SMOTE-ADNet: 향상된 CNN 과 SMOTE 를 활용한 알츠하이머병 및 초기 단계 정확한 분류

손소영[1], 추현승[2]
[1]성균관대학교 전자전기컴퓨터공학과 박사과정
[2]성균관대학교 전자전기컴퓨터공학과 교수

alvy.sun@g.skku.edu, choo@skku.edu

SMOTE-ADNet Leveraging Enhanced CNN and SMOTE for Accurate Classification of Alzheimer's Disease and Early Stages

Sun Xiaoying[1], Hyunseung Choo[1]
[1] Dept. of Electrical and Computer Engineering, Sungkyunkwan University

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by gradual cognitive decline and memory loss, with subtle changes in brain structure that make accurate classification particularly challenging. This study presents SMOTE-ADNet, an innovative Convolutional Neural Network (CNN) model designed to enhance classification performance for Alzheimer's disease by integrating advanced CNN techniques with the Synthetic Minority Over-sampling Technique (SMOTE). The SMOTE-ADNet architecture includes multiple convolutional layers, dropout regularization, and a final dense layer optimized for multi-class classification, aimed at differentiating between five stages of Alzheimer's disease: Alzheimer's disease (AD), Cognitive Normal (CN), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), and Mild Cognitive Impairment (MCI). Given the challenge of distinguishing between subtle variations in brain structure during these stages, SMOTE-ADNet effectively balances the dataset using SMOTE and leverages advanced CNN layers to achieve a remarkable accuracy of 98%. This result demonstrates the model's capability to manage the inherent difficulty of classifying subtle structural differences and its potential for improving diagnostic precision and aiding early intervention in Alzheimer's disease.

1. Introduction

Alzheimer's disease (AD) is a progressive neuro degenerative disorder and the most common cause of dementia, affecting millions of people worldwide. As of 2020, over 55 million people globally were living with dementia, with the majority of cases attributed to Alzheimer's disease. By 2050,this number is projected to reach nearly 139 million, emphasizing the urgent need for early detection and intervention (Alzheimer's Disease and Dementia). Early and accurate diagnosis is crucial for delaying disease progression and improving the quality of life for those affected, as well as easing the burden on caregivers and healthcare systems.

Current diagnostic approaches for AD, which rely on clinical assessments, cognitive tests, and brain imaging, often face challenges due to the subtle and gradual onset of symptoms, especially in the early stages of the disease. There are structural changes in the brain during the progression from mild cognitive impairment (MCI) to advanced Alzheimer's, but these changes are often subtle and difficult to detect. This makes classification between disease states particularly challenging for both clinicians and automated systems.

Recent advances in artificial intelligence (AI) and deep learning have shown promising results in medical image analysis. Convolutional neural networks (CNNs) have become a popular choice for tasks such as image classification due to their ability to capture complex spatial patterns in data. However, many medical datasets suffer from class imbalance, as in the case of Alzheimer's disease datasets, where advanced stages of the disease are overrepresented compared to early-stage or healthy samples. This imbalance can lead to biased models that underperform in predicting minority classes.

To address this, we introduce SMOTE-ADNet, a CNN-based model enhanced by the Synthetic Minority Over-sampling Technique (SMOTE) to tackle the issue of class imbalance. SMOTE generates synthetic samples of underrepresented classes, improving the model's ability to generalize across all stages of the disease. By leveraging CNNs' powerful feature extraction capabilities and SMOTE's data handling approach, our model aims to improve classification accuracy across five distinct stages of Alzheimer's disease, ranging from cognitively normal (CN) to advanced AD.

This research not only contributes to the growing body of work in automated AD classification but also highlights the potential of integrating data augmentation techniques like SMOTE in healthcare applications. The model's enhanced performance in distinguishing subtle brain structural differences at various stages of AD progression showcases its utility for real-world clinical diagnostics. Our study aims to provide a robust and scalable solution for improving early-stage diagnosis, which is critical for effective disease management.

2. Related Work

Recent advancements in Alzheimer's disease (AD) classification increasingly utilize machine learning and deep learning techniques to enhance diagnostic accuracy. Early approaches relied on traditional algorithms and hand-crafted features, but these were limited by feature extraction quality and brain complexity. The rise of convolutional neural networks (CNNs) has revolutionized this field by automating feature extraction and classification from raw imaging data.

Related works such as Arafa's deep learning framework for early diagnosis of AD on MRI images (2024) showcase the utility of CNNs in analyzing MRI data, highlighting improvements in diagnostic accuracy through enhanced feature extraction[1]. Similarly, Chen's graph neural network approach (2023) introduces a learnable subdivision technique for functional brain network analysis, offering greater interpretability in cognitive disorder diagnosis[2]. These studies underscore the significance of deep learning advancements and inform the methods explored in this research.

Despite these improvements, class imbalance remains a challenge. AD datasets often have disproportionate representations of disease stages, leading to biased models. Techniques like Synthetic Minority Over-sampling Technique (SMOTE) address this by generating synthetic samples for underrepresented classes[3], improving model fairness and generalization. Integrating SMOTE with CNNs, as explored in this study, provides a promising solution to tackle class imbalance and enhance the robustness of AD classification models, potentially offering more reliable diagnostic tools for early detection and effective management of Alzheimer's disease.

3. Methodology

This section outlines the methodology of our study, focusing on image preprocessing, model architecture, and

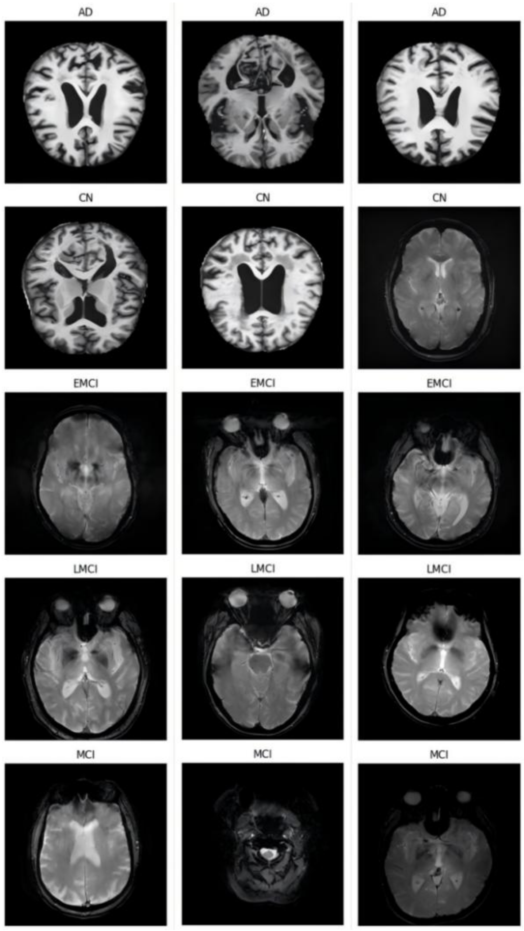

Figure 1. Random Sample Images from Each Alzheimer Class

training and evaluation. We begin with preprocessing techniques that ensure data uniformity and quality. Next, we describe the SMOTE-ADNet model architecture for feature extraction and classification. Finally, we discuss the training and evaluation strategies used to optimize and assess model performance.

3.1. Dataset

The dataset used for this study is derived from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and comprises a total of 16,200 images. These images are categorized into five distinct classes, representing different stages of Alzheimer's disease and cognitive states:

CN	*EMCI*	*LMCI*	*MCI*	*AD*	*Total*
7,430	*240*	*72*	*922*	*7,536*	*16,200*

<Table 1> ADNI Dataset

The dataset's distribution shows a significant imbalance, with AD and CN classes being well-represented, while MCI, EMCI, and LMCI classes have fewer samples. To address this imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. SMOTE generates synthetic samples for the minority classes, resulting in the following balanced distribution:

CN	*EMCI*	*LMCI*	*MCI*	*AD*	*Total*
7,536	*7,536*	*7,536*	*7,536*	*7,536*	*37,680*

<Table 2> Post-SMOTE ADNI Dataset

Post-SMOTE, the dataset was expanded to a total of 37,680 images, with each class now containing 7,536 images. This balanced dataset improves model training by providing a fair representation of all classes, enhancing the model's ability to generalize and perform effectively across different stages of Alzheimer's disease.

3.2. Image Preprocessing

Resizing and Normalization: To ensure uniformity across the input data, images are resized to a fixed dimension of 224 ×224 pixels. This standardization is crucial for compatibility with the convolutional neural network (CNN) model, which requires a consistent input size. Pixel values are normalized to a [0,1] .

Data Augmentation Techniques: To mitigate overfitting and enhance the model's generalization capability, various data augmentation techniques are employed. These include random rotations, horizontal and vertical flips, and random shifts.

3.3. Model Architecture

The SMOTE-ADNet model is specifically designed to extract relevant features from brain scan images and classify them into five Alzheimer's disease stages: AD, CN, EMCI, LMCI, and MCI. The architecture emphasizes capturing subtle differences in brain structures to enhance classification accuracy.

The model comprises several key components:

● **Convolutional Layers (Conv2D):** These layers apply filters to detect local features within the input images.

● **Max Pooling Layers (MaxPool2D):** Max pooling reduces the spatial dimensions of the feature maps by selecting the maximum value within a defined window.

● **Dense Layers:** Fully connected layers aggregate the features extracted by the convolutional layers.

● **Dropout Layers:** Dropout is used to prevent overfitting by randomly setting a fraction p of the neurons to zero during each training iteration.

● **Activation Functions and Output Layer:** The ReLU (Rectified Linear Unit) activation function is applied in the hidden layers to introduce non-linearity. In the output layer, the softmax function is employed to produce a probability distribution over the five target classes.

3.4. Training and Evaluation

Traning: The proposed model is trained using the Adam optimizer with categorical cross-entropy loss. The training process is designed to be adaptive, employing early stopping to halt training when the validation performance plateaus. The model utilizes a batch size of 32, and the number of epochs is determined based on early stopping criteria. A learning rate scheduler dynamically adjusts the learning rate during training to optimize convergence.

During training, the model achieved a final loss of 0.0176 and an accuracy of 99.41% on the training set, while the validation set resulted in a loss of 0.0969 and an accuracy of 98.12%. These metrics indicate a high level of performance and generalization ability.

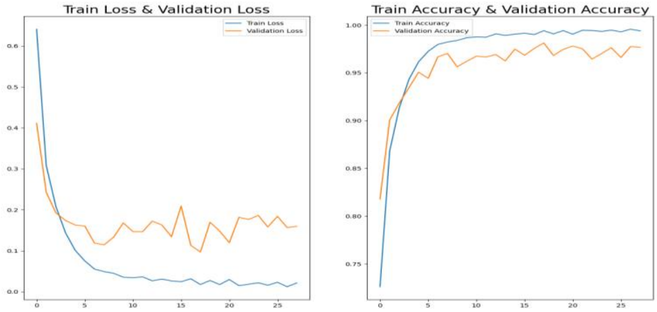

Figure 2. Train and Validation Loss & Accuracy

The dataset is divided into training and testing sets with an 80-20 split. To address class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) is applied, which generates synthetic samples for minority classes. The SMOTE algorithm creates new samples using the following formula:

$$x_{new} = x_{base} + \lambda(x_{neighbor} - x_{base}) \quad (3)$$

where x_{base} *is the original sample,* $x_{neighbor}$ *is a selected nearest neighbor, and* λ *is a random number between 0 and 1.*

Evaluation: The performance of the model is assessed using several key metrics to ensure a comprehensive evaluation of its efficacy.

Accuracy measures the overall correctness of the model's predictions. Accuracy provides a general measure of how often the model's predictions match the true labels, reflecting its overall reliability.

The evaluation metrics used in this study provide a thorough assessment of the model's ability to classify Alzheimer's disease states. The confusion matrix offers a detailed breakdown of the model's performance across the five classes—AD, CN, EMCI, LMCI, and MCI—by showing the counts of true positives and false positives for each category. This matrix is essential for understanding how well the model differentiates between the classes and identifying areas for potential improvement.

Complementing the confusion matrix, the classification report includes precision, recall, and F1-score metrics for each class are defined as follows:

$$\begin{cases} \text{Precision} = \dfrac{\text{TP}}{\text{TP} + \text{FP}} \\ \text{Recall} = \dfrac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{F1} - \text{Score} = 2 \cdot \dfrac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \end{cases} \quad (4)$$

where:

TP denotes True Positives, the number of correctly predicted positive instances.

FP denotes False Positives, the number of incorrectly predicted positives instances.

FN denotes False Negatives, the number of missed positive instances.

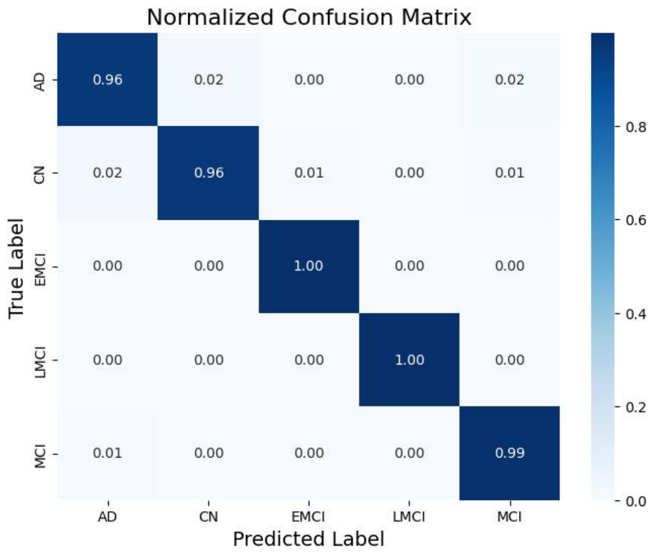

Figure 3. Normalized Confusion Matrix

These metrics collectively confirm that the model performs with high accuracy and balanced effectiveness across all categories. Precision values indicate that the model's positive predictions are reliable, while recall values demonstrate its capability to identify relevant instances. The F1-scores further highlight the model's balanced performance. Overall, these evaluation tools underscore the model's robustness, its ability to handle class imbalance effectively, and its capacity to deliver accurate predictions crucial for classifying Alzheimer's disease stages.

4. Results

The SMOTE-ADNet model achieved impressive results in classifying Alzheimer's disease states, as evidenced by both the test loss and accuracy metrics. The model recorded a test loss of 0.0826 and a test accuracy of 0.9823, demonstrating its high reliability and effectiveness.

The detailed evaluation further highlights the model's performance. The classification report reveals consistently high precision, recall, and F1-scores across all classes. Specifically, precision values range from 0.97 for AD to 1.00 for LMCI, while recall values span from 0.95 for CN to 1.00 for EMCI and LMCI. The F1-scores also show strong performance, ranging from 0.97 to 1.00. Overall accuracy stands at 98%, reflecting the model's robust ability to correctly classify instances of Alzheimer's disease. The macro and weighted averages reinforce this finding, both at 0.98, confirming the model's balanced performance across different classes.

5. Conclusion

This study demonstrates that the SMOTE-ADNet model achieves outstanding performance in classifying Alzheimer's disease states with notable precision and reliability. The model excels across all five classes, with precision ranging from 0.97 for AD to 1.00 for LMCI, and recalls of 0.95 for CN and 1.00 for EMCI and LMCI. These metrics reflect a robust ability to accurately differentiate between the various stages of Alzheimer's disease. The high F1-scores and an overall accuracy of 0.98 further confirm the model's effectiveness, illustrating its capability to handle both class imbalance and precise classification.

The promising results underscore the potential of SMOTE-ADNet as a significant advancement in medical image classification. Moving forward, our research will focus on further enhancing the model's performance and exploring innovative methodologies to broaden its applicability. We are preparing to submit these findings to a peer-reviewed journal, aiming to contribute valuable insights to the field and support ongoing efforts in the early detection and classification of Alzheimer's disease.

	precision	*recall*	*f1-score*	*support*
AD	*0.97*	*0.96*	*0.96*	*1,131*
CN	*0.97*	*0.96*	*0.97*	*1,130*
EMCI	*0.99*	*1.00*	*0.99*	*1,130*
LMCI	*1.00*	*1.00*	*1.00*	*1,131*
MCI	*0.98*	*0.99*	*0.98*	*1,130*
accuracy			*0.98*	*5,652*
macro avg	*0.98*	*0.98*	*0.98*	*5,652*
weighted avg	*0.98*	*0.98*	*0.98*	*5,652*

<Table 3> ADNI Classification Report

References

[1] Doaa Ahmed Arafa etc. "A deep learning framework for early diagnosis of Alzheimer's disease on MRI images" Multimedia Tools and Applications, 83(2), 3767--3799

[2] Dongdong Chen etc. "Learnable Subdivision Graph Neural Network for Functional Brain Network Analysis and Interpretable Cognitive Disorder Diagnosis" MICCAI Canada 2023, pp 56–66

[3] Blagus, Rok, Lusa, Lara "SMOTE for high-dimensional class-imbalanced data" BMC bioinformatics, volume 14, pp 1–16

ACKNOWLEDGEMENT

This work was supported in part by the BK21 FOUR Project (50%) and the Korea government (MSIT), IITP, Korea, under the ICT Creative Consilience program (RS-2020-II201821, 25%), Development of Brain Disease (Stroke) (RS-2024-00459512, 25%).