Analysis of ’Early diagnosis of Alzheimer’s disease from MRI images with deep learning model’

Ben Decker
CS 6680
Utah State University
Logan, Utah
a02269549@usu.edu
Github Code and Data.

***Abstract*—Early detection of Alzheimer’s disease is crucial for effective intervention and management of the condition. This paper reviews the study ’Early diagnosis of Alzheimer’s disease from MRI images with deep learning model’, which explores the use of machine learning models for diagnosing Alzheimer’s disease from MRI images. While the original study demonstrates the potential of deep learning in medical diagnostics, it makes some questionable assumptions and employs inefficient practices. In this paper, we identify these issues and propose corrections to improve the accuracy and efficiency of the diagnostic model.**

***Index Terms*—Alzheimer’s, Dementia, Deep Learning, MRI, Medical Imaging, Machine Learning**

I. Introduction

Alzheimer’s disease is a progressive neurodegenerative disorder that affects memory, thinking, and behavior. It is the most common cause of dementia, a general term for a decline in cognitive function severe enough to interfere with daily life. Alzheimer’s disease accounts for 60-80% of dementia cases. Symptoms of Alzheimer’s disease typically develop slowly and worsen over time, leading to severe memory loss and cognitive decline. The exact cause of Alzheimer’s disease is not fully understood, but it is believed to involve a combination of genetic, environmental, and lifestyle factors. There is currently no cure for Alzheimer’s disease, but treatments are available to help manage symptoms and improve quality of life for patients and their families. Early diagnosis of Alzheimer’s disease is crucial for effective intervention and management of the condition. Medical imaging techniques, such as magnetic resonance imaging (MRI), can be used to detect structural changes in the brain associated with Alzheimer’s disease. Machine learning models have shown promise in diagnosing Alzheimer’s disease from MRI images, offering a non-invasive and cost-effective approach to early detection.

In September 2024, ’Early diagnosis of Alzheimer’s disease from MRI images with deep learning model’ [1] was published. It claims impressive results, but it seems to lack novel innovation. It builds on the Dementia Network (DEMNET) convolutional neural network, proposed in ’DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images’ [2]. The DEMNET model is designed to diagnose Alzheimer’s disease and dementia from MRI images, and it uses a deep learning approach to achieve high accuracy. The model is trained on a dataset of MRI images from patients with Alzheimer’s disease and healthy controls, and it is evaluated on a separate test set to assess its performance. The results show that the DEMNET model can accurately diagnose Alzheimer’s disease and dementia from MRI images, with high sensitivity and specificity. It concludes that deep learning models can be effective tools for early diagnosis of Alzheimer’s disease and dementia, and it highlights the potential of machine learning in medical imaging. In [1], the authors attempt to utilize Google’s Inceptionv3 model to extract features from the input images, and then pass the features to the existing DEMNET model for classification. However, there are several assumptions and practices in the study that raise concerns about its validity and generalizability. In this paper, we critically review the study and propose improvements to address these issues. Code and data for improvements are found at github.

II. Related Work

It has been shown that convolutional neural networks (CNNs) can be effective in diagnosing Alzheimer’s disease from MRI images. ’Alzheimer’s Disease Detection Through Whole-Brain 3D-CNN MRI’ [3] uses a 3D convolutional neural network to classify MRI cubes into Alzheimer’s disease and control groups. The model achieves high accuracy and demonstrates the potential of deep learning in medical imaging. Similarly, feature extraction methods have been used to extend image classification models. ’Feature Extraction for Medical Image Classification: A Novel Statistical Approach’ [4] uses feature extraction to improve the performance of a CNN model to classify diabetic retinopathy. The study shows that feature extraction can enhance the model’s ability to learn discriminative features from the input images. Google’s Inceptionv3 network has been widely used in image classification tasks. ’Rethinking the Inception Architecture for Computer Vision’ [5] introduces the Inceptionv3 architecture, which uses a combination of convolutional layers with different filter sizes to capture spatial hierarchies in the input images. The model achieves state-of-the-art performance on the ImageNet dataset and demonstrates the effectiveness of deep learning in image recognition tasks. It can be modified to extract features from the input images and pass them to another model for classification, as demonstrated in 'OCTNet: A Modified Multi-Scale Attention Feature Fusion Network with InceptionV3 for Retinal OCT Image Classification' [6]. [2] proposes the DEMNET model, which utilizes several convolutional layers to classify MR images of dementia patients into one of four groups: Moderate, Mild, Very Mild, and Non Demented. In this paper we will incrementally implement the DEMNET model and compare the results with the original study [2], then we will discuss the results and propose improvements to the model.

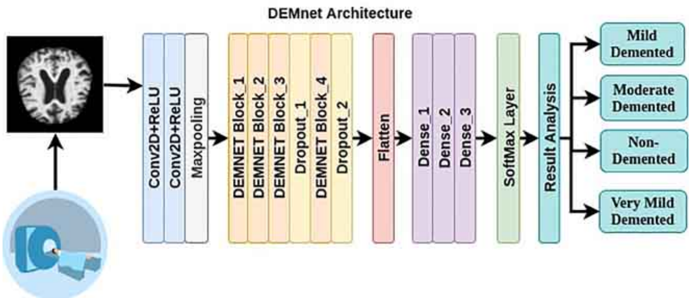

Fig. 1. High-level architecture of the DEMNET model. [2]

III. Methodology

The proposed DEMNET model consists of the following types of layers:

- Convolutional layer: These layers apply a set of filters to the input image to extract features. The filters are learned during the training process. The output of a convolutional layer is a feature map that represents the presence of specific features in the input image. They are useful for detecting patterns in the 2D input images.
- Pooling layer: These layers downsample the feature maps to reduce the spatial dimensions.
- Fully connected layers: These layers connect every neuron in one layer to every neuron in the next layer.
- Dropout layer: This layer randomly sets a fraction of the input units to zero during training to prevent overfitting.
- Softmax layer: This layer computes the probability distribution over the 4 output classes.
- DEMNET block: This is a custom layer that consists of two convolutional layers followed by a batch normalization and maxpooling layer.

The sizes of the 2D convolutional filters are never revealed, but the authors reveal the number of trainable parameters in each layer, from which the size of the filters can be inferred (See Figure 2).

The dataset used for training and testing the DEMNET model is found on Kaggle.com, and consists of 6400 MR images of dementia patients. The dataset is labled with one of 4 lables: Moderate, Mild, Very Mild, and Non Demented. The number of images in each class is shown in Figure 3. There is an obvious class imbalance issue, which can lead to poor performance of the model. To address this, the authors use a method called Synthetic Minority Oversampling Technique (SMOTE) [7] to generate synthetic samples for the minority classes. The only detail the authors provide about the use of SMOTE is the random seed that was used, which was 42. In our application of SMOTE, we use class `SMOTE` in `imblearn.over_sampling` because it is simple and customizable. It should be noted that SMOTE is not often used for oversampling of image data because it can introduce artifacts in the synthetic samples. Images are spatially correlated, and generating synthetic samples can lead to unrealistic patterns in the data. Because the brain MR images in the dataset are fixed in perspective, the negative effects of SMOTE may be less pronounced. The authors do not justify the use of SMOTE in this context. This may be a potential point of failure for the DEMNET model. Applying SMOTE gives us a dataset with 12800 images, as shown in Figure 4.

Layer Type	Output Shape	Parameters
Conv2D+ReLU	(None, 176, 176, 16)	448
Conv2D+ReLU	(None, 176, 176, 16)	2320
Maxpooling	(None, 88, 88, 16)	0
DEMNET Block_1	(None, 44, 44, 32)	14016
DEMNET Block_2	(None, 22, 22, 64)	55680
DEMNET Block_3	(None, 11, 11, 128)	221952
Dropout_1	(None, 11, 11, 128)	0
DEMNET Block_4	(None, 5, 5, 256)	886272
Dropout_2	(None, 5, 5, 256)	0
Flatten	(None, 6400)	0
Dense_1	(None, 512)	3279360
Dense_2	(None, 128)	66176
Dense_3	(None, 64)	8512
SoftMax Layer	(None, 4)	260
Total parameters: 4,534,996		
Trainable parameters: 4,532,628		
Non-trainable parameters: 2,368		

Fig. 2. Trainable parameters of DEMNET model. [2]

Class	No of Images
Mild Demented (MID)	896
Moderate Demented (MOD)	3200
Non-Demented (ND)	2240
Very Mild Demented (VMD)	64

Fig. 3. Number of images per class in Kaggle dataset. [2]

Class	No of Images
Mild Demented (MID)	3200
Moderate Demented (MOD)	3200
Non-Demented (ND)	3200
Very Mild Demented (VMD)	3200

Fig. 4. Number of images per class in dataset after SMOTE. [2]

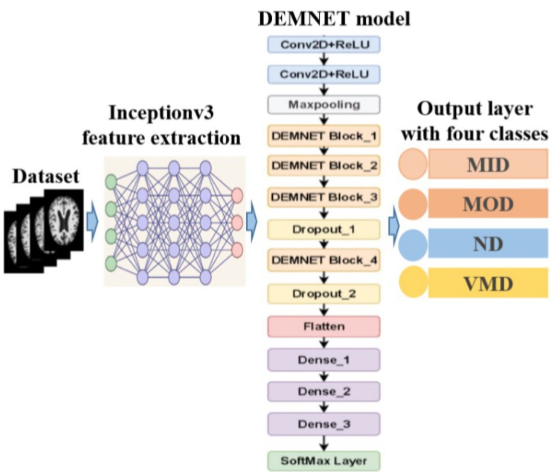

Fig. 5. Inceptionv3/DEMNET combined model. [1]

The model is built using the Keras API with TensorFlow backend. All of the convolutional layers use padding to maintain the spatial dimensions of the feature maps. The activation function used in the convolutional layers is ReLU, which is a common choice for deep learning models. The final layer uses the softmax activation function to compute the probability distribution over the 4 output classes. The model is compiled with the Adam optimizer and categorical crossentropy loss function. The model is trained for 20 epochs with a batch size of 64. The training and validation accuracy and loss are monitored during training to evaluate the model's performance. The model is evaluated on a separate test set to assess its accuracy and generalization performance. The authors of 'Early diagnosis of Alzheimer's disease from MRI images with deep learning model' [1] propose to use Google's Inceptionv3 model to extract features from the input images, and then pass the features to the DEMNET model for classification. The Inceptionv3 model is a deep convolutional neural network that has been pre-trained on the ImageNet dataset. It is capable of extracting high-level features from the input images, which can be useful for image classification tasks. The authors do not provide details about how the Inceptionv3 model is integrated with the DEMNET model, but they mention that the features extracted by Inceptionv3 are passed to the DEMNET model for classification. There are several potential issues with this approach that we will enumerate here. A diagram of the combined Inceptionv3/DEMNET model is shown in Figure 5.

- The Inceptionv3 model is pre-trained on the ImageNet dataset, which consists of natural images. Because the Inceptionv3 model was imported with pre-trained weights (and those weights were not updated),the features extracted by Inceptionv3 may not be relevant for classifying brain MR images. Some form of transfer learning or further fine-tuning of the Inceptionv3 model on the brain MR images dataset may be more effective.
- The DEMNET model, as described above, is intended on being trained on 2D images. [2] The 2D convolutional layers in the model are designed to extract features from the input images. The Inceptionv3 network will product a feature vector, but that vector will no longer be spacially correlated. The DEMNET model would not be able to extract features from the feature vector, and the model would not be able to learn the spatial hierarchies in the input images. This may lead to poor performance of the model. The 2D convolutional layers will not be able to function on the feature vector. The authors do not explain what, if any, modifications were made to the DEMNET model to operate with Inceptionv3. In our attempt to reproduce the authors' results, we will remove the 2D convolutional layers from the DEMNET model.
- The authors do not justify the use of the Inceptionv3 model in the context of diagnosing Alzheimer's disease from MRI images. The choice of model architecture should be based on the characteristics of the input data and the requirements of the task.

Despite these issues, the authors claim that the combined model achieves high accuracy (98.67%) [1] in diagnosing the correct stage of Alzheimer's disease from MRI images. The authors do not provide details about the training process or hyperparameters used in the study. This lack of transparency makes it difficult to reproduce the results and evaluate the validity of the study. The authors also do not compare the performance of the combined model with the DEMNET model alone, which makes it difficult to assess the contribution of the Inceptionv3 model to the overall performance. We will attempt to reproduce the results of the study and evaluate the performance of the combined model. We will also compare the performance of the combined model with the DEMNET model alone to assess the contribution of the Inceptionv3 model to the overall performance. In the combined model, the data was preprocessed in the same way as the DEMNET model. The images were resized to 256x256 pixels and normalized to the range [0, 1], with the class imbalance addressed with SMOTE. Neither study provided details about the application of SMOTE, so we used the same random seed (42) as the authors. The Inceptionv3 model was imported from the Keras applications module and the weights were not updated during training. We removed the final softmax layer from the Inceptionv3 model and used the output of the penultimate layer as the feature vector. We then passed the feature vector to the DEMNET model for classification. Because it made no sense to include the 2D convolutional layers in DEMNET, they were removed.The parameter count of the combined Inceptionv3/DEMNET model is shown in Figure 6. Both networks

Layer (type)	Output Shape	Param #
input_layer_43 (InputLayer)	(None, 176, 176, 3)	0
functional_2 (Functional)	(None, 4, 4, 2048)	21,802,784
global_average_pooling2d_16 (GlobalAveragePooling2D)	(None, 2048)	0
sequential_21 (Sequential)	(None, 4)	1,123,268

Fig. 6. Parameter count of the combined Inceptionv3/DEMNET model. The 'functional' layer is Inceptionv3, and the 'sequential' layer is DEMNET.

(DEMNET and combined) were trained for 20 epochs with a batch size of 64.

To address the issues with the combined model, we propose the following improvements:

- Use SMOTE data augmentation on the feature vectors, not the images.
- Modify the DEMNET model to operate on the feature vector, removing the 2D convolutional layers.
- Use the smaller VGG16 [8] model instead of Inceptionv3 to extract features from the input images. VGG16 is a simpler model that may be more suitable for the task of diagnosing Alzheimer's disease from MRI images.

IV. Results - DEMNET

The DEMNET model was trained on the preprocessed dataset and evaluated on a separate test set. The model training took 66 minutes total. The training and validation accuracy are shown in Figure 7 After training, the model was evaluated

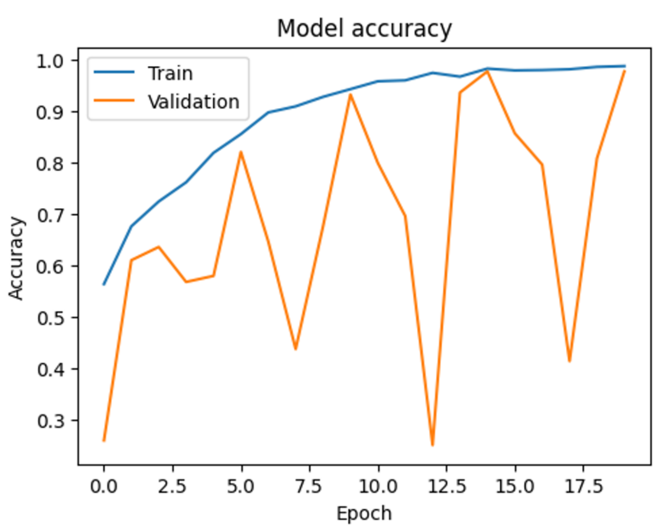

Fig. 7. Training and validation accuracy of DEMNET model.

against the test set and the DEMNET model achieved an accuracy of 0.9314 and a loss of 0.2440. This is comparable to the results achieved by the original authors of DEMNET, with an advertised accuracy of 93%. [2]. Achieving the same

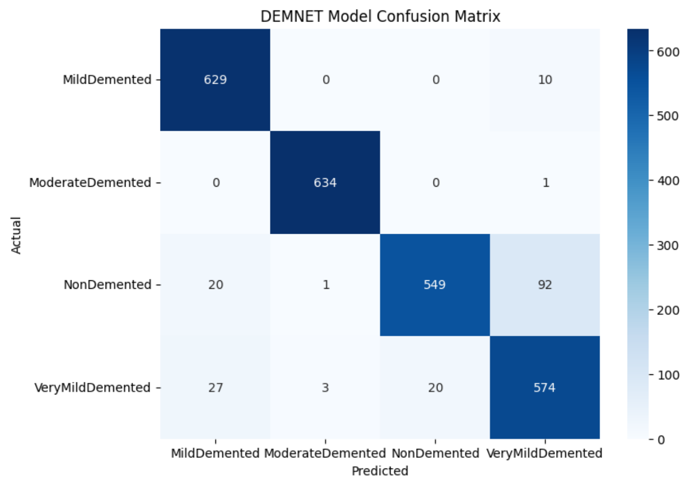

Fig. 8. Confusion matrix of DEMNET model validation.

results as the original authors is a good sign that the model was implemented correctly.

V. Results - Inceptionv3/DEMNET

The combined Inceptionv3/DEMNET model was trained on the preprocessed dataset and evaluated on a separate test set. The model training took 80 minutes total. The training and validation accuracy and loss are shown in Figure 9 and Figure **??**. After training, the model was evaluated against the test

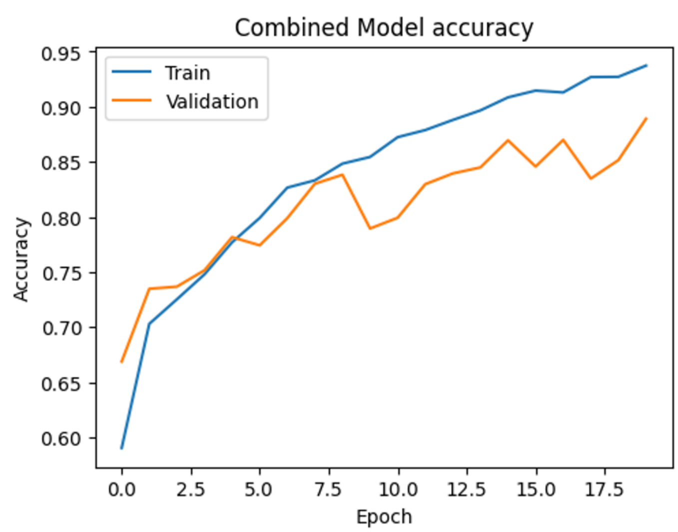

Fig. 9. Training and validation accuracy of Inceptionv3/DEMNET model.

set and the combined Inceptionv3/DEMNET model achieved an accuracy of 0.8716, which is much less than the advertised value by the original authors of 98.67% [1]. The loss was 0.3145. The confusion matrix of the combined model is shown in Figure 10.

VI. Results - Improved Model

The improved model was trained on the preprocessed dataset and evaluated on a separate test set. The model training

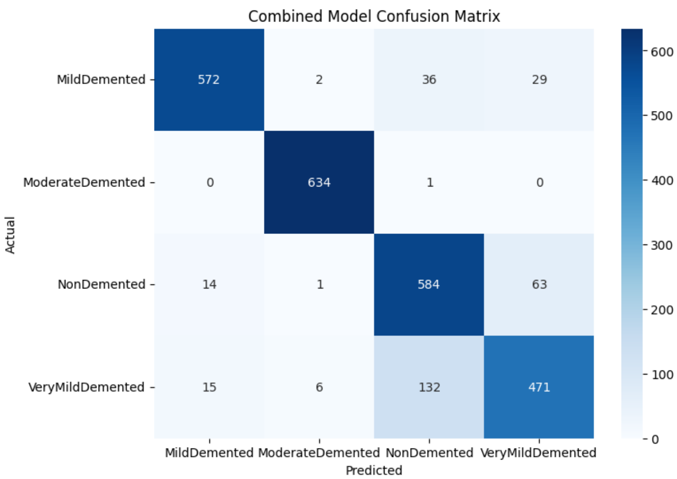

Fig. 10. Confusion matrix of combined model validation.

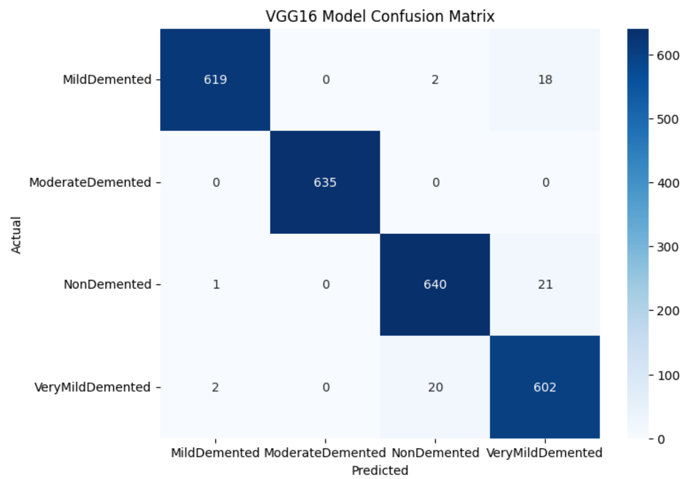

Fig. 12. Confusion matrix of improved model validation.

took only 2 minutes total. The code and data can be found at github. The training and validation accuracy are shown in Figure 11 The confusion matrix of the improved model is

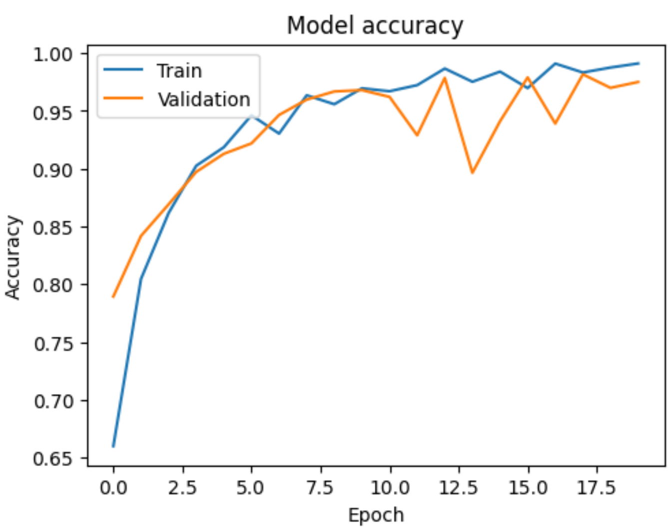

Fig. 11. Training and validation accuracy of improved model.

shown in Figure 12.

VII. Discussion

In this study, we evaluated the performance of the DEMNET model and a combined model consisting of InceptionV3 and DEMNET. Our results indicate that the DEMNET model alone was successful in achieving the desired outcomes. However, the combined model (InceptionV3 + DEMNET) was unsuccessful. This lack of success could be attributed to two potential reasons: either the authors failed to include enough detail to accurately represent the combined model, or they were unjustified in making the claims they did regarding its effectiveness, falling victim to the issues with the combined model that were outlined previously. The improved model, which used the VGG16 model instead of InceptionV3, was able to achieve an accuracy of 0.9750, which is superior to the DEMNET model. This suggests that the InceptionV3 model may not be well-suited for the task of diagnosing Alzheimer's disease from MRI images. The VGG16 model is a simpler architecture that may be more appropriate for this task. The results of the improved model demonstrate the importance of selecting an appropriate model architecture for the task at hand. The DEMNET model alone was able to achieve high accuracy in diagnosing Alzheimer's disease from MRI images, which highlights the potential of deep learning in medical imaging. The combined model, however, was not successful in improving the performance of the DEMNET model. This suggests that the InceptionV3 model may not be well-suited for this task, and further investigation is needed to identify the most appropriate model architecture for diagnosing Alzheimer's disease from MRI images. Future research should focus on providing a more detailed representation of the combined model and critically evaluating the claims made about its performance. This will help in understanding the limitations and potential improvements for the combined approach.

VIII. Conclusion

In this paper, we reviewed the study 'Early diagnosis of Alzheimer's disease from MRI images with deep learning model' [1] and identified several issues with the proposed approach. We attempted to reproduce the results of the study and evaluated the performance of the DEMNET model and a combined InceptionV3/DEMNET model. Our results indicate that the DEMNET model alone was successful in diagnosing Alzheimer's disease from MRI images, achieving an accuracy of 0.9314. However, the combined model was not successful in improving the performance of the DEMNET model, achieving an accuracy of 0.8716. We proposed an improved model that used the VGG16 model instead of InceptionV3 to extract features from the input images. The improved model achieved an accuracy of 0.9750, which is superior to the DEMNET model. This suggests that the InceptionV3 model may not be well-suited for the task of diagnosing Alzheimer's disease from MRI images. The results of the improved model demonstrate the importance of selecting an appropriate model architecture for the task at hand. Future research should focus on identifying the most appropriate model architecture for diagnosing Alzheimer's disease from MRI images and providing a more detailed representation of the combined model to evaluate its performance accurately.

REFERENCES

[1] S. A. Javid and M. M. Feghhi, "Early diagnosis of alzheimer's disease from mri images with deep learning model," *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*, pp. 1–5, Feb. 2024. [Online]. Available: http://dx.doi.org/10.1109/AISP61396.2024.10475240

[2] S. Murugan, C. Venkatesan, M. G. Sumithra, X.-Z. Gao, B. Elakkiya, M. Akila, and S. Manoharan, "Demnet: A deep learning model for early diagnosis of alzheimer diseases and dementia from mr images," *IEEE Access*, vol. 9, pp. 90 319–90 329, 2021.

[3] G. Folego, M. Weiler, R. F. Casseb, R. Pires, and A. Rocha, "Alzheimer's disease detection through whole-brain 3d-cnn mri," *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020. [Online]. Available: https://www.frontiersin.org/journals/bioengineering-and-biotechnology/articles/10.3389/fbioe.2020.534592

[4] M. A. C. Y., "Feature extraction for medical image classification: A novel statistical approach," *International Information and Engineering Technology Association*, vol. 40, no. 2, pp. 727–733, 2023. [Online]. Available: https://iieta.org/journals/ts/paper/10.18280/ts.400232

[5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *CoRR*, vol. abs/1512.00567, 2015. [Online]. Available: http://arxiv.org/abs/1512.00567

[6] I. Khalil, A. Mehmood, H. Kim, and J. Kim, "Octnet: A modified multi-scale attention feature fusion network with inceptionv3 for retinal oct image classification," *Mathematics*, vol. 12, no. 19, 2024. [Online]. Available: https://www.mdpi.com/2227-7390/12/19/3003

[7] H. Mansourifar and W. Shi, "Deep synthetic minority over-sampling technique," 2020. [Online]. Available: https://arxiv.org/abs/2003.09788

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:14124313