information about the pre-processing stages of the data set and the proposed model. The practical applications of the model and the test results obtained from the proposed model are evaluated and discussed in the 3rd chapter, the discussion section. Finally, the work done in the 4th chapter is summarized, and suggestions for future studies are shared.

Materials & Methods

MRI dataset

The KAGGLE online community obtained the dataset [Dubey, 2023]. The dataset obtained from KAGGLE consists of 5121 axial images collected from different websites. The images in the dataset were collected and labeled into four different classes. These are no dementia, very mild dementia, mild dementia, and moderate dementia. There is no information about the age of the patients in the MR images obtained from the patients. The training dataset consists of MR images of 2560 healthy (ND), 1792 very mild dementia (vmD), 717 mild dementia (miD), and 52 moderate dementia (mD) individuals. The test dataset consists of MR images of 640 healthy (NC), 448 very mild dementia (vmD), 179 mild dementia (miD), and 12 moderate dementia (mD) individuals. The resolution of the images is 176x208. Samples are shown in Fig. 1. No other circumstances of the MR images obtained were specified.

Fig. 1. Samples of axial images in the Kaggle data set

Dataset Pre-processing

Various image restoration techniques are used in the Kaggle AD dataset to increase the model's training and test accuracy. First, the Non-Local Means (NLM) algorithm was applied to the images. Then, the image noise was de-noised using the weighted average of the pixel neighborhoods with similarity. Also, to improve performance when calculating the similarity of each pixel, instead of considering just one pixel, a small area of pixels is chosen around it, given by the small_window parameter. Patch_Size=2 and Patch_Distance=1 was selected to increase the noise removal performance of the NLM algorithm. Finally, Estimate_sigma is fused with the NLM algorithm to improve the performance of the proposed model.

The resulting images were passed through the CLAHE algorithm. Contrast-limited histogram equalization is performed by dividing images into small blocks called tiles in CLAHE in the OpenCV library. As a result of the experimental studies, the clipping level in CLAHE was chosen as 2.6. Samples of 4 different classes are shown in Fig. 2., showing the MR images in the Kaggle dataset before and after the noise removal and CLAHE. In Fig. 2., the pictures in the top line are the MR images before the image pre-processing techniques, and the images in the bottom line are the MR images obtained after the image pre-processing methods. The pre-processed training data set was doubled by applying random shifting, random rotation, random rescaling, and random horizontal and vertical flip methods.

Fig. 2. Samples of image pre-processing techniques in the Kaggle data set: the pictures in the top line are the MR images before the image pre-processing techniques, and the images in the bottom line are the MR images obtained after the pre-processing image methods.

The Proposed Categorical Generalized Focal Dice Loss Function

The proposed Hybrid Loss function is a fused model of the Generalized Dice Loss (GDL) and Focal Loss(FL) [Sudre et al., 2017; Lin et al., 2022]. The FL function formula is shown in Eq. 1.

$$FL = -\alpha_t(1-p_t)^{\gamma}\log(p_t) \quad (1)$$

As seen in Eq. 1., Contrary to Cross-entropy (CE), Loss of Focus prevents suppression of CE Loss in large class imbalances. Thus, the Negatives that make up most of the Loss are reduced. Adding α to Focal Loss in Balanced Cross Entropy balances positive and negative samples. In addition, the adjustable focusing parameter $\gamma \geq 0$ and the cross-entropy Loss modulating factor $(1-p_t)$ was added to reduce the weight of easy-to-learn samples and to enable the deep-learning model to focus on difficult-to-train samples. As a result of the experimental studies performed in this study, we chose γ=2 and α=4, where p is the model's predicted probability for the class [Lin et al., 2022].
The most significant disadvantage of Dice Loss is that although it has a good training score, its test score is low due to its poor response to class imbalances. As a solution to this, Sudre et al. transformed the Generalized Dice Score (GDS) function proposed by Crum et al. to score multi-

class segmentation into a loss function named Generalized Dice Loss(GDL) [Crum et al., 2006]. The GDL function formula is shown in Eq. 2.

$$GDL = 1 - 2\frac{\sum_{l=1}^{2} w_l \sum_{n} r_{ln}\, p_{ln} + \in}{\sum_{l=1}^{2} w_l \sum_{n} r_{ln} + p_{ln} + \in} \quad (2)$$

Where r_l is the ground truth, and p_l is the predicted value. $\in$ is a smoothing value for fine-tuning. Here w_l is used to provide immutability to different labeled sets where w_l is formulated as shown in E.q. 3.

$$w_l = {}^{1}\Big/_{(\sum_{n=1} r_{ln})^2} \quad (3)$$

Where r_l is the ground truth. In this study, FL and GD, seen as robust models in multi-class segmentation, were fused to predict four different classes of AD, as shown in E.q. 4.

Fusion_Loss = (1 - lamda) * GDL + lamda * FL (4)

The lambda value was chosen as 0.5 because of experimental studies.

Performance Metrics

Accuracy, Sensitivity, and Specificity metrics were used to measure the performance of the proposed deep learning-based classification model. Accuracy is the ratio of correctly predicted samples to the sum of all correct and incorrectly guessed samples; Sensitivity, or recall, is the ratio of positively correctly predicted samples to the sum of negative incorrectly predicted samples and correctly predicted positive samples. Specificity (Spec) represents the ratio of correctly predicted negative samples to the sum of correctly predicted negative samples and incorrectly predicted positive samples.

The Proposed Methodology

The end-to-end automatically AD classifier block diagram is shown in Fig. 3. The section including CLAHE covers the pre-processing stage of the dataset, and the section after CLAHE covers the training and testing of the model.

Fig. 3 End-to-end automatically DCNN-based AD classifier block diagram

The Proposed Deep Learning Model

In the proposed deep learning architecture, 2xConvolutional2D blocks with 3x3 filters are used in all layers before the fully connected layer. In addition, the Xavier was used as the weight initializer function [Glorot et al., 2010]. Group Normalization (GN) is the normalizer, and ReLU is the activation function of the proposed model [Wu and He, 2020]. Wu and He demonstrated with various experiments how superior the group normalization technique is to other normalization methods.

Besides, ADAM was used as the optimizer to optimize the weights of the proposed model [Kingma et al., 2015]. In the last layer, Softmax was used as an activation function to find out which class the MR image belongs to the AD class. If we pay attention to the proposed model, it is seen that the U-shaped network is formed by taking the deconvolution part and adding the Fully connected network. Deep Neural networks are profound networks due to the multitude of layers. Therefore, the information learned in the first layers can be forgotten. The residual blocks(ResNet Block) ensure that the convolutional layers do not forget the information they learned in the first layers of the network. In the proposed model, Residual blocks are followed by 2×2 maximum pooling(Max-pool), in which the feature maps' spatial dimensions (height and width) are reduced by half. Max-pool reduces the computational cost by reducing the number of trainable parameters. The layers of the proposed deep learning architecture are shown in Fig. 4. in detail.

Fig. 4. Detailed representation of the proposed architecture

Proposed Model's Experimental Studies

This section explains the experimental studies of the proposed deep learning architecture on the KAGGLE dataset in detail.

Employed hardware materials

Experimental studies of the proposed architecture were conducted with a computer equipped with Intel(R) Core(TM) i5-8300H @ 2.30 GHz CPU, 32 GB RAM, and NVIDIA GTX1050 4 GB GPU. In addition, the deep learning ecosystem consists of artificial intelligence libraries with Python 3.7 programming language based on Anaconda. Libraries and ecosystems are entirely open source. The dataset consists of training, validation, and test datasets. 20% of the training dataset was used as a validation dataset using the scikit-learn library. We performed 5-fold cross-validation to measure the training dataset's heterogeneity and the model's fit.

Comparison of parameters

The model used in the proposed study and the models used in the other literature were analyzed comparatively regarding the number of parameters. As can be seen from the comparisons in Table 1, the proposed model is the one with the smallest parameter. It can be seen from Table 1 that the model with the smallest parameter is the recommended model.

Table1. Comparison of parameters and their computational efficiency

Comparison of loss functions

The proposed fusion loss function has been analyzed in comparison with the most used Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Cross-Entropy, FL, and GDL loss function for classification in the literature. As can be seen from Table 2, the proposed fusion loss function is the most suitable loss function for the proposed architecture.

Table2. Effects of different loss functions on the proposed architecture.

Settings and studies on the KAGGLE dataset

Using our deep learning architecture, we performed binary classification, i.e., distinguishing NC and AD, and multi-class classification (NC, very mild AD, mild AD, moderate AD) experimental studies on the KAGGLE dataset. The deep learning network was trained for 15 thousand epochs for these two studies. After 15.000 epochs, the validation accuracy of the network remained constant at around 0.99. The minibatch size is 32, and ADAM is chosen as the optimizer. Numerical information about the data set used in training the proposed model is listed in Table 3. As can be seen from Table 3, there is a significant imbalance between classes. In this case, the benefits of data augmentation methods will be limited. Therefore, GDL and FL, which are successful and robust loss functions in class imbalances, are proposed as a solution to the class imbalance of the data set.

Table 3. Slice-based MR image counts used for train, validation, and testing of the KAGGLE Data set on which the proposed model is trained and tested.

Experimental results and analysis

In this section, experimental results are shared, and comparative analyzes are made with other studies in the literature. In addition, information is given about the deep learning models used for detecting AD and how they train the dataset on the performance criteria they use. A slice-based data set was used in this study. The Kaggle dataset is a low-resolution and slice-based dataset collected on different websites. In addition, class imbalances are also very high in the dataset. Also, the Kaggle dataset's having axial views is another challenge. For these reasons, the Kaggle dataset is very challenging. The most significant difficulty of all the literature and proposed research studies is distinguishing AD from NC in the MCI stage. If AD is successfully detected in the MCI stage, the effects of this neurodegenerative and irreversible disease can be slowed down. In addition, one of the most critical challenges is to detect NCs (very MCI) that may have AD. This study focused on two different experimental studies: i) distinguishing NC from AD at the MCI stage; ii) To successfully classifying MR images with NC, vMCI, MCI, and MC. For this

reason, articles in the literature that made these two different experimental studies and analyzed their results were examined. Validation Accuracy and Validation Loss values obtained using the Categorical Generalized focal Membrane loss function of the proposed model are shown graphically in Figure 5 and Figure 6.

Fig. 5. ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

Fig. 6. ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

The comparative performance results of our proposed fusion loss-based deep learning architecture are shown in Table 4 and Table 5. The proposed deep learning architecture achieved accuracy values of 0.973 in binary classification and 0.982 in multi-class classification. At the same time, it obtained high sensitivity and specificity values compared to other studies in the literature.

Table 4. Experimental results for NC and AD binary classification

Table 5. Experimental results for multi-class classification (NC, vmiD, miD, mD)

Discussion

This study proposes a deep learning-based model with fusion loss for diagnosing and classifying AD. The Kaggle dataset, the most challenging AD dataset, was used to compare the proposed architecture with other studies in the literature. High accuracy values of up to 0.99 were obtained in other publicly available datasets (OASIS and ADNI) used in the literature. The Kaggle dataset has four classes: vmiD, miD, mD, and NC. The proposed fusion loss deep learning model showed a higher performance than other studies in the literature. In addition, Ensemble-based deep learning algorithms are generally emphasized in the literature. The deep learning architecture proposed in this paper also dramatically reduces computation. However, the minibatch could not be increased enough due to hardware limitations. In addition, resizing was not done because the resolution of the data line was low. Therefore, the number of epochs determined for training the proposed model

may be higher than those used by other models. However, the parameter number of the proposed architecture is considerably smaller than the parameter numbers of different architectures in the literature.

Conclusions

This study proposes a fusion loss deep learning model using the group normalization technique with residual blocks to detect three different classes (vmiD, miD, mD) of AD disease. The proposed model is trained on the KAGGLE dataset, a very challenging dataset that includes three other dementia classes (vmiD, miD, mD) and normal cognitive status. We tried to solve the large class imbalance in the Kaggle dataset using the Categorical Generalised Focal Dice Loss function proposed in this study. In the proposed model, FL and GDL loss functions are used successfully in multi-class segmentation by fusing them. FL and GDL were used separately in the proposed model, and 84% to 88% accuracy was achieved, while an accuracy value of 0.982 was obtained in the model with the fused loss function. Compared to the latest technology studies in the literature, the proposed model has achieved very high success in binary classification (NC and AD) and multi-class classification (vmiD, miD, mD, NC). Although the proposed architecture showed high performance on a challenging dataset, it must be tested on an MR device in real-time.

Acknowledgements

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through project number (IF-PSAU-2021/01/18563).

Ethical approval

This article does not contain any data, or other information from studies or experimentation, with the involvement of human or animal subjects.

Conflicts of interest

The authors declare that there is no conflict to interest related to this paper.

Raw Data Availability

Dubey, S. Alzheimer's dataset (4 class of images) [online] Available: https://www. kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images (Accessed 17 July 2023).

Source code is available at GitHub:
https://github.com/kpolat14/data-weighting-model-and-classification

References

Alzheimer's disease facts and figures. Alzheimer's Dement., 18: 700-789, 2022. https://doi.org/10.1002/alz.12638.

Alberdi, A., Weakley, A., Schmitter-Edgecombe, M., Cook, D.J., Aztiria, A., Basarab, A., Barrenechea, M.(2018) Smart home-based prediction of multidomain symptoms related to Alzheimer's disease. IEEE J. Biomed. Health Inform. 2018;22:1720–31.

Alorf, A., Khan, M.U.G.(2022) Multi-label classification of Alzheimer's disease stages from resting-state fMRI-based correlation connectivity data and deep learning. Comput Biol Med. 2022 Dec;151(Pt A):106240. doi: 10.1016/j.compbiomed.2022.106240. Epub 2022 Nov 9. PMID: 36423532.

Armstrong, N., Nugent, C., Moore, G., Finlay, D.(2009) In International conference on smart homes and health telematics; Springer, pp 25-31.

Babalola, K.O., Patenaude, B., Aljabar, P., Schnabel, J., Kennedy, D., Crum, W., Smith, S., Cootes, T., Jenkinson, M., Rueckert, D.(2009). Anevaluationoffour automatic methods of segmenting the subcortical structures in the brain. Neuroimage 2009;47(4): 1435–47.

Ba, J., Kingma, DP, , 2015. Adam: a method for stochastic optimization. in: International Conference on Learning Representations (ICLR), pp.1–11.

Chandra, A., Dervenoulas, G., Politis, M.(2019). Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment. J. Neurol. 2019;266:1293–302.

Coppola, J.F., Kowtko, M.A., Yamagata, C., Joyce, S.(2013). Applying mobile application development to help dementia and Alzheimer patients. Wilson Center Soc. Entrepreneur. 2013;16. Paper.

Crum, W., Camara, O., Hill, D.: Generalized Overlap Measures for Evaluation and Validation in Medical Image Analysis. IEEE TMI 25(11), 1451{1461 (nov 2006)

Dubey, S. Alzheimer's dataset (4 class of images) [online] Available: https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images. (Accessed 07 Feb 2023).

Frizzell, T.O., Glashutter, M., Liu, C.C., Zeng, A., Pan, D., Hajra, S.G., D'Arcy, R.C.N., Song X.(2022). Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: A systematic review. Ageing Res Rev. 2022 May;77:101614. doi: 10.1016/j.arr.2022.101614. Epub 2022 Mar 28. PMID: 35358720.

Glorot, X., Bengio, Y, 2010. Understanding the difficulty of training deep feedforward neural networks. In Aistats, volume 9, pages 249–256.

Gopinadhan, A., Angeline Prasanna, G., Anbarasu, S.(2022). AD-EHS: Alzheimer's disease severity detection using efficient hybrid image segmentation,Advances in Engineering Software, Volume 173, 2022, 103234, ISSN 0965-9978, https://doi.org/10.1016/j.advengsoft.2022.103234.

Jung, E., Luna, M., Park, S.H.(2023). Conditional GAN with 3D discriminator for MRI generation of Alzheimer's disease progression. Pattern Recognition, Volume 133, 2023,109061, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2022.109061.

Lahmiri, S.(2023). Integrating convolutional neural networks, kNN, and Bayesian optimization for efficient diagnosis of Alzheimer's disease in magnetic resonance images. Biomedical Signal Processing and Control, Volume 80, Part 2, 2023, 104375, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2022.104375.

Lin, T.Y., Goyal, P., Girshick, R., He K., Dollár, P.(2020). Focal Loss for Dense Object Detection. in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 318-327, 1 Feb. 2020, doi: 10.1109/TPAMI.2018.2858826.

Loddo, A., Buttau, S., Di Ruberto, C.(2021) Deep learning based pipelines for Alzheimer's disease diagnosis: A comparative study and a novel deep-ensemble method. Comput Biol Med. 2022 Feb;141:105032. doi: 10.1016/j.compbiomed.2021.105032. Epub 2021 Nov 21. PMID: 34838263.

Mehmood, A., Yang, S., Feng, Z., Wang, M., Ahmad, A.S., Khan, R., Maqsood, M., Yaqub, M.(2021). A Transfer Learning Approach for Early Diagnosis of Alzheimer's Disease on MRI Images. Neuroscience, Volume 460, 2021, Pages 43-52.

Raghavaiah, P., Varadarajan, S.(2021) Novel deep learning convolution technique for recognition of Alzheimer's disease, Materials Today: Proceedings, Volume 46, Part 9, 2021, Pages 4095-4098, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.02.626.

426 Sathish Kumar, L., Hariharasitaraman, S., Narayanasamy, K., Thinakaran, K., Mahalakshmi, J.,
427 Pandimurugan. V.(2022). AlexNet approach for early stage Alzheimer's disease
428 detection from MRI brain images. Materials Today: Proceedings, Volume 51, Part 1,
429 2022, Pages 58-65, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.04.415.
430 Sharma, S., Guleria, K., Tiwari, S., Kumar, S.(2022). A deep learning based convolutional neural
431 network model with VGG16 feature extractor for the detection of Alzheimer Disease
432 using MRI scans, Measurement: Sensors, Volume 24, 2022, 100506, ISSN 2665-
433 9174, https://doi.org/10.1016/j.measen.2022.100506.
434 Shanmugam, J.V., Duraisamy, B., Simon, B.C., Bhaskaran, P.(2022). Alzheimer's disease
435 classification using pre-trained deep networks. Biomedical Signal Processing and
436 Control, Volume 71, Part B, 2022, 103217, ISSN 1746-8094,
437 https://doi.org/10.1016/j.bspc.2021.103217.
438 Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M. (2017). Generalised Dice
439 Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. In:
440 , et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical
441 Decision Support . DLMIA ML-CDS 2017 2017. Lecture Notes in Computer Science(),
442 vol 10553. Springer, Cham. https://doi.org/10.1007/978-3-319-67558-9_28.
443 Wu, Y., He, K. Group Normalization. Int J Comput Vis 128, 742–755 (2020).
444 https://doi.org/10.1007/s11263-019-01198-w

Table 1(on next page)

Comparison of parameters and their computational efficiency

Comparison of parameters and their computational efficiency

1 **Table1.** Comparison of parameters and their computational efficiency

Architectures	Parameter(M)
VGG16[Sharma et al.]	134.27
Alexnet[Loddo et al.]	62.38
Resnet 101[Loddo et al.]	44.5
Inception ResNetv2[Loddo et al.]	54.5
Proposed Architecture	5.45

2

3

Table 2(on next page)

Effects of different loss functions on the proposed architecture.

Effects of different loss functions on the proposed architecture.

1 **Table2.** Effects of different loss functions on the proposed architecture.

Methods	Performance metrics		
	Accu	Sens	Spec
Cross-entropy	0.966	0.966	0.982
MAE[Loddo et al.]	0.932	0.932	0.934
RMS[Loddo et al.]	0.956	0.959	0.957
FL[Loddo et al.]	0.927	0.930	0.928
GDL	0.965	0.968	0.966
Our Fusion Loss	**0.982**	**0.982**	**0.989**

2

Table 3(on next page)

Slice-based MR image counts used for train, validation, and testing of the KAGGLE Data set on which the proposed model is trained and tested.

Slice-based MR image counts used for train, validation, and testing of the KAGGLE Data set on which the proposed model is trained and tested.

Table 3. Slice-based MR image counts used for train, validation, and testing of the KAGGLE Data set on which the proposed model is trained and tested.

Data set name	NC	vmiD	miD	mD
Training data set	2050	1075	573	42
Train+Data Augmented	7170	4659	2007	146
Validation data set	510	717	287	10
Test data set	640	448	179	12

Table 4(on next page)

Experimental results for NC and AD binary classification

Experimental results for NC and AD binary classification

1 **Table 4.** Experimental results for NC and AD binary classification

Methods	Performance metrics		
	Accu	Sens	Spec
Deep-Ensemble[Loddo et al.]	0.966	0.966	0.982
AlexNet[Loddo et al.]	0.897	0.898	0.897
ResNet-101[Loddo et al.]	0.961	0.961	0.961
Inception-ResNet-v2[Loddo et al.]	0.912	0.914	0.912
(Our Methodology)	**0.973**	**0.975**	**0.988**

2

3

Table 5(on next page)

Experimental results for multi-class classification (NC, vmiD, miD, mD)

Experimental results for multi-class classification (NC, vmiD, miD, mD)

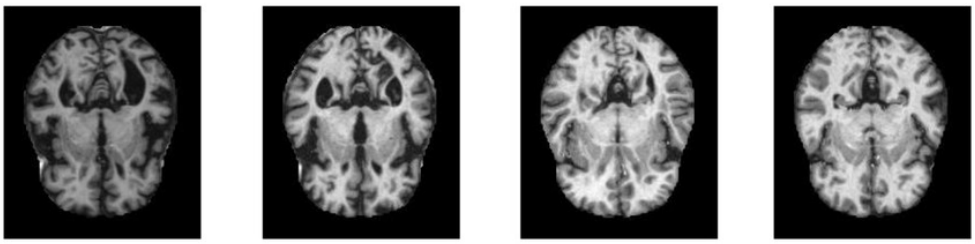
Table 5. Experimental results for multi-class classification(NC, vmiD, miD, mD)

Methods	Performance metrics		
	Accu	Sens	Spec
Deep-Ensemble[Loddo et al.]	0.971	0.967	0.982
Neural Nets with VGG16 [Sharma et al.,]	0.904	0.905	0.904
AlexNet[Loddo et al.]	0.893	0.906	0.817
ResNet-101[Loddo et al.]	0.965	0.978	0.961
Inception-ResNet-v2[Loddo et al.]	0.897	0.901	0.856
(Our Methodology)	**0.982**	**0.982**	**0.989**

Figure 1

Samples of axial images in the Kaggle data set

Samples of axial images in the Kaggle data set



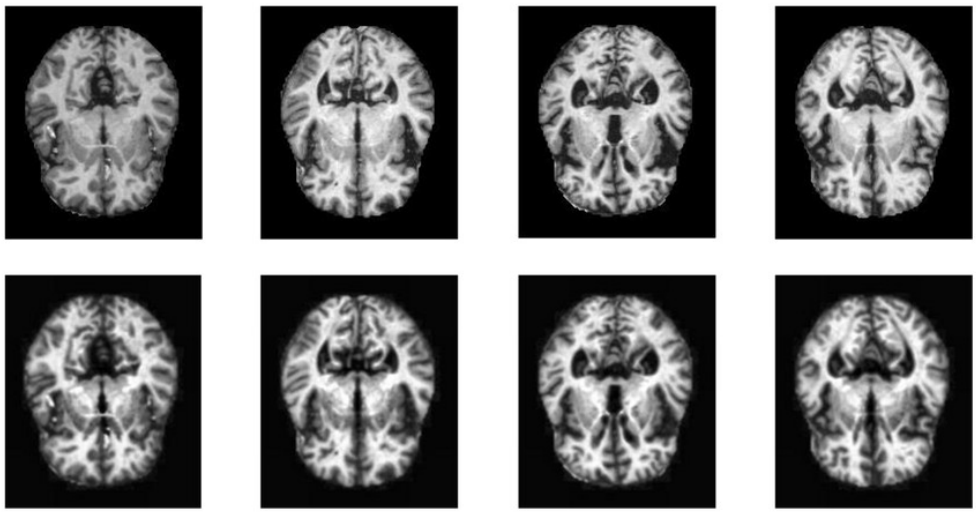
(a) No dementia (b) Very mild dementia (c) Mild dementia (d) Moderate dementia

Fig. 1. Samples of axial images in the Kaggle data set

Figure 2

Samples of image pre-processing techniques in the Kaggle data set

Samples of image pre-processing techniques in the Kaggle data set



No dementia (b) Very mild dementia (c) Mild dementia (d) Moderate dementia

Fig. 2. Samples of image pre-processing techniques in the Kaggle data set: the pictures in the top line are the MR images before the image pre-processing techniques, and the images in the bottom line are the MR images obtained after the pre-processing image methods.

Figure 3

End-to-end automatically DCNN-based AD classifier block diagram

End-to-end automatically DCNN-based AD classifier block diagram

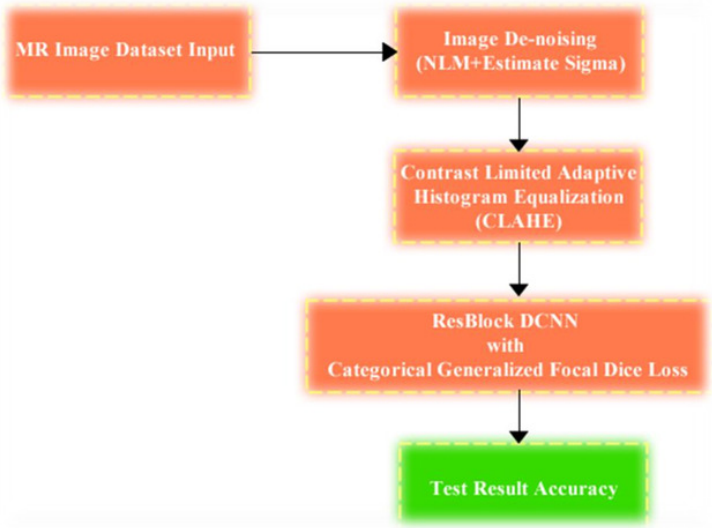


Fig. 3 End-to-end automatically DCNN-based AD classifier block diagram

Figure 4

Detailed representation of the proposed architecture

Detailed representation of the proposed architecture

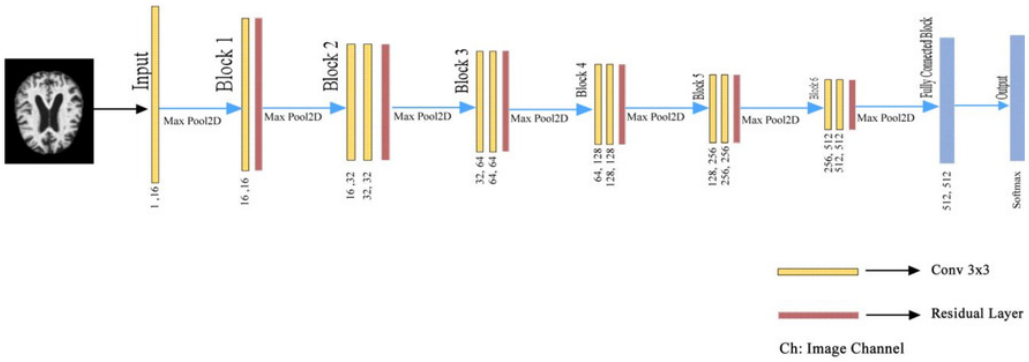


Fig. 4. Detailed representation of the proposed architecture

Figure 5

ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

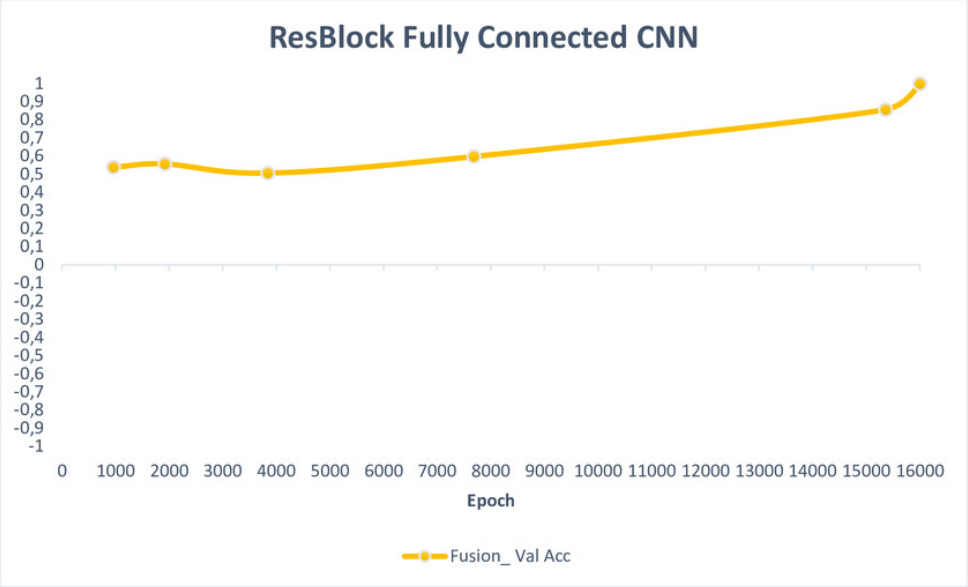


Fig. 5. ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

Figure 6

ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results

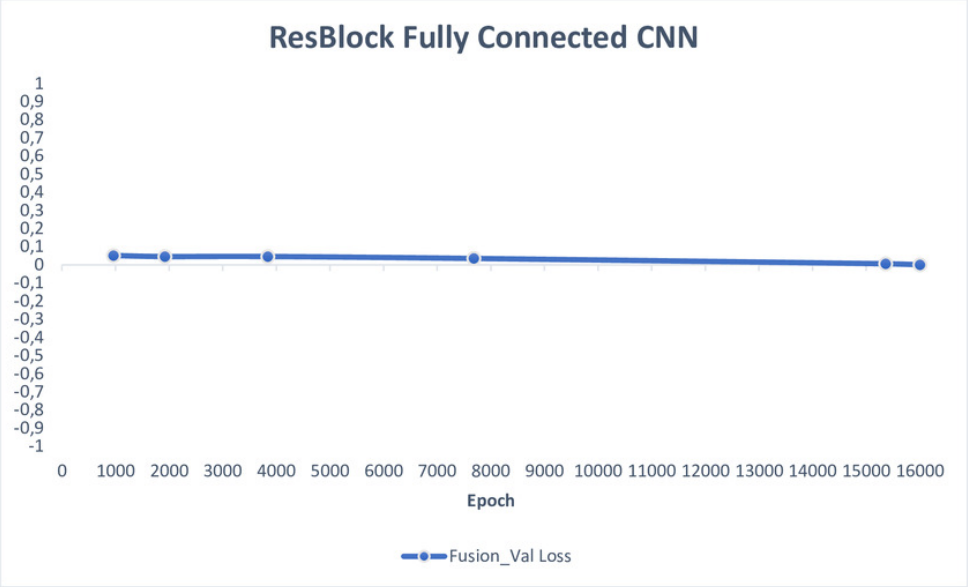


Fig. 6. ResBlock Fully Connected CNN Validation Loss and Validation Accuracy Results