Smart Media Journal / Vol.13, No.10 / ISSN:2287-1322
https://dx.doi.org/10.30693/SMJ.2024.13.10.28

# Comparison and analysis of CNN models to Address Skewed Data Issues in Alzheimer's Diagnosis

Faizaan Fazal Khan, Goo-Rak Kwon

**Abstract**

Alzheimer's disease is a form of dementia that can be managed by identifying the disease in its initial phases. In recent times, numerous computer-aided diagnostic techniques utilizing magnetic resonance imaging (MRI) have demonstrated promising outcomes in the categorization of Alzheimer's disease (AD). The OASIS MRI dataset was utilized which has 80,000 brain MRI images. It is suggested to resample this dataset as it is highly imbalanced and posed a challenge in preventing bias toward majority class while employing the convolution neural network (CNN) model for classification. This paper examines and extracts patterns and features of 461 patients taken from the OASIS dataset. The research has aimed at utilizing the Base Model of EfficientNetV2B0 with custom classification layers, and simplified custom CNN model, also exploring Multi-class classification across four distinct classes: Non-Demented, Very Mild Demented, Mild Demented, Moderate Demented in addition to binary classification as Non-Demented and treating other classes as demented. Furthermore, different dataset sizes were experimented with 5,000 and 20,000 for each class to be discussed in this paper. The experiment results indicate that EfficientNetV2B0 achieved the accuracy of 98% in binary classification, 99% in multiclass. Whereas custom sequential CNN model in multiclass classification presents the accuracy of 96% for 20,000 dataset size and 98% for 80,000 dataset size.

Keywords: Alzheimer's disease | Non-Demented | Very Mild Demented | Mild Demented | Moderate Demented | Convolution Neural Network | EfficientNetV2B0 | MRI | OASIS dataset

## I. INTRODUCTION

AD affects 10% of adults over 65, 50% over 85. Need for early prediction and detection techniques are mentioned in the report[1]. Approximately 6.5 million Americans age 65+ have dementia caused by Alzheimer's. Approximately 16 billion hours of care without compensation will be supplied in 2021[2]. 6.7 million Americans over the age of 65 living with dementia caused by Alzheimer's. Around 13.8 million will be impacted by 2060 without medical improvements[3]. Alzheimer's disease detection is achieved through a deep learning approach using a 3D convolutional neural network (3D CNN) with 78.07% accuracy, also converting 3D images to 2D for analysis[4]. Much study is being conducted but there is no perfect solution, hence the only approach is postponed aging process of the condition. Therefore, it's crucial to identify

* This study was supported by research funds from Chosun University, 2024

Manuscript : 2024.07.08
Revised : 2024.08.20
Confirmation of Publication:2024.09.18
Corresponding Author : Goo-Rak Kwon e-mail : grkwon@chosun.ac.kr

Alzheimer's. Alzhiermer's can be divided into different stages, according to dataset we have it has following stages (Non-demented, very mild demented, mild demented, and moderate demented). In the past few years, convolutional neural networks have delivered greater results in the image classification field. In the current study we dealt with unbalanced dataset while exploring and comparing results of EfficientNetV2B0 and custom CNN model with focus on time consumption and improvements of accuracy and other params[5]-[8].
The dataset was separated into training and testing to avoid data leakage, so that we can test model on unseen data, then it was resampled to accommodate class imbalance using random sampling. Also, to keep the learning process stable and efficient normalization was performed.
To achieve better accuracy different parameters and dataset resampling sizes were used, we will discuss 5000 and 20000 results in this study.

# II. METHODOLOGY

## 2.1 OASIS-1 Dataset

For our study, OASIS-1(Open Access Series of Imaging Studies)[9]. Cross-sectional MRI data was used, OASIS-1 is a publicly available dataset created to facilitate studies on Alzheimer's disease progression and related cognitive impairments. It includes a comprehensive array of neuroimaging data such as MRI scans, along with associated clinical and demographic information. The data from OASIS-1 are from OASIS data repository, providing a valuable resource for analyzing and understanding Alzheimer's disease through neuroimaging analysis.

## 2.2 Data demographics

In our study, the referenced OASIS-1 dataset consists of 416 subjects, patient data classification was done based metadata and Clinical Dementia Rating (CDR). The z-axis was used to slice the brain pictures into 256 parts, with 100- 160 slices per patient chosen at random. This strategy resulted in a 80,000 2D images for study. The class distribution of resulting dataset in shown in Figure 1.

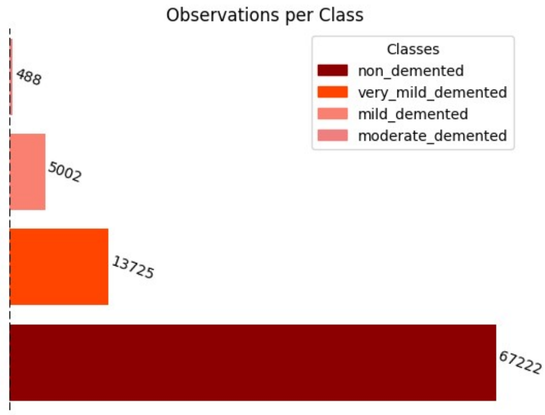


Figure 1 Number of resulting images per class

The dataset is highly skewed and poses a challenge of avoiding overfitting and bias of majority Non-Demented as seen above.

## 2.4 Proposed Model

In the proposed study, we have assessed and compared two alternative models, EfficientNetV2B0 and custom CNN sequential model for the categorization of Alzheimer's illness. For preventing overfitting and increasing accuracy appropriate measures were employed. Balanced data, ensured no data leakage through splitting testing data from training and validation data in earlier stages, added dropout layer, also used earlier stopping with patience of 5.

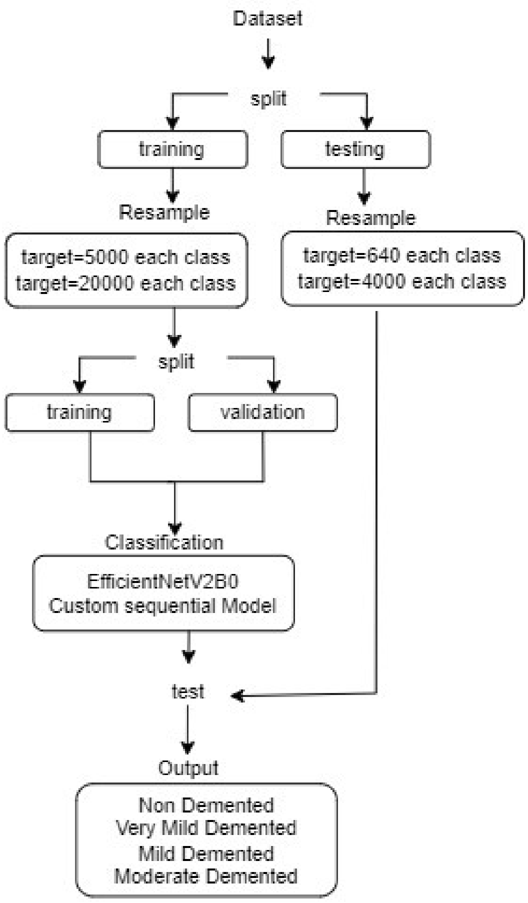


Figure 2 Architecture diagram showing the flow

The flow followed is shown in Figure 2. The data obtained from OASIS dataset if first split into testing and training and resampling was performed using random sampling and choices to balance dataset. We explored small and large dataset from 20,000 images (5000 each class) and 80,000 (20000 each class).

Output of multiclass classification is depicted in figure above but for binary output would be either Non-Demented or Demented.

EffiecientNetV2B0 model is pretrained on ImageNet dataset so we implied transfer learning by excluding top classification layer. And add custom layers to reduce and avoid overfitting and increase accuracy. The model is summarized in Figure 3.

For Custom CNN Sequential Model, we kept the layer simple as dataset was resampled, model may see same data again and again that may lead to overfitting. Custom CNN sequential Model is shown in Figure 4.

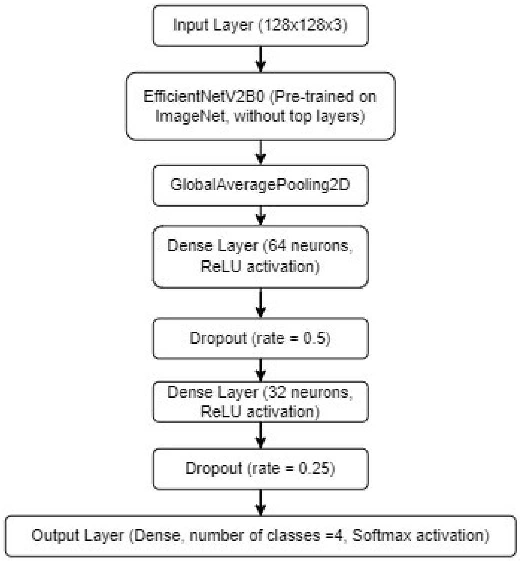


Figure 3 EfficientNetV2B0 Model Summary

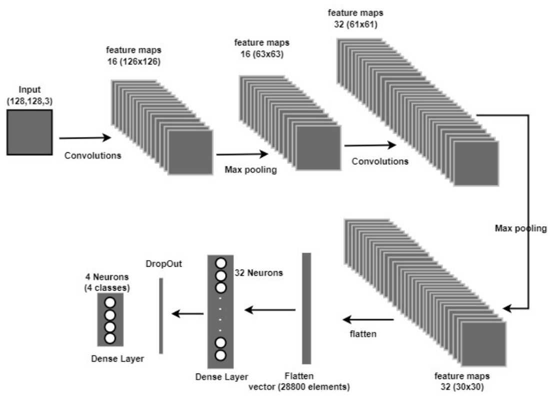


Figure 3 Custom Sequential Model Architecture

## III. EXPERIMENT RESULT & DISCUSSION

### 3.1 Training and Evaluation Parameters

In this study we have used the multi−class classification and binary class classification also to evaluate the model performance by using the EfficientNetV2B0 and Custom Sequential CNN model.

We can see model loss of custom CNN model for 5000 each sample size in figure 5. which uses Categorical Cross−Entropy, given by

$$Loss = -\sum_{i=1}^{N} y_i \log(p_i) \quad (1)$$

Where $y_i$ s the actual label (one−hot encoded), and $p_i$ is the predicted probability for the class.

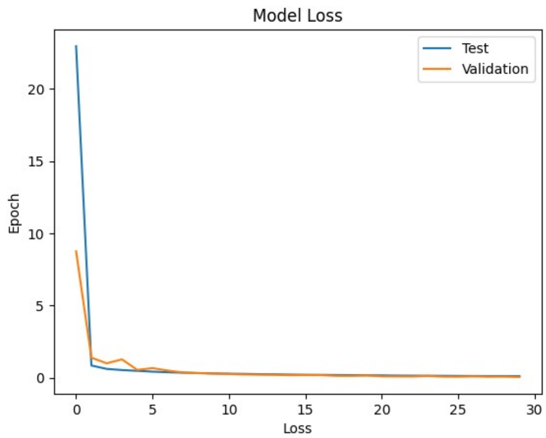


Figure 4 Model loss of custom CNN model

We can see that validation and testing loss decrease side by side and converge well.

EfficientNetV2B0 using same loss function for 5000 each sample size, Model accuracy and model loss are shown in Figures 6 and 7 respectively.

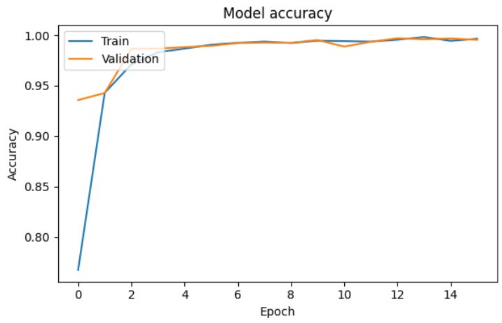


Figure 5 Model Accuracy of EffiecientNetV2B0 for 5000 each class sample size

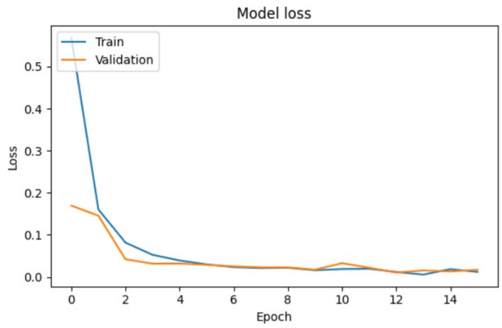


Figure 6 Model Loss of EffiecientNetV2B0 for 5000 each class sample size

Overall performance of model can be seen from above Figure 6. And Figure 7 shows good initial learning and shows satisfactory convergence with stabilized and minimal losses. We can also perceive the model is generalizing well based on the narrow gap between training and validation loss which suggests little overfitting.

Table 1. Evaluation Parameters

| | Prediction classification | | | | |
|---|---|---|---|---|---|
| **Actual classification** | classes | ND | V-MID | MI-D | MO-D |
| | ND | **TP** | $F_{AL}$ | $F_{AE}$ | $F_{AH}$ |
| | V-MID | $F_{LA}$ | **TP** | $F_{LE}$ | $F_{LH}$ |
| | MI-D | $F_{EA}$ | $F_{EL}$ | **TP** | $F_{EH}$ |
| | MO-D | $F_{HA}$ | $F_{HL}$ | $F_{HE}$ | **TP** |

We used Confusion matrix for evaluation, in ND(Non Demented), V−MID(Very Mild Demented), MI−D(Mild Demented),

M-OD(Moderate Demented) are the classes.
All of these classifiers are accountable for the prediction of the outcome class appropriately. To calculate we have utilized the parameters for the number of valid output form n the matrix by true positive(TP), true negative(TN), false positive(FP), and false negative(FN). To understand this by utilizing Table 1.
Other parameters which are used to predict are accuracy, precision, and recall from equations 1, 2, and 3.

$$\text{Accuracy} = \frac{\textit{Number of Correct Predictions}}{\textit{Total Number of Predictions}} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

And for binary classification we only have Non-Demented and Demented data, we have also merged all classes apart from Non-Demented and labeled them as demented to resolve class imbalance issue without resampling.
In medical Diagnosis where accuracy may not always be the most appropriate metric, we use precision and recall score in addition to accuracy for better evaluation.

### 3.2 Experiment Result

The experiment was conducted in the Python 3.11 environment. All the strategies have achieved high performance, with accuracy rates of 99%, 98%, 98%, and 96%. Due to the significant skewness in the data of classes, the accuracy may not be reliable. In addition, precision and recall are also provided as shown in Table 2. We tried to find best results with different sizes of resampling as each model complexity is different and for small size samples complex models were easy to get overfitted.

Table 2. Classification Results

| Classification | classes | Data size total | Acc % | Pre % | Rec % |
|---|---|---|---|---|---|
| EfficientNetV2 B0 | 4 | 20K | 99% | 99% | 99% |
| EfficientNetV2 B0 | 2 | 80K | 98% | 98% | 99% |
| Custom model | 4 | 80K | 98% | 98% | 98% |
| Custom model | 4 | 20K | 96% | 97% | 97% |

Acc% = Accuracy percentage,

Rec% = Recall percentage,

Pre% = precision percentage.

### 3.3 Discussion

Researchers have used several research methodologies to classify AD using deep learning algorithms. Acharya et al[10] have proposed transfer learning models for Alzheimer's multiclass classification like VGG16, ResNet-50 and AlexNet along with convolution neural networks.
In classification tools like free surfer are being used for preprocessing and feature extraction that help in disease detection[11].

## IV. CONCLUSION

We conducted a comparative analysis of CNN models, each consisting of a single convolutional layer, with the purpose of classifying Alzheimer's disease on highly skewed data. According to the results in Table2, it seems that the model is not suffering from overfitting and is performing well on both the training and validation datasets. The high level of test accuracy provides further evidence of the model's strong performance. OASIS processed data was used for classification. Accuracy of the research is contingent upon the specific kind of data used. In the future, we will persist in suggesting categorization models that are based on various datasets, while also taking into account other forms of medical imaging data.

## ACKNOWLEDGMENT

This study was supported by research funds from Chosun University, 2024. Data collection and sharing for this project was funded by the OASIS (Open Access Series of Imaging Studies) [6]. Available at: https://sites.wustl.edu/oasisbrains/.

## REFERENCES

[1] R. Khilar, T. Subetha, and P. Kayal, "A report on early prediction, detection & diagnosis of Alzheimer's disease," *J. Stat. Manag. Syst.*, vol. 26, no. 1, pp. 213–223, 2023.

[2] "2022 Alzheimer's disease facts and figures," *Alzheimers Dement.*, vol. 18, no. 4, pp. 700–789, Apr. 2022.

[3] K. Dhana, T. Beck, P. Desai, R. S. Wilson, D. A. Evans, and K. B. Rajan, "Prevalence of Alzheimer's disease dementia in the 50 US states and 3142 counties: A population estimate using the 2020 bridged-race postcensal from the National Center for Health Statistics," *Alzheimers Dement.*, vol. 19, no. 10, pp. 4388–4395, Oct. 2023.

[4] C. Wang, "A Review on 3D Convolutional Neural Network," in *2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA)*, Shenyang, China: IEEE, pp. 1204–1208, Jan. 2023.

[5] V. Ramineni, G.-R. Kwon, "An Implementation of Effective CNN Model for AD Detection," Korean Institute of Smart Media, Vol. 13, No. 6, Jun. 2024.

[6] Bijen Khagi, G.-R. Kwon, "Analysis of Weights and Feature Patterns in Popular 2D Deep Neural Networks Models for MRI Image Classification," J. Multimed Inf. Syst., Vol. 9, No. 3, pp. 177-182, Sep. 2022.

[7] V. Ramineni, G.-R. Kwon, "A Comparative Study of the CNN Model for AD Diagnosis," Korean Institute of Smart Media, Vol. 12, No. 7, pp. 52-58, Aug. 2023.

[8] V. Ramineni, G.-R. Kwon, "Diagnosis of Alzheimer's Disease using Wrapper Feature Selection Method," Korean Institute of Smart Media, Vol. 12, No. 3, pp. 30-37, Apr. 2023.

[9] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented

Older Adults," *J. Cogn. Neurosci.*, vol. 19, no. 9, pp. 1498–1507, Sep. 2007.

[10] H. Acharya, R. Mehta, and D. Kumar Singh, "Alzheimer Disease Classification Using Transfer Learning," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India: IEEE, pp. 1503–1508, Apr. 2021.

[11] R. Prajapati and G.-R. Kwon, "A Binary Classifier Using Fully Connected Neural Network for Alzheimer's Disease Classification," *J. Multimed. Inf. Syst.*, vol. 9, no. 1, pp. 21–32, Apr. 2022.

Authors

Faizaan Fazal Khan

He has completed his bachelor' s degree at the department of CS & IT from Air University, Islamabad, Pakistan, in 2023. Currently, he is pursuing his Master studies in the department of Information and Communication Engineering at Chosun University, Gwangju, Republic of Korea. His research interests involve digital image processing, Machine learning utilizing Medical Imaging.

Goo-Rak Kwon

He got a Ph.D. from the Department of Mechatronic Engineering, Korea University, in 2007. He has been a professor at Chosun University, since 2017. His research focus includes medical image analysis, A/V signal processing, video communication, and applications. He is a senior member of the IEEE.