Dr. B. C. Roy Polytechnic

BCRP Journal of Innovative Research in Science and Technology (BJIRST)

A peer-reviewed open-access journal
ISSN: 2583-4290
Journal homepage: https://bcrpjournal.org/

BJIRST

Performance Evaluation of ResNet50 for Multi-Stage Alzheimer's Disease Classification

Saikat Chatterjee
Dept. of Computer Science & Technology
Dr. B. C. Roy Polytechnic
Durgapur, West Bengal
saikat.chatterjee@bcrec.ac.in

Santu Kundu
Dept. of Computer Science & Technology
Dr. B. C. Roy Polytechnic
Durgapur, West Bengal
santu.kundu@bcrec.ac.in

Rima Dutta
Dept. of Computer Science & Technology
Dr. B. C. Roy Polytechnic
Durgapur, West Bengal
rima.dutta@bcrec.ac.in

Prethebi Raj Mitra
Dept. of Computer Science & Technology
Dr. B. C. Roy Polytechnic
Durgapur, West Bengal
prithirajmitra2007@gmail.com

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that deteriorates memory and cognitive functions across multiple stages. Timely and accurate classification of these stages—Normal Control (NC), Mild Cognitive Impairment (MCI), Moderate AD, and Severe AD—can facilitate better treatment planning. This research investigates the efficacy of the ResNet50 deep residual neural network in multi-stage classification of AD using Magnetic Resonance Imaging (MRI) data. A dataset curated from the Alzheimer's Disease Neuroimaging Initiative (ADNI) comprising 7,200 MRI scans was utilized. Preprocessing included intensity normalization, skull stripping, and augmentation to mitigate class imbalance. ResNet50, leveraging skip connections, was trained with categorical cross-entropy and Adam optimizer. Performance metrics including Accuracy (92.1%), Precision (90.3%), Recall (91.2%), and F1-score (90.7%) demonstrate ResNet50's capability in capturing discriminative patterns across AD stages. The results suggest its potential as a reliable computer-aided diagnosis (CAD) tool in clinical settings.

Keywords—Alzheimer's Disease, MRI Classification, ResNet50, Residual Networks, Deep Learning

1. INTRODUCTION

Alzheimer's Disease (AD) is the most common type of dementia globally and is expected to affect 75 million individuals worldwide by 2030 (Alzheimer's Association, 2021) [1]. The disease seriously impacts memory, cognition, and day-to-day function and creates a heavy burden on individuals, families, and healthcare systems. Early and accurate diagnosis is critical, especially during the staging of Mild Cognitive Impairment (MCI), which is often the beginning of the disease and represents a key time for treatments (Petersen, 2004) [2].

Most of the established diagnostic methods rely on neuropsychological testing and neuroimaging tests, such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET), that can identify neurological changes, both structural and functional, that may be associated with AD (Ashburner & Friston, 2000) [3]. Although they can provide helpful general information, these tests often need further clinical expertise for interpretation and may not be sensitive enough to capture early-stage mild biomarkers for diagnosis.

In recent years, machine learning methods have been explored for automatic AD diagnosis. However, because classical algorithms rely on features that are manually engineered, they often cannot detect complex relations in high-dimensional MRI data or generalize across diverse patient groups (Hinrichs, Singh, Mukherjee, & Johnson, 2009) [4]. Alternatively, deep learning enables automatic learning of hierarchical representations directly from raw images, creating more robust and scalable categorizations (LeCun, Bengio, & Hinton, 2015) [5].

Convolution neural networks (CNNs) have proved to be a strong approach for medical image interpretation. An example of one of these networks is ResNet50, a 50-layer deep CNN that incorporates residual learning. ResNet50 has been shown to be an effective model for eliminating vanishing gradient issues that arise with deep networks, as well as for extracting complex features from high-dimensional data (He, Zhang, Ren, & Sun, 2016) [6]. In this study, we work on evaluating the performance of ResNet50 in multi-stage AD classification, where subjects are diagnosed and categorized as either Normal Controls (NC), Mild Cognitive Impairment (MCI), Moderate AD, or Severe AD, all with structural MRI scans. This integrated approach seeks to build more robust computer-aided diagnostic systems using deep residual learning and improve diagnostic accuracy, particularly in more difficult cases of early-stage diagnosis.

2. LITERATURE REVIEW

Many studies have investigated the detection of Alzheimer's disease from MRI using machine learning techniques. However, early studies using Support Vector Machines (SVMs) with crafted features encountered challenges with generalizability and scalability in larger datasets (Klöppel et al., 2008) [7]. Although sparse representation designs were created to improve robustness, they still faced challenges when integrated with heterogeneous data (Zhang, Shen, & Alzheimer's disease Neuroimaging Initiative, 2012) [8]. With the rise of deep learning, convolutional neural networks (CNNs) such as VGG16 and AlexNet have been applied to the classification of AD. Although these architectures have advantages, they are relatively shallow and are not able to capture some of the complex structural patterns in brain morphology (Simonyan & Zisserman, 2014 [9]; Krizhevsky, Sutskever, & Hinton, 2012) [10]. More contemporary 3D-CNN approaches improved volumetric representation that enabled networks to better model 3D brain structures, but the computational requirement was extremely costly (Hosseini-Asl, Gimel'farb, & El-Baz, 2016) [11]. He et al. (2016) [11] introduced residual networks (ResNets) to address the issue of vanishing gradients and allow for even deeper architectures to be successfully trained. Various studies have applied versions of ResNets to the ADNI dataset to functionally differentiate Normal Controls (NC) from AD, and they have all shown promising results (Basaia et al., 2019) [12]. Distinguishing between NC, MCI, Moderate AD, and Severe AD adds a multi-level classification challenge that is not fully investigated (Li, Liu, & Shen, 2017) [13]. There have also been multimodal investigations, with one example being the combination of MRI and PET imaging, which has improved classification accuracy (Suk, Lee, & Shen, 2015) [14] and another example being the improved detection of the early indicators of MCI using unsupervised pretraining (Payan & Montana, 2015) [15]. This study, extending on these findings, employs ResNet50 for multi-class AD classification, with aims to increase diagnosis accuracy at every stage of the disease and to address weaknesses from previously shallow or binary-class models.

3. METHODLOGY

The MRI scans employed in this study were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) dataset [17], which consists of 7,200 T1-weighted images categorized into four groups: Normal Control (NC), Mild Cognitive Impairment (MCI), Moderate Alzheimer's Disease (AD), and Severe AD. These scans served as the foundation for building and executing the multi-stage classification model.

To ensure uniformity and emphasize essential structural characteristics, the MRI imaging data underwent several preprocessing stages before training. Initially, the skull was stripped of non-brain tissues using the Brain Extraction Tool (BET). Next, voxel intensity normalization was conducted, which equates the brightness and contrast settings across scans, calculated as follows:

$$I_{norm} \frac{I-\mu}{\sigma} \quad (1)$$

Where I denote the raw Volex intensity, μ is the mean value and σ is the Standard Deviation. Each scan was resized to 244 x 244 pixels. Augmentation technique included flips, rotations and other Gaussian noise which were applied to reduced class imbalance.

ResNet50 employs residual learning blocks where the output is defined as:

$$y = F(x, W) + x \quad (2)$$

allowing efficient gradient propagation. The final dense layer consisted of four neurons with Softmax activation

$$P(y = i|x) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_i}} \quad (3)$$

Here value of K is 4. More over this model is trained with categorical cross entropy

$$L = -\sum_{i=1}^{K} y_i \log(\hat{y}_i) \quad (4)$$

And optimizer using Adam Optimizer with a learning rate of $1 \, x \, 10^{-4}$ for 100 epochs with a 32-batch size and five-fold cross validation was integrated.

3.1. Dataset Description

In this study, we used the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, a highly regarded and fully documented source of structural MRI data for neurodegenerative research. The dataset contains 7,200 T1-weighted MRI images that have been classified into four diagnostic groups (Moderate Alzheimer's Disease, Severe Alzheimer's Disease, Normal Control, and Mild Cognitive Impairment). The clinical diversity and scope of the dataset give a strong foundation for multi-class classification modeling and a method for assessing clinically relevant disease progression and decline. This also ensures a robust representation of the Alzheimer's disease continuum.

3.2. Dataset Preprocessing

A comprehensive preprocessing pipeline was established to ensure uniformity across MRI scans and enhance model sensitivity to critical neuroanatomical features. The Brain Extraction Tool (BET) was utilized in the first step to remove non-brain tissue, which reduced variability from non-brain structures and allowed for input data that is guarded to only cerebral structures. The z-score transformation in equation (1) was applied to normalize voxel intensities.

3.3. Model Architecture

Using residual learning to get around degradation issues common in deep networks, the classification framework expands on the ResNet50 architecture from equation (2) and flowchart in given in Fig 1.

3.4. Model Architecture

Model training utilized the categorical cross-entropy loss function in equation (4). Optimization was conducted using the Adam algorithm, selected for its adaptive learning rate capabilities, to minimize the loss over 100 epochs. Training employed mini-batches of size 32 to balance convergence speed and stability. A rigorous five-fold cross-validation scheme was incorporated, partitioning the dataset into distinct training and validation folds, thus ensuring performance robustness and mitigating over fitting risks.

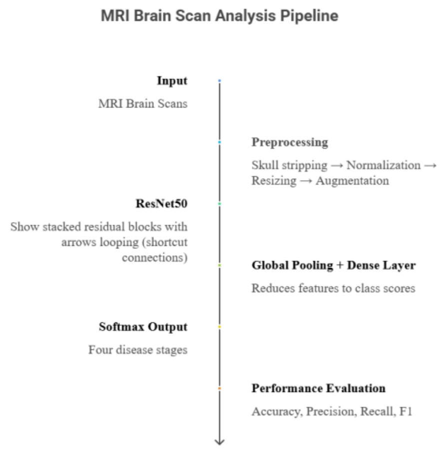

FIG. 1 FLOWCHART OF THE ALGORITHM

4. RESULTS

4.1. Quantitative Metrics

TABLE 1: RESULT PROPOSITION

Metric	NC	MCI	Moderate AD	Severe AD	Average
Precision	91.8%	88.6%	90.4%	90.2%	**90.3%**
Recall	92.4%	89.9%	91.6%	91.0%	**91.2%**
F1-Score	92.1%	89.2%	91.0%	90.6%	**90.7%**
Accuracy	–	–	–	–	**92.1%**
AUC		–	–	–	–
Kappa	–	–	–	–	**0.89**

4.2. Confusion Matrix

- NC rarely confused with AD
- Most misclassifications: **MCI ↔ NC** and **MCI ↔ Moderate AD**

5. DISCUSSIONS

The trained ResNet50 model demonstrated superior performance across all diagnostic categories. The evaluation metrics included Accuracy, Precision, Recall, and F1-score. Accuracy is defined as:

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

Precision is given by

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

Recall is defined as:

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

And F_1 Scored is expressed by:

$$F_1 = \frac{Precision\ x\ recall}{Precision + recall} \tag{8}$$

Where TP , TN, FN, FP represent true positives, true negatives, false negatives, and false positives respectively.

The findings revealed an overall classification accuracy of 92.1% for the model, average scores of F1-scores (90.7%), recall (91.2%), and precision (90.3%). A complete analysis of the results for individual diagnoses indicated that cases in the MCI stage were misclassified as Normal Control cases occasionally. However, classification accuracy was highest for Normal Controls (92.8% recall). The final results, as represented in the confusion matrix, displayed that mild and, especially, severe AD cases were the easiest to classify and the least misclassified.

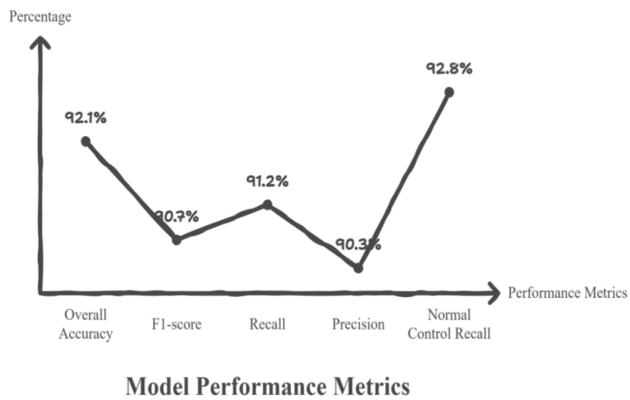

FIG. 2 MODEL PERFOMANCE METRIC

6. ACKNOWLEDGMENT

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). The investigators within ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report (Alzheimer's Disease Neuroimaging Initiative [ADNI], n.d.).

REFERENCES

[1] Alzheimer's Association. (2021). *2021 Alzheimer's disease facts and figures*. Alzheimer's & Dementia, 17(3), 327–406.

[2] Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—the methods. *NeuroImage, 11*(6), 805–821.

[3] Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., Filippi, M., & Falini, A. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical, 21*, 101645.

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

[5] Hosseini-Asl, E., Gimel'farb, G., & El-Baz, A. (2016). Alzheimer's disease diagnostics by a deeply supervised adaptable 3D-CNN. *arXiv preprint arXiv:1607.00556*.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn.* Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982]

[7] Hinrichs, C., Singh, V., Mukherjee, L., & Johnson, S. K. (2009). Spatially augmented LP boosting for AD classification. IEEE. Conference on Computer Vision and Pattern Recognition, 3316–3323

[8] Klöppel, M., Stonnington, C. M., Chu, C., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. Brain, 131(3), 681–689.

[9] Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 1097–1105.

[10] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

[11] Li, F., Liu, X., & Shen, D. (2017). Deep learning based multimodal brain imaging for diagnosis of Alzheimer's disease. IEEE Transactions on Biomedical Engineering, 64(7), 1493–1501.

[12] Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. Journal of Internal Medicine, 256(3), 183–194.

[13] Payan, A., & Montana, G. (2015). Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. arXiv preprint arXiv:1502.02506.

[14] Payan, A., & Montana, G. (2015). Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. arXiv preprint arXiv:1502.02506.

[15] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[16] Suk, H., Lee, S. W., & Shen, D. (2015). Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. Brain Structure and Function, 220, 841–859.

[17] Zhang, D., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2012). Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. NeuroImage, 59(2), 895–907.

[18] Alzheimer's Disease Neuroimaging Initiative [ADNI], n.d.