Jurnal AL-AZHAR INDONESIA SERI SAINS DAN TEKNOLOGI
Vol. 11, Nomor 01, January 2026, hal. 61-69
DOI http://dx.doi.org/10.36722/sst.v11i1.5311 https://jurnal.uai.ac.id/index.php/SST

An EfficientNetV2-Based for Alzheimer's Disease Classification

M Sadewa Wicaksana Wibowo[1*], Khairul Umam[2]

[1] Medical Informatics, Faculty of Health Sciences, Universitas Muhammadiyah Lamongan,
[2] Computer Engineering, Faculty of Science, Technology, and Education, Universitas Muhammadiyah Lamongan, Jl. Plalangan No.KM, RW.02, East Java, 62218.

Corresponding author/E-mail: sadewawicaksana@umla.ac.id

Abstract – **In Indonesia, Alzheimer's disease has emerged as a critical public health priority. This neurodegenerative disorder is characterized by the gradual erosion of memory, linguistic capabilities, and problem-solving skills resulting from irreversible neuronal damage. Magnetic Resonance Imaging (MRI) is commonly used for early diagnosis; however, manual interpretation of MRI scans is time-consuming and subject to inter-observer variability among medical professionals. Recent advances in artificial intelligence have enabled automated analysis of MRI images for Alzheimer's disease detection, yet many existing approaches rely on deep learning architectures with high computational complexity. To address this limitation, this study proposes a lightweight deep convolutional network based on EfficientNetV2 for Alzheimer's disease classification using brain MRI images. Data augmentation techniques, including random rotation, affine transformation, horizontal and vertical flipping and normalization are applied to enhance model generalization. Two EfficientNetV2 variants, EfficientNetV2_s and EfficientNetV2_m, are evaluated and compared using accuracy, precision, recall, and F1-score metrics. Experimental results demonstrate that EfficientNetV2_s achieves superior performance, attaining an accuracy, precision, recall, and F1-score of approximately 0.90, while EfficientNetV2_m achieves corresponding values of approximately 0.81, indicating lower generalization capability. These results confirm that the smaller EfficientNetV2_s model provides more accurate and reliable classification performance despite its reduced computational complexity.**

Keywords - *Alzheimer's Disease, Classification, Convolutional Neural Networks, Deep Learning.*

INTRODUCTION

One of the most urgent public health issues is dementia, which has gained significant attention in international study. It is typified by a multifaceted, gradual loss of cognitive abilities brought on by biological injury to the central nervous system. Millions of people worldwide suffer with Alzheimer's disease, the most prevalent kind of dementia. Due to damage to neurons in certain brain areas, Alzheimer's disease is characterized by a gradual decline in memory, language, problem-solving skills and other cognitive capacities [1]. After the age of 60, the prevalence of Alzheimer's disease in the US doubles every five years, rising from around 1% in those between the ages of 60 and 64 to almost 40% in those 85 and older [2]. According to estimates, there were 1.2 million Alzheimer's patients in Indonesia in 2016; by 2030, that number is expected to increase to 2 million, and by 2050, it is expected to reach 4 million [3]. Although several therapy approaches have been found to decrease the course of the disease, early identification is still crucial to improve patient outcomes. Using Magnetic Resonance Imaging (MRI) to examine structural alterations in the brain is one medical strategy for early diagnosis [4].

Medical professionals can use MRI to assess, recognize and diagnose neurological disorders in order to choose the best course of therapy. However, because MRI interpretation is difficult, diagnosis results sometimes differ amongst practitioners, which can cause delays in managing Alzheimer's disease [5]. The recent modern medical methods are increasingly using artificial intelligence (AI)-based

Received: 27 December 2025, Accepted: 20 January 2026, Published: 26 January 2026

systems to automatically evaluate brain imaging data in order to overcome these issues and enable earlier and more accurate diagnosis. AI methods have been used in a number of earlier research to classify Alzheimer's disease using MRI pictures. In order to examine brain slices and categorize them into three groups, Yadav et al. on the research suggested a 2D-based classification method utilizing ResNet50 [6]. In a different work, Buvaneswari et al. showed that deep learning methods and pretrained models greatly enhance classification performance by using a SegNet-based deep learning methodology to discover local brain morphological characteristics [7]. Using Support Vector Machines, Random Forests, and Logistic Regression, Baglat et al. developed a hybrid machine learning framework and found that deep learning techniques produce better classification accuracy [8]. Additionally, with an accuracy of up to 99%, El-Assy et al. developed a modified Convolutional Neural Network (CNN) architecture for Alzheimer's disease classification and early detection. Nevertheless, the evaluation was limited to overall accuracy, without presenting a confusion matrix or class-specific performance metrics [9]. Austin et al, applying ConvNeXT for identifying Alzheimer's disease from brain MRI images, on the findings exposes the accuracy 75% and demonstrates the results performance on the confusion matrix, however on this study, the difficulty is that the model design requires a large amount of computational resources [10].

The majority of current research uses deep learning architectures, which demand significant computer resources, despite these encouraging outcomes. In order to enable implementation on systems with limited resources, lightweight deep learning has evolved as an alternative paradigm that stresses fewer model parameters, lower latency, decreased memory usage and enhanced energy efficiency [11]. Shahriar et al studied the comparative analysis of lightweight deep learning models for memory-constrained devices, it shows that EfficientNetV2-s is one the model lightweight and the results shows that EfficientNetV2 have the highest accuracy than the other lightweight deep learning models [12]. EfficientNet basically is a well-known lightweight design that uses the *Mobile Inverted Bottleneck Convolution* (MBConv) block, which was first presented in MobileNetV2. By reversing the traditional bottleneck structure found in topologies like ResNet, this design greatly lowers computing complexity. Furthermore, the *Squeeze-and-Excitation* (SE) module improves channel-wise feature representation with little computational cost, leading to better accuracy with effective resource use.

EfficientNetV2 is used in this study to classify Alzheimer's disease. EfficientNetV2 has performed well in a variety of non-medical applications, including conventional food image classification [13], mango leaf disease detection [14] and weather classification [15]. However, it has also performed well in medical applications, including brain tumor classification [16], breast cancer classification [17] and diabetic retinopathy detection [18]. The primary improvement in the EfficientNetV2 model lies in the architectural redesign of its building blocks, specifically the substitution of regular MBConv blocks with Fused Inverted Residual Blocks (Fused-MBConv) to increase training efficiency. The proposed method will assist physicians in analyzing MRI images more efficiently, accelerating early diagnosis, and enhancing clinical judgment for the management of Alzheimer's disease.

METHOD

Research Stages

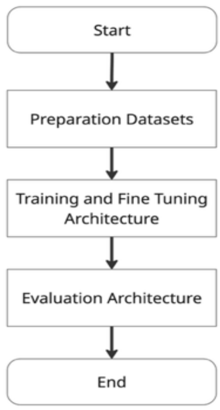


Figure 1. Research Flow Diagram

This study takes a methodological approach, beginning with a thorough assessment of similar past works, followed by an examination of the architectural models offered in earlier research. The procedure then moves on to dataset preparation, model training and fine tuning on the model architecture, and evaluation performance model of the EfficientNetV2 architecture. This evaluation assesses the proposed model's categorization accuracy. The whole research methodology used in this study is shown in Figure 1.

Preparation Dataset

The Alzheimer's Dataset (4-Class Images), licensed by the Alzheimer's Disease Neuroimaging Initiative (ADNI), was utilized in this investigation [19], [20]. The ADNI study is a global research project that gathers longitudinal data from people with Alzheimer's disease, Mild Cognitive Impairment (MCI) and normal cognitive function with the goal of enabling early identification by MRI image processing. The datasets can download on this link https://www.kaggle.com/datasets/marcopinamonti/alzheimer-mri-4-classes-dataset/data. Previous study on Alzheimer's disease used these databases as the basis for their research [21], [22]. There are four types of Alzheimer's disease are included in this dataset are Non-Demented, Very Mild Demented, Mild Demented and Moderate Demented. The MRI brain scans are displayed as 2D image slices. The Clinical Dementia Rating (CDR) scale, which indicates the degree of cognitive impairment in afflicted persons, is the source of the illness designations. The collection has 6,400 photos in total, all of which have a consistent resolution of 208 × 176 pixels. There are 3,200 non-demented photos, 2,240 very mildly demented images, 896 mildly demented images and 64 moderately demented images in the class distribution.

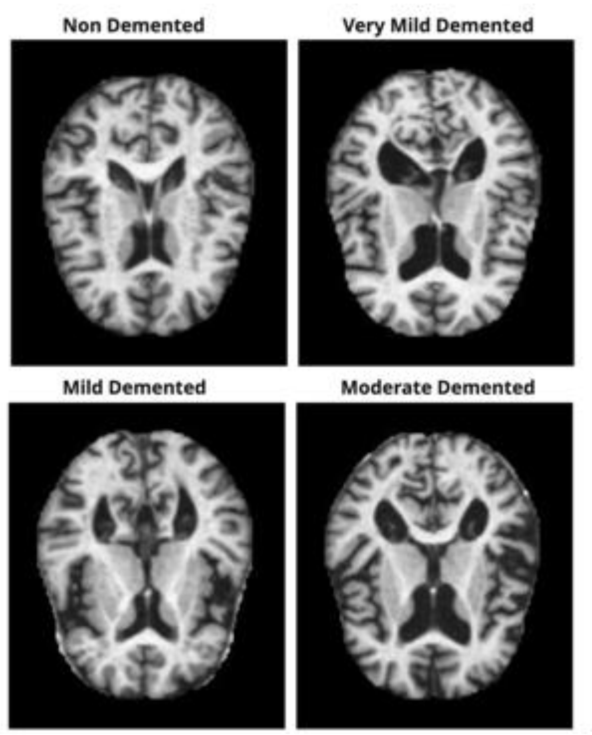


Figure 2. Alzheimer Images on Each Class

The illustration image of each class can be showed on the Figure 2. The dataset for this study is separated into two subsets are training data and testing data. To guarantee consistency, the data is split at an 60:20:20 ratio for training, validation and testing, utilizing the scikit-learn package and a fixed random state of 42. Based on this ratio, the datasets are divided into 3,840 samples are allotted for training, 1,280 for validation and 1,280 for testing. Furthermore, to make the picture classification process easier, one-hot encoding is performed to each class labels.

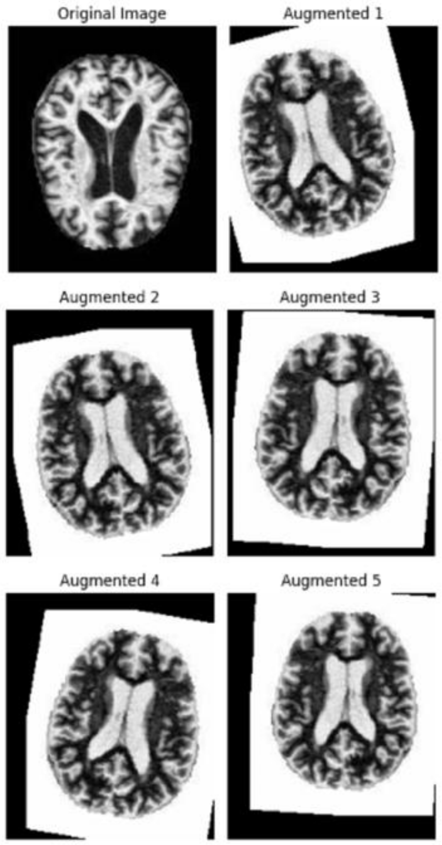


Figure 3. Visualization Data Augmentation

After dividing the dataset into training, validation and test sets, data augmentation is used on the training data to increase model generalization and decrease overfitting. This phase is very crucial in Alzheimer's disease classification, as MRI datasets are frequently restricted and unbalanced. In keeping with past research, the augmentation pipeline includes both horizontal and vertical flipping to promote training variation and resilience [22]. Given that Alzheimer's disease causes worldwide and essentially symmetric brain structural alterations, the classification job does not need rigid anatomical orientation, making flip-based augmentation appropriate for this application. In addition to flipping, additional geometric modifications are used, including random rotation (±15°) to allow for head position fluctuation and random affine scaling

(0.9-1.1) without translation to maintain anatomical alignment. To achieve steady optimization, input photos are normalized with a mean of 0.5 and a standard deviation of 0.5. Overall, the augmentation approach includes random rotation, affine transformation, horizontal and vertical flipping (p = 0.5), and normalization. Figure 3 depicts graphic examples of data augmentation approaches.

Training and Fine Tuning Architecture

The architecture proposed in this study is based on EfficientNetV2, which a new family of CNNs and it produces a higher performance accuracy and have a short period of time for training. The work employed a mix of training-aware neural architecture search and scaling to enhance both training speed and parameter efficiency rather than the previous versions [23]. The foundation of EfficientNetV2 is made up of two key building blocks: Fused-MBConv and MBConv (Figure 4). MBConv (Mobile Inverted Bottleneck Convolution) is used because of its lightweight design and computational effectiveness in image processing jobs. The phrase inverted bottleneck refers to a channel design in which the number of channels is initially increased and subsequently decreased in the final convolutional layer. In contrast, the Fused-MBConv block is used to speed up training and inference, especially for input resolutions ranging from tiny to medium. Fused-MBConv combines the MBConv block's initial 1×1 and 3×3 depthwise convolutions into a single 3×3 operation [24].

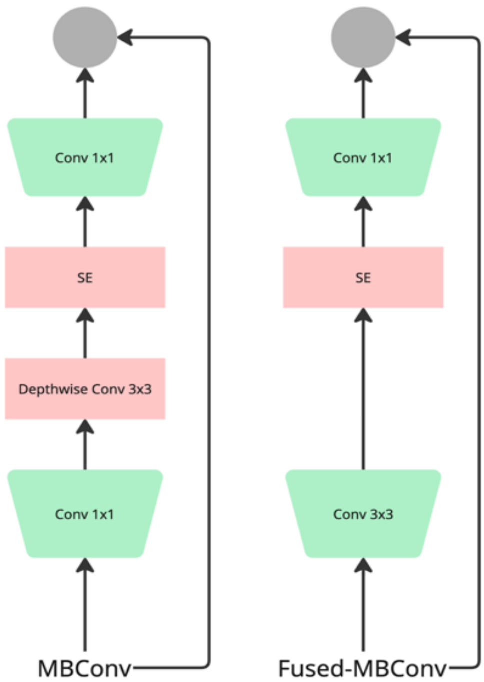
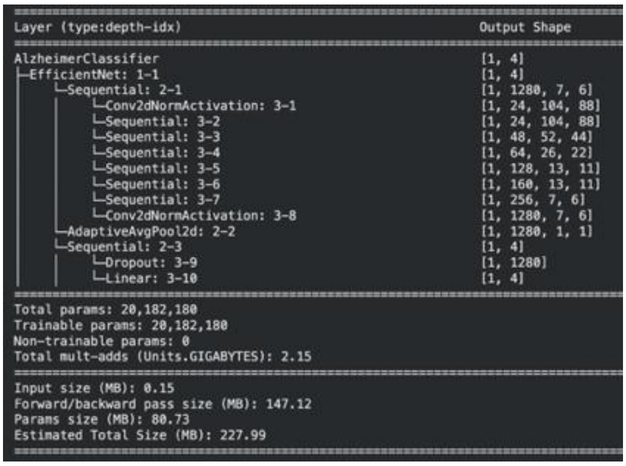


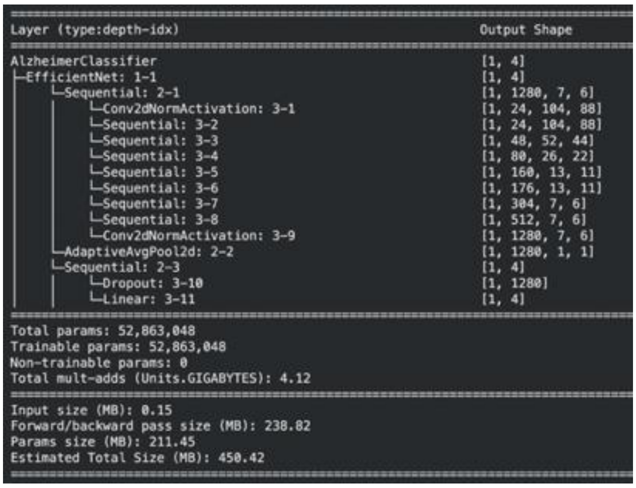
Figure 4. Differentiations MBConv and Fused-MBConv

This study employed with two EfficientNetV2 models, *EfficientNetV2_s* and *EfficientNetV2_m*. Figure 5 and Figure 6 depict the design model for each model, which does partial fine tuning by modifying the final classification layer. Postfix s and m on that words denotes the size and scale of the model architecture. In EfficientNetV2, the suffix 'S' and 'M' denote 'Small' and 'Medium' variants relative to the model family, containing approximately 22 million and 54 million parameters, respectively. Although these models are not lightweight enough for direct deployment on ultra-low-power edge devices, they are significantly more parameter-efficient than conventional deep CNN architectures commonly used in medical image analysis, such as ResNet101 or DenseNet201. Therefore, EfficientNetV2-S and EfficientNetV2-M can be considered resource-efficient models suitable for clinical environments with limitations computational resources, such as hospital workstations or centralized inference servers.



```
Layer (type:depth-idx)                        Output Shape
AlzheimerClassifier                           [1, 4]
├─EfficientNet: 1-1                           [1, 4]
│    └─Sequential: 2-1                        [1, 1280, 7, 6]
│    │    └─Conv2dNormActivation: 3-1         [1, 24, 104, 88]
│    │    └─Sequential: 3-2                   [1, 24, 104, 88]
│    │    └─Sequential: 3-3                   [1, 48, 52, 44]
│    │    └─Sequential: 3-4                   [1, 64, 26, 22]
│    │    └─Sequential: 3-5                   [1, 128, 13, 11]
│    │    └─Sequential: 3-6                   [1, 160, 13, 11]
│    │    └─Sequential: 3-7                   [1, 256, 7, 6]
│    │    └─Conv2dNormActivation: 3-8         [1, 1280, 7, 6]
│    └─AdaptiveAvgPool2d: 2-2                 [1, 1280, 1, 1]
│    └─Sequential: 2-3                        [1, 4]
│    │    └─Dropout: 3-9                      [1, 1280]
│    │    └─Linear: 3-10                      [1, 4]
Total params: 20,182,180
Trainable params: 20,182,180
Non-trainable params: 0
Total mult-adds (Units.GIGABYTES): 2.15
Input size (MB): 0.15
Forward/backward pass size (MB): 147.12
Params size (MB): 80.73
Estimated Total Size (MB): 227.99
```

Figure 5. Model *EfficientNetV2_s* with Partial Fine Tuning



```
Layer (type:depth-idx)                        Output Shape
AlzheimerClassifier                           [1, 4]
├─EfficientNet: 1-1                           [1, 4]
│    └─Sequential: 2-1                        [1, 1280, 7, 6]
│    │    └─Conv2dNormActivation: 3-1         [1, 24, 104, 88]
│    │    └─Sequential: 3-2                   [1, 24, 104, 88]
│    │    └─Sequential: 3-3                   [1, 48, 52, 44]
│    │    └─Sequential: 3-4                   [1, 80, 26, 22]
│    │    └─Sequential: 3-5                   [1, 160, 13, 11]
│    │    └─Sequential: 3-6                   [1, 176, 13, 11]
│    │    └─Sequential: 3-7                   [1, 304, 7, 6]
│    │    └─Sequential: 3-8                   [1, 512, 7, 6]
│    │    └─Conv2dNormActivation: 3-9         [1, 1280, 7, 6]
│    └─AdaptiveAvgPool2d: 2-2                 [1, 1280, 1, 1]
│    └─Sequential: 2-3                        [1, 4]
│    │    └─Dropout: 3-10                     [1, 1280]
│    │    └─Linear: 3-11                      [1, 4]
Total params: 52,863,048
Trainable params: 52,863,048
Non-trainable params: 0
Total mult-adds (Units.GIGABYTES): 4.12
Input size (MB): 0.15
Forward/backward pass size (MB): 238.82
Params size (MB): 211.45
Estimated Total Size (MB): 450.42
```

Figure 6. Model *EfficientNetV2_m* with Partial Fine Tuning

The model that has already been initiated will be added with the image input. The image input size for the models is 208x176 and more fine adjustment is available at the end of the layer, which includes a thick layer with four classes. The model additionally employs *Adam Optimizer* with hyperparameter learning to decrease error throughout the training process. Hyperparameter loss function on this study also implemented with cross entropy loss to improving model convergence and classification performance. Detail informations about number hyperparameters used can be see on the Table 1.

Table 1. Hyperparameter Configurations

Hyperparameter	Value
Learning Rate	0.001
Epoch	128
Batch Size	8

Evaluation Architecture

This study assessment procedure is based on the values of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) as determined by the confusion matrix. The confusion matrix is a performance evaluation tool that shows the number of properly and mistakenly categorized samples for each class by comparing actual labels to model predictions. TP is the number of real positive samples accurately predicted as positive by the model, whereas TN represents the number of actual negative samples correctly forecasted as negative. FP is the number of actual negative samples that were improperly forecasted as positive, whereas FN represents the number of actual positive samples that were incorrectly projected as negative. The four values will be used for calculate *precision*, *recall*, and *f1-score* of the models which the formula as you can see on the formula (1-3).

$$precision = \frac{TP}{TP+FP} \quad (1)$$

$$recall = \frac{TP}{TP+FN} \quad (2)$$

$$F1 - Score = 2 * \frac{precision * recall}{precision + recall} \quad (3)$$

RESULTS AND DISCUSSIONS

The model obtained during the training phase with the assistance of *runpod* to tailor GPU use gained some information and *pytorch* for the library. The training and validation results of the model are presented in the form of training and validation loss, recall, precision and F1-score curves.

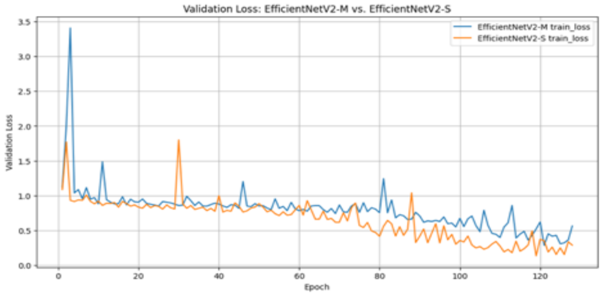


Figure 7. Validation Loss on Each Epoch

The Figure 7 gives information about validation loss during training with data validation each epoch by comparing between model EfficientNetV2_s with model EfficientNetV2_m. The training and validation loss curves are used to analyze the learning behavior and generalization capability of the proposed model throughout the training process. Overall, both models exhibit a decreasing trend in validation loss, showing that generalization performance improves as training continues. However, EfficientNetV2-s consistently achieves lower validation loss and has more steady convergence than EfficientNetV2-m. In the early epochs, both models exhibit noteworthy swings, with EfficientNetV2-m exhibiting a high validation loss spike, indicating initial instability and sensitivity to model complexity. Although the value spikes are still present during training, EfficientNetV2-s shows reduced variability and reaches stability more gradually.

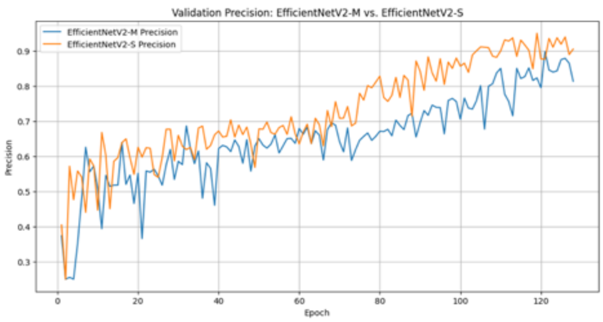


Figure 8. Validation Precision on Each Epoch

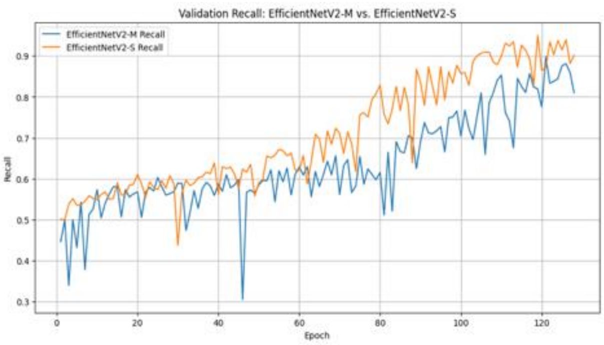


Figure 9. Validation Recall on Each Epoch

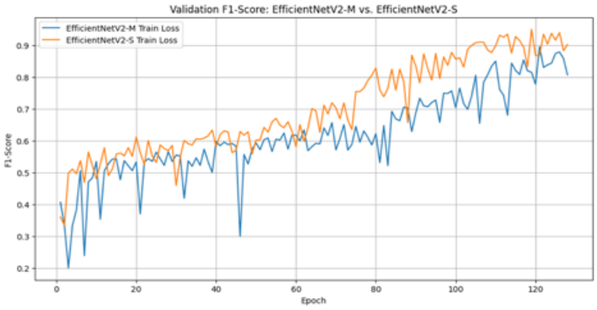


Figure 10. Validation F1-Score on Each Epoch

In the other hand, Figure 8, Figure 9 and Figure 10 present a validation performance comparison of EfficientNetV2-m and EfficientNetV2-s in terms of precision, recall and F1-score across training epochs. The precision, recall and F1-score curves are used to completely assess the model's classification performance throughout training and validation. Precision assesses the model's ability to properly identify positive predictions while reducing false positives, recall indicates the model's sensitivity in finding real positive situations and the F1-score gives a balanced assessment by combining precision and recall. Overall, both models show continuous improvements in all assessment measures, demonstrating that classification performance is improving with time. However, EfficientNetV2-s regularly beats EfficientNetV2-m on all measures and has a more steady convergence tendency. In the early phases of training, both models exhibit considerable variations, notably in precision and recall, indicating initial instability in feature learning. As training advances, EfficientNetV2-s converges more gradually and attains greater precision values, suggesting more confidence in good predictions and fewer false positives. Simultaneously, EfficientNetV2-s achieves greater recall with fewer sudden dips, indicating superior sensitivity in detecting affirmative cases.

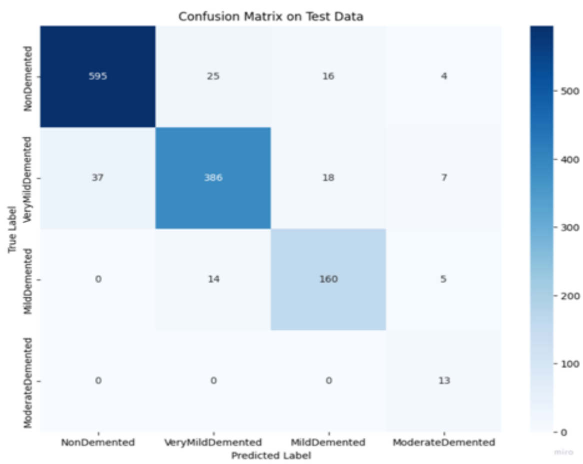


Figure 11. Confusion Matrix EfficientNetV2_s

The F1-score reflects the combined benefit of enhanced precision and recall, as EfficientNetV2-s regularly produces higher values with lower variation than EfficientNetV2-m. In the last epochs, EfficientNetV2-s achieves F1-scores better than 0.90, but EfficientNetV2-m remains lower and demonstrates more performance fluctuations. These findings suggest that the smaller model strikes a better compromise between sensitivity and specificity, resulting in improved generalization on previously encountered data. Overall, the data show that EfficientNetV2-s provides higher and more robust validation performance despite its reduced model complexity for classify the data validation during training phase.

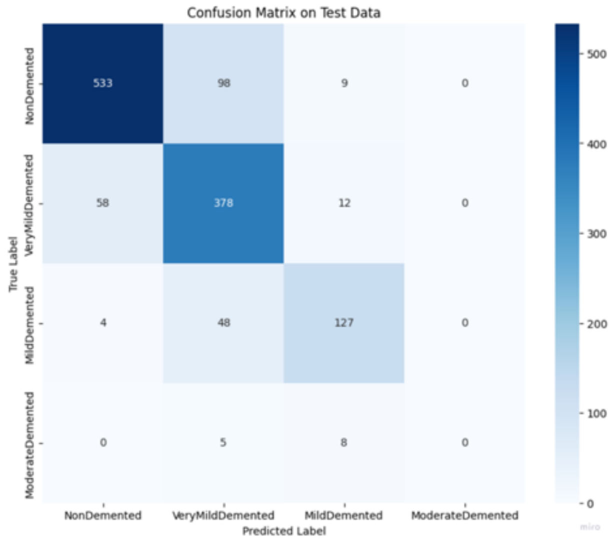


Figure 12. Confusion Matrix EfficientNetV2_m

While Figure 11 and Figure 12 shows the confusion matrices produced by evaluating the proposed model on the test dataset. The matrices show the distribution of true versus predicted labels and give information on class-specific performance and misclassification tendencies. The initial confusion matrix shows robust class discriminating throughout all dementia stages. The Non Demented class has a high true positive rate (595 samples) and low misclassification to adjacent classes. The Very Mild Demented class is likewise mainly properly diagnosed (386 samples) with little misunderstanding with Non Demented, indicating a slight clinical shift between the two stages. Similarly, the Mild Demented class performs well (160 accurate predictions) with little misclassification largely to Very Mild Demented. Importantly, the Moderate Demented class is properly categorized, with all samples accurately recognized.

In contrast, the second confusion matrix shows decreased class separability. Although Non Demented predictions remain quite high (533 samples), misclassification as Very Mild Demented grows significantly. Confusion between Very Mild Demented and nearby classes worsens, and the Mild Demented class becomes less accurate, with frequent reclassification into Very Mild Demented. Notably, the model does not reliably detect any Moderate Demented patients in this setting. Overall, the results show that the first model configuration provides more balanced and clinically reliable performance across all dementia stages, particularly Moderate Demented, whereas the second model causes significant confusion between adjacent severity levels, emphasizing the importance of robust feature representation for accurate dementia staging.

Besides that, Table 2 highlights the test performance comparison of model EfficientNetV2-s and model EfficientNetV2-m using common evaluation criteria. As demonstrated in the table, EfficientNetV2-s consistently beats EfficientNetV2-s across all measures, with the performance matrix precision, recall, and F1-score values of about 0.90. In comparison, EfficientNetV2-m achieves poorer performance, with precision, and recall values of 0.81 and an F1-score of 0.80. The higher F1-score produced by EfficientNetV2-s suggests a better balanced trade-off between precision and recall, implying greater resilience and generalization capabilities on previously encountered test data.

Table 2. Result Precision, Recall and F1-Score

Metrics	EfficientNetV2_s	EfficientNetV2_m
Precision	**0.90**	0.81
Recall	**0.90**	0.81
F1-Score	**0.90**	0.80
Accuracy	**0.90**	0.81

Furthermore, to provide a comprehensive comparison between EfficientNetV2 and other architectures, this study incorporates accuracy as an additional evaluation metric and compares the results with those obtained using the ConvNeXT model. The results of the comparison model accuracy can be see on the Table 3.

Table 3. Comparison of Model Accuracy

Model	Accuracy
EfficientNetV2-s	**0.90**
EfficientNetV2-m	**0.81**
ConvNeXT [10]	**0.75**

In addition, Table 4 presents the class-wise precision (P), recall (R) and F1-score (F1) for model EfficientNetV2_s and model EfficientNetV2_m across four Alzheimer's disease categories are Mild Demented, Moderate Demented, Non Demented, and Very Mild Demented. Based on that table, EfficientNetV2_s demonstrates superior and more consistent performance across all classes compared to EfficientNetV2_m. For the Mild Demented class, EfficientNetV2_s has a high accuracy (0.94), recall (0.92) and F1-score (0.93), showing great discriminative ability. In contrast, the architecture EfficientNetV2_m performs worse, notably in recall (0.83), resulting in an F1-score of 0.86. Similar tendencies can be seen in the Moderate Demented class, where EfficientNetV2_s achieves an F1-score of 0.88, surpassing EfficientNetV2_m, which has an F1-score of 0.77.

Table 4. Precision, Recall and F1-Score on each Class

Class	EfficientNetV2_s			EfficientNetV2_m		
	P	**R**	**F1**	**P**	**R**	**F1**
Mild Demented	0.94	0.92	0.93	0.89	0.83	0.86
Moderate Demented	0.90	0.86	0.88	0.71	0.84	0.77
Non Demented	0.82	0.89	0.85	0.81	0.70	0.75
Very Mild Demented	0.44	1	0.61	0	0	0

Notably, in the Very Mild Demented class, EfficientNetV2_s has perfect recall (1.00) but low precision (0.44), resulting in an F1-score of 0.61. This pattern indicates that, while the model correctly detects all Very Mild Demented patients, it also generates a significant number of false positives. In contrast, EfficientNetV2_m fails to accurately identify this class, resulting in 0 precision, recall, and F1-score, most likely owing to class imbalance and inadequate feature representation. Overall the class-wise analysis shows that EfficientNetV2_s has more robust and reliable classification performance, particularly for early-stage Alzheimer's disease detection, whereas EfficientNetV2_m struggles with minority classes, demonstrating the effectiveness of lightweight architectures in imbalanced medical imaging datasets.

Table 5. Comparison Parameter Size and FLOPs Models

Model	Parameters	FLOPs
EfficientNetV2-s	**22 M**	**8.8 G**
EfficientNetV2-m	**54 M**	**24 G**
ConvNeXT Tiny	**28 M**	**4.5 G**
ConvNeXT Small	**50 M**	**8.7 G**
ConvNeXT Base	**89 M**	**15.4 G**

After analyzing the performance of EfficientNetV2, this study further compares the parameter sizes of the models and FLOPs (Floating Point Operations per Second) to identify the architecture with the lowest number of parameters, thereby ensuring computational efficiency. Table 5 shows a comparison of model complexity between the EfficientNetV2 and ConvNeXT variations in terms of parameter count and floating-point operations (FLOPs). Among the models tested, EfficientNetV2-s has the smallest parameter size (22 million) and requires 8.8 GFLOPs, indicating its lightweight architecture. In contrast, the architecture of EfficientNetV2-m, with 54 million parameters and 24 GFLOPs, is the most computationally costly model in the comparison. The ConvNeXT variations show differing trade-offs between parameter size and computational cost. ConvNeXt-Tiny has 28 million parameters and a processing need of 4.5 GFLOPs, but ConvNeXt-Small and ConvNeXt-Base gradually rise in complexity, reaching 50 million parameters with 8.7 GFLOPs and 89 million parameters with 15.4 GFLOPs, respectively. Overall, this comparison shows that EfficientNetV2-s provides the best mix of model compactness and computational efficiency, with a low parameter count and mild FLOPs. Although ConvNeXt-Tiny uses fewer FLOPs, EfficientNetV2-s achieves better classification performance with fewer parameters, making it more suited for deployment in resource-constrained and clinical situations where accuracy and efficiency are essential.

CONCLUSIONS

The objective of this study is to evaluate the effectiveness of EfficientNetV2 in classifying Alzheimer's disease. EfficientNetV2 is an improved version of its predecessor, EfficientNet, designed to achieve high classification performance while enabling faster data processing during the training phase. Although EfficientNetV2 is a lightweight model commonly deployed on resource-constrained devices, the experimental results demonstrate its strong capability in Alzheimer's disease classification, achieving competitive performance compared to other models that require significantly larger computational resources.

This study has various obstacles due to a lack of data for specific labels, resulting in an unbalanced dataset. This imbalance has a substantial impact on the performance of the model being evaluated. The findings reveal a better performance matrix for the model EfficientNetV2-s compared to the model EfficientNetV2-m, which has a large layer and performs poorly. The proposed model has to be improved and refined further. As a result, future research may include additional image enhancement techniques, such as Contrast Limited Adaptive Histogram Equalization (CLAHE), to improve class separability, as well as data balancing methods like the Synthetic Minority Over-sampling Technique (SMOTE), to address dataset imbalance.

REFERENCES

[1] A. G. M. Sianturi, "Stadium, Diagnosis, dan Tatalaksana Penyakit Alzheimer: Stages, Diagnosis, and Management of Alzheimer's Disease," *Majalah Kesehatan Indonesia*, vol. 2, no. 2, pp. 39–44, Oct. 2021, doi: 10.47679/makein.202132.

[2] J. L. Cummings and G. Cole, "Alzheimer Disease," *JAMA*, vol. 287, no. 18, pp. 2335–2338, May 2002, doi: 10.1001/jama.287.18.2335.

[3] A. Arfina, "Pengaruh Edukasi terhadap Pengetahuan Masyarakat Tentang Deteksi Dini Alzheimer di Kelurahan Labuh Baru Pekanbaru," *Health Care: Jurnal Kesehatan*, vol. 10, no. 2, pp. 256–261, 2021. https://doi.org/10.36763/healthcare.v10i2.170.

[4] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nat Rev Neurol*, vol. 6, no. 2, pp. 67–77, 2010, doi: 10.1038/nrneurol.2009.215.

[5] C. Ferrari and S. Sorbi, "The complexity of Alzheimer's disease: an evolving puzzle," *Physiol Rev*, vol. 101, no. 3, pp. 1047–1081, Jan. 2021, doi: 10.1152/physrev.00015.2020.

[6] K. Yadav and K. Miyapuram, "A Novel Approach Towards Early Detection of Alzheimer's Disease Using Deep Learning on Magnetic Resonance Images," 2021, pp. 486–495. doi: 10.1007/978-3-030-86993-9_43.

[7] P. R. Buvaneswari and R. Gayathri, "Deep Learning-Based Segmentation in Classification of Alzheimer's Disease," *Arab J Sci Eng*, vol. 46, no. 6, pp. 5373–5383, 2021, doi: 10.1007/s13369-020-05193-z.

[8] P. Baglat, A. W. Salehi, A. Gupta, and G. Gupta, "Multiple Machine Learning Models for Detection of Alzheimer's Disease Using OASIS Dataset," in *Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation*, S. K.

Sharma, Y. K. Dwivedi, B. Metri, and N. P. Rana, Eds., Cham: Springer International Publishing, 2020, pp. 614–622. https://link.springer.com/book/10.1007/978-3-030-64861-9.

[9] A. M. El-Assy, H. M. Amer, H. M. Ibrahim, and M. A. Mohamed, "A novel CNN architecture for accurate early detection and classification of Alzheimer's disease using MRI data," *Sci Rep*, vol. 14, no. 1, p. 3463, 2024, doi: 10.1038/s41598-024-53733-6.

[10] "Klasifikasi Penyakit Alzheimer Dari Scan Mri Otak Menggunakan Convnext," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 6, pp. 1223–1232, Dec. 2024, doi: 10.25126/jtiik.2024118117.

[11] H.-I. Liu *et al.*, "Lightweight deep learning for resource-constrained environments: A survey," *ACM Comput Surv*, vol. 56, no. 10, pp. 1–42, 2024. https://doi.org/10.48550/arXiv.2404.07236.

[12] T. Shahriar, "Comparative Analysis of Lightweight Deep Learning Models for Memory-Constrained Devices," *arXiv preprint arXiv:2505.03303*, 2025. https://doi.org/10.48550/arXiv.2505.03303.

[13] E. E. Citra, D. H. Fudholi, and C. K. Dewa, "Implementasi Arsitektur EfficientNetV2 Untuk Klasifikasi Gambar Makanan Tradisional Indonesia," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 2, pp. 766–777, 2023. DOI: https://doi.org/10.30865/mib.v7i2.5881.

[14] A. H. Nurfauzi, Y. Azhar, and D. R. Chandranegara, "Penerapan Model EfficientNetV2-B0 pada Benchmark IP102 Dataset untuk Menyelesaikan Masalah Klasifikasi Hama Serangga," *Jurnal Repositor*, vol. 5, no. 3, 2023. DOI: https://doi.org/10.22219/repositor.v5i3.32075.

[15] M. A. Mutasodirin and F. M. Falakh, "Efficient Weather Classification Using DenseNet and EfficientNet," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 9, no. 2, pp. 173–179, 2024. DOI: https://doi.org/10.30591/jpit.v9i2.7539.

[16] K. Gencer, "A Comparative Analysis of EfficientNetB0 and EfficientNetV2 Variants for Brain Tumor Classification Using MRI Images," *International Journal of Innovative Engineering Applications*, vol. 9, no. 1, pp. 1–7, 2025, doi: 10.46460/ijiea.1523782.

[17] M. Al Moteri, T. R. Mahesh, A. Thakur, V. Vinoth Kumar, S. B. Khan, and M. Alojail, "Enhancing accessibility for improved diagnosis with modified EfficientNetV2-S and cyclic learning rate strategy in women with disabilities and breast cancer," *Front Med (Lausanne)*, vol. Volume 11-2024, 2024, [Online]. Available: https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2024.1373244

[18] M. Tang, H. Y. Ting, A. Atanda, and K. Ting, "Transfer Learning with EfficientNetV2 forDiabetic Retinopathy Detection," *International Journal of Engineering and Manufacturing*, vol. 14, pp. 54–60, Dec. 2024, doi: 10.5815/ijem.2024.06.05.

[19] S. G. Mueller *et al.*, "Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI)," *Alzheimer's & Dementia*, vol. 1, no. 1, pp. 55–66, 2005. DOI: 10.1016/j.jalz.2005.06.003.

[20] C. R. Jack Jr. *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, Apr. 2008, doi: https://doi.org/10.1002/jmri.21049.

[21] M. Asaduzzaman, Md. K. Alom, and Md. E. Karim, "ALZENET: Deep learning-based early prediction of Alzheimer's disease through magnetic resonance imaging analysis," *Telematics and Informatics Reports*, vol. 17, p. 100189, 2025, doi: https://doi.org/10.1016/j.teler.2025.100189.

[22] A. A. A. El-Latif, S. A. Chelloug, M. Alabdulhafith, and M. Hammad, "Accurate Detection of Alzheimer's Disease Using Lightweight Deep Learning Model on MRI Data," *Diagnostics*, vol. 13, no. 7, 2023, doi: 10.3390/diagnostics13071216.

[23] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, PMLR, 2019, pp. 6105–6114. https://doi.org/10.48550/arXiv.1905.11946.

[24] M. Tan and Q. Le, "Efficientnetv2: Smaller models and faster training," in *International conference on machine learning*, PMLR, 2021, pp. 10096–10106. https://doi.org/10.48550/arXiv.2104.00298.